Group Housing Allocation: Collusion, Fairness and Efficiency

Justin Burkett, Francis X. Flanagan, and Amanda L. Griffith*

February 15, 2015

Abstract

We model student housing allocation mechanisms at universities as a one-sided matching market in which students have preferences over roommates and rooms. We show that in general there is no best "fair" mechanism, which means there is no mechanism which treats similar students equivalently and always results in a weakly more efficient allocation than all other fair mechanisms. We also show that, due to the trade-off between room quality and roommate quality, certain mechanism allow, and encourage, groups of students to collude. We apply these results to data collected from the housing allocation process at a specific university and show that certain "collusive" groups are able to improve the room allocation of their members, at the cost of non-collusive groups.

I Introduction

In the competitive market for post-secondary education in the U.S., colleges and universities are always interested in how to allocate their limited resources optimally to attract the highest quality students. Although one might think this would mean spending more on enhancing academic quality, recent research suggests that students' decisions about which universities to attend in the US are more sensitive to the non-academic amenities that universities offer, such as student housing. Universities can then compete for students by offering more attractive consumption goods like luxury on-campus housing instead of by improving their academic quality (Jacob et al., 2013; Kirp, 2005;

*Burkett: Department of Economics, Wake Forest University, Box 7505, Winston-Salem, NC 27109; burketje@wfu.edu. Flanagan: Department of Economics, Wake Forest University, Box 7505, Winston-Salem, NC 27109; flanagfx@wfu.edu. Griffith: Department of Economics, Wake Forest University, Box 7505, Winston-Salem, NC 27109; griffial@wfu.edu.

Griffith and Rask, 2013). There is also some evidence that this trend in student preferences is growing, and the issue has attracted the attention of the media.[1]

Although universities can react to this trend by building more attractive housing, it will still be the case that there will be some on-campus housing that is more attractive to students than others for a variety of reasons like age of capital stock or location on campus. If the most attractive housing is scarce on campus, the university must decide how to allocate this housing, and its choice of allocation mechanism may ultimately influence the students' perception of the university. For luxury dormitories to influence the matriculation decisions of high-quality students it must be the case that all (or most) students feel that they have a fair chance at scoring a room in the coveted dorm. If however, students are consistently frustrated in their attempts to live in the luxury dorm, this will likely dissuade future students that would have been drawn to the university by this dorm from enrolling.

We study one university's housing allocation problem through a theoretical analysis and by making use of detailed data from the university. We are interested in this particular allocation problem for a couple of reasons. The first is that it is an instance of a matching problem with features that to our knowledge have not yet been studied in the matching literature. Specifically, students are placed into suites of varying sizes (from 1 to 11 in our case), and hence are matched to other students as well as the suite in which they will live. The roommate problem, in which students are matched to roommates, and the housing allocation problem, in which students are matched to "houses", have been studied extensively, but neither incorporates both elements into the same problem.

Our second reason for interest in this problem is that our empirical results indicate that while almost all individual student's react to the incentives of the mechanism, only some student *groups* do. It is clear from the design of the mechanism that there may be opportunities for larger groups of students to profitably coordinate their actions (i.e., to collude), and we observe this effect in the data. However, it does not seem to be the case that every organization is able to coordinate effectively. This turns out to have important consequences for the ex-post outcomes of the mechanism, even if ex-ante the mechanism is clearly fair.[2] These results pose a challenge for a university who is concerned about the perceived fairness of an allocation mechanism.

[1] For example, see "How to Get College Tuition Under Control" (The Wall Street Journal, October 8, 2013).

[2] We will say that a mechanism is fair if similar students are treated equivalently by the mechanism.

We introduce a tractable model in which students have preferences over other students and where they live. The simplifying assumptions, such as the assumption that all students have the same preferences over houses, are supported by our data from a survey of students' preferences and the actions taken by the students in the mechanism we study. This model provides a framework for thinking about the university's problem, allowing for an analysis of the consequences of the apparent collusive behavior and potential alternative mechanisms.

For example, a simple modification to the current mechanism would eliminate the fairness concern, as it would eliminate the incentive to collude (at least in the current manner). However, it is not clear that this mechanism is better than the existing one, as it does not necessarily perform better on the basis of a natural efficiency criterion in our model. The idea behind the modification is to, in contrast to the existing mechanism, separate the decisions of who to live with and where to live. This illustrates a basic trade-off between efficiency and fairness in this environment.

In the case we study, the university's problem is to allocate housing that accommodates a variety of group sizes with the most common being two, four and six. The basic process used at the university can be thought of as a random serial dictatorship mechanism with priority based on seniority and modified to accommodate groups of students selecting housing.[3] Initially, the university assigns lottery numbers to all students that determine the time at which each student can select housing. Since multiple lottery numbers correspond to the same selection time, we use the term lottery rank to represent the time slot that each lottery number allows a student to pick in, allowing for ties. The lottery ranks are independently distributed within each class year with more senior students always receiving better ranks than their juniors.[4]

Selection of housing occurs in real time at a later date, and each selection made is effectively final with the selected housing immediately removed from the pool of available housing. A group of students can select housing for the group at the earliest time that a member of the group can. For example, if two students who received the best and worst lottery ranks intend to live together they can use the best one and select from any of the initially available housing options. The students are allowed an ample amount

[3]Section VI describes this mechanism in more detail, while Section IV provides a definition based on the model.

[4]The distribution is not uniform across ranks, because the school assigns more students to certain time slots to balance the load on the system.

of time after they receive their lottery ranks to organize into groups before housing is selected. This procedure is described in more detail in Section VI.

This design presents the students with clear trade-offs between whom they live with and where they live. For a simple example, suppose that two students, s_1 and s_2, are close friends but both receive bad lottery ranks. If s_3, a casual acquaintance of s_1, receives a very good lottery rank, s_1 might prefer to live with s_3 in a more desirable location. The students with good lottery ranks are effectively pushed up the preference lists of the other students due to their access to more housing options.

The possibility for collusion arises because organizations of students can manipulate this process to increase the surplus of the organization by arranging its members to make the best use of its available lottery ranks. To see this, suppose that s_1, s_2, s_3, and s_4 are members of an organization. If the close friends s_1 and s_2 receive the two best lottery ranks and live together, they must "waste" a good lottery rank for the organization. By placing s_1 with s_3 and s_2 with s_4, the organization can improve the rooms received by s_3 and s_4, who might have received bad lottery ranks. Whether or not this improves the allocation for the organization depends on how willing each student is to live with the person they are placed with, but in principle the organization could compensate s_2 for any loss in surplus associated with splitting from s_1.

Based on the absence of a divisible good being traded in the mechanism and the participants limited future interaction with the mechanism, one might expect that it is difficult for groups to implement collusive outcomes, and hence that collusion is not a significant concern. But this argument would hinge on restricting groups to using instruments provided by the mechanism to redistribute surplus and implement a collusive outcome. In a setting were groups of students likely interact frequently outside of the mechanism, it seems more likely that some (but potentially not all) groups have the means to compensate those who take sub-optimal actions in order to implement a collusive outcome.

In the data, we observe the lottery ranks received by each participant in this process and the one that was used for the group that they selected housing with. If an organization is effective at rearranging members to make better use of the lottery ranks they receive, we should see that the lottery ranks used by a member in that organization should tend to be better (lower in this case) than those used by students outside of the organization.

In Figure 1, we plot the density estimates for the used lottery ranks. For each of three

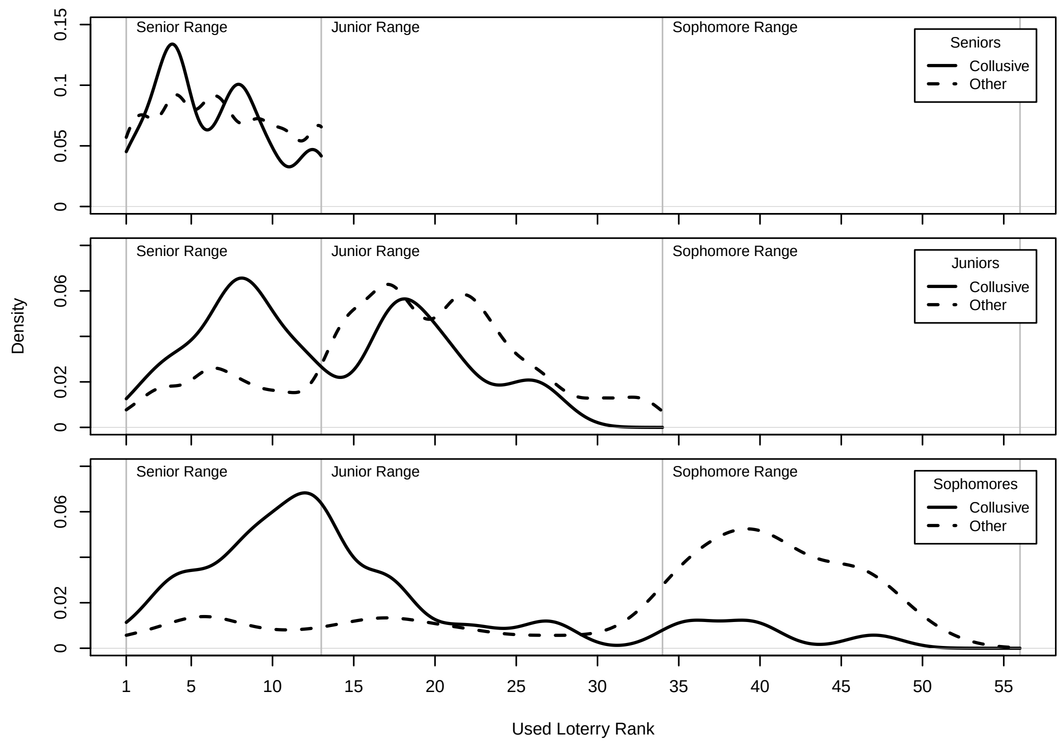

Figure 1: Estimated Densities of Used Lottery Ranks

class years (freshmen are assigned housing according to a different process), we compare the lottery ranks used by those in groups that we identify later in the paper as collusive to all others. Note that each curve integrates to one by construction. The lottery ranks are partitioned into ranges corresponding to the interval of lottery ranks that a member of each class may initially receive.[5] Using the ranges as a reference it is immediately clear that it is relatively common for a sophomore to live with a junior or a senior, especially in the collusive groups.

Statistical tests support the pattern seen in the graph, that junior and sophomore members of the collusive organizations tend to use better lottery ranks than non-members.[6] We also can confirm that there is no significant difference between the lottery

[5]These ranges are disjoint, and students with more seniority based on the number of semesters they have live on campus always receive better lottery ranks than those with strictly less seniority.

[6]Using Kolmogorov-Smirnoff tests at the 1% level, we reject the null hypothesis that the distribution of used lottery ranks is the same across collusive and non-collusive groups for either the juniors or the

numbers received by the members of the collusive and non-collusive groups,[7] so the observed effect is due entirely to how the lottery numbers are used. Although strongly suggestive, this graph does not control for potential differences in preferences, which we do later in Section VIII.

We present our model in Section III, following a discussion of the related literature in Section II. Section IV compares the existing mechanism to alternatives. Section V explores the theoretical implications of colluding student organizations. In Section VI, we describe the data we use, while Section VII presents our initial empirical results. Section VIII gives the evidence that these outcomes represent collusion, and Section IX concludes.

II Related Literature

The theoretical model we introduce in this paper unites two distinct subjects in the matching literature: housing allocation problems and roommate problems. The *house allocation problem*, introduced by Hylland and Zeckhauser (1979), is characterized by a set of n agents who each must be assigned exactly one object (house) from a set of n indivisible objects. This is closely related to the *housing market problem*, introduced by Shapley and Scarf (1974), which is a house allocation problem in which each agent initially owns one object. Abdulkadiroglu and Sönmez (1999) introduce a hybrid model, the *house allocation problem with existing tenants*, in which some, but not all, objects are initially owned. Our model begins as a housing allocation market without existing tenants: initially the objects (rooms) are unassigned. The main difference between previous literature on house allocation and our model is that we wish to assign objects to coalitions, rather than individuals. Dogan et al. (2011) looks at the housing market problem with couples, however the authors assume that coalitions are determined exogenously. In the present paper we are explicitly interested in the coalition formation part of the game, and our theoretical results mostly focus on coalitions of size two, which is why our model is also closely related to the roommates problem.

The *roommates problem*, introduced by Gale and Shapley (1962), is a one-sided matching problem: there are $2n$ agents who must form n coalitions. Gale and Shapley (1962)

sophomores.

[7] This was confirmed for each of the seniors, juniors and sophomores using Kolmogorov-Smirnoff tests at the 5% level.

introduced this as an example of a market in which there is no guaranteed stable match. Thus in order to guarantee a stable roommate match we must make further restrictions on preferences. In our model students exist in a metric space and prefer roommates who are closer to themselves. This restriction was originally proposed by Bartholdi and Trick (1986) as one for which there is always a stable match, and Chung (2000) shows that this is part of a more general category of preferences, called *no odd rings* preferences, which always guarantee a stable match in the roommates problem. Our empirical results are drawn from a market in which students are allowed to form coalitions larger than two, however, as mentioned, we mostly restrict our theoretical results to the situation in which students can only form pairs. One reason for this restriction is because even if preferences are generated by distances in a metric space stability is not guaranteed if the size of the roommate coalition is expanded beyond couples (Arkin et al. (2009)).

Lastly, our paper fits into the large and diverse literature on collusion in games (Cf. Graham and Marshall (1987), McAfee and McMillan (1992), Pesendorfer (2000)). Although we do not explicitly model transferable utility, we sometimes allow certain organizations to dictate their members' behavior, with the interpretation that they have some exogenous means of transferring utility among members. Unlike the previous example, collusion in our model is not explicitly against the rules of the mechanisms we study. Furthermore, we find that collusion has an ambiguous effect on efficiency, although of course collusion harms individuals who, for whatever reason, cannot collude with other students.

III Model

There are N rooms, each with capacity 2. Let S be the finite set of students, $|S| = 2N$, and let (S, d) be a metric space. A roommate *pair* C is a subset of S such that $|C| = 2$. For $\{s, s'\} = C$, let $d(C) = d(s, s')$. Denote the set of all pairs $\mathcal{C}$. A *pair structure* ψ is a partition of S into pairs. For any pair structure ψ, let $C(s)$ denote the pair in ψ which contains student s. Thus $s' \in C(s) \iff C(s) = C(s')$.

A pair structure ψ is *stable* if there does not exist a $C' \in \mathcal{C}$ such that $d(C') < d(C(s_i))$ for each $s_i \in C'$. In this model agents' preferences over roommates will be determined by distance to the roommate, so this is the usual notion of stability in a model in which agents only have preferences over other agents.[8] It is well known that when students' preferences over roommates are determined by distance then a stable pair structure exists.

Theorem 1. *(Bartholdi and Trick (1986), Chung (2000)) A stable pair structure exists.*

Proof. The proof of this theorem is constructive. First, find a pair $C \subset S$ which satisfies $\min_C d(C)$. This pair would never wish to deviate from any pair structure which it is part of. Let this pair be part of ψ and remove it from the set of students. Continue this process until the set of students is fully partitioned; the resulting pair structure is stable. □

This proof also shows that if each student has strict preferences over roommates, meaning $d(s_i, s_j) \neq d(s_i, s_k)$ for any $s_i, s_j, s_k \in S$, then the stable pair structure is unique.

Corollary 1. *If $d(s, s') \neq d(s, s'')$ for any $s, s', s'' \in S$, then there exists a unique stable pair structure.*

Proof. Similar to the construction of ψ in the proof of Theorem 1, find all pairs $C \subset S$ which satisfy $\min_C d(C)$. Since $d(s, s') \neq d(s, s'')$ for any $s, s', s'' \in S$, there is no student who is named in more than one pair in this set. Let all of these pairs be part of ψ and remove them from S. Continue this process. The resulting pair structure is the unique stable pair structure. □

Besides stability of the pair structure, we also want to explore how allowing preferences over rooms affects the stability of the overall allocation of rooms across roommate pairs. It is easy to show that allowing very general preferences over rooms makes this model intractable, in the sense that there is no way to guarantee that a stable allocation exists. Thus, in order to analyze the game, and specifically isolate the affect the chosen mechanism has on room distribution, we assume that each student has the exact same, strict, cardinal preferences over rooms. That is, there exists a common value function $v : R \to \mathbb{R}^+$ such that $v(r_i)$ represents the value any student would receive from room r_i. Therefore we can label the rooms as $r_1, r_2, \ldots, r_n$ such that $v(r_i) > v(r_j) \iff i < j$. The complete utility function for student s for being matched with roommate s' and room r is therefore $u_s(r, s') = v(r) - d(s, s')$. Each student s has a preference relation $\succeq_s$ which agrees with u_s.

[8] We allow for preferences over rooms as well as roommates, but it will be useful to examine the stability characteristics of the pair structure.

Although a strong assumption, our data suggest that this assumption reasonably applies in our empirical analysis. In a survey conducted prior to the students selecting housing, we asked which suite sizes they were considering choosing and then conditional on these responses we asked them to rank all of the available buildings. In the appendix we show that these survey responses are consistent with the hypothesis that most students have very similar preferences over buildings. These reported preferences also closely correspond to the order in which the suites disappear during the selection process.

An *allocation* $A = (\psi, M)$ is a pair structure ψ and a mapping $M : \psi \to R$, where $M(\{s,s'\})$ is the unique room assigned to pair $\{s,s'\} \in \psi$. We will say roommate pair $\{s,s'\}$ *occupies* room $r = M(\{s,s'\})$. For any collection of pairs $\mathcal{C}'$, let $R(\mathcal{C}') = \cup_{c \in \mathcal{C}'} M(C)$ be the set of rooms occupied by pairs in $\mathcal{C}'$ according to M, and let $S(\mathcal{C}') = \{s | \exists C \in \mathcal{C}' \text{ s.t. } s \in C\}$, be the set of students named in any pair in $\mathcal{C}'$. The stability concept we use for allocations is strong-stability. An allocation A is strongly-stable if there is no collection of pairs that can rearrange rooms and roommates such that every student is at least as well off as in A and at least one student is better off. Such a set of pairs is called a *blocking group*.[9]

Definition 1. An allocation is *strongly-stable* if there does not exist a set of pairs and rooms $\{\mathcal{C}', R(\mathcal{C}')\}$, $\mathcal{C}' \subseteq \psi$, such that there exists a set of roommate pairs $\mathcal{C}''$ and a room mapping $M' : \mathcal{C}'' \to R(\mathcal{C}')$ such that:

1. $S(\mathcal{C}') = S(\mathcal{C}'')$,
2. $R(\mathcal{C}') = R(\mathcal{C}'')$, and
3. for each $s \in S(\mathcal{C}')$, for $C' \in \mathcal{C}'$ and $C'' \in \mathcal{C}''$ such that $s \in C'$ and $s \in C''$, we have that $(C'', M'(C'')) \succeq_s (C', M(C'))$, and for at least one student $(C'', M'(C'')) \succ_s (C', M(C'))$.

Theorem 2. *A strongly-stable allocation exists.*

Proof. We will show that any stable pair structure ψ and any room distribution $M(\psi)$ is a strongly-stable allocation. Suppose this were not true. Then there exists a blocking group $\mathcal{C}' \subseteq \psi$. Let r'_n be the worst room in $R(\mathcal{C}')$. After the block this room must be occupied by

[9]We will assume that every student in the blocking group has either a different roommate, room, or both after the block.

a new student pair, $\{s, s'\}$. However, since this is the least valued room by all students, and each student s and s' must have a weakly higher payoff after the block, it must be that $d(s,s') < \min\{d(C(s)), d(C(s'))\}$. Thus $\{s, s'\}$ would block ψ, contradicting that it is a stable pair structure. $\square$

Corollary 2. *Given any stable pair structure ψ and any room mapping $M(\psi)$, $A = (\psi, M(\psi))$ is a strongly-stable allocation.*

Thus stable pair structures will always define stable allocations as well, however the converse is not true, as will be shown later. The next section describes some characteristics and desired outcomes of various mechanisms used to assign students to rooms.

IV Comparison of Mechanisms

A *mechanism* will be any procedure used to assign students, or student pairs, to rooms. We will refer to the mechanism designer as the "housing department," H, of a college or university.[10] We assume that this office values stability, efficiency, and fairness. One allocation is more *efficient* than another if it has higher total welfare, and a mechanism is *fair* if, ex ante, each student has equal probability of receiving any room. In this vein we assume that H attempts to maximize *roommate-welfare*, $W = -\sum_{C \in \psi} d(C)$, subject to ψ being stable.[11] For now assume that there is a unique roommate-welfare maximizing stable pair structure, ψ^*.[12] Therefore the ideal mechanism for H, the *H-mechanism*, partitions students into $\psi = \psi^*$, and assigns each pair $C \in \psi^*$ room r_i with probability $1/N$.

Anecdotal evidence suggests that the most common way colleges distribute rooms is through some sort of lottery. That is, there is some procedure in which students form roommate coalitions, are assigned lottery numbers, and then choose rooms based on the priority structure created by the lottery numbers. However even within this simple framework there can be many variations. For example, some colleges may require roommate coalitions to be formed first and assign each coalition a single lottery number; other colleges may have lottery numbers given to each student first, and then have the

[10] At our university this is the department of Residence Life and Housing.

[11] Since $\sum_{C \in \psi} v(M(C))$ is the same for all pair structures, maximizing roommate-welfare is equivalent to maximizing total welfare.

[12] This would be true if students have strict preferences over roommates, since this would mean there is a unique stable pair structure.

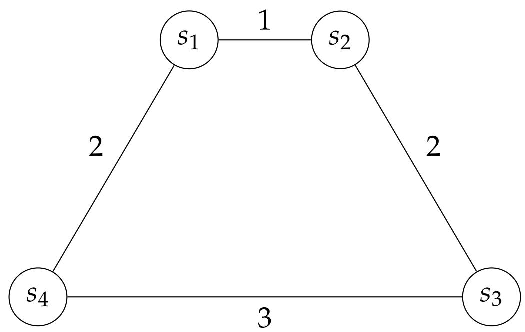

Figure 2: Spatial Distribution of Students

students form coalitions. In the latter case the lottery number of the group may be the lowest number of the group, or an average, etc.

The mechanism in which each student first receives a lottery number, and then students form pairs, we will call the *random serial dictatorship with pair formation mechanism (RSDP-mechanism)*, and is as follows: Let the set of lottery numbers be $L = \{1, \ldots, 2N\}$. Each student is randomly assigned a unique lottery number $l \in L$.[13] After lottery numbers are assigned, students form pairs. For the pair $\{s, s'\}$, the lottery number is the lower of the two lottery numbers assigned to the students in the pair. After pairs are formed rooms are chosen in order dictated by the lottery numbers for each pair. For now we allow cooperation among any subset of students and all students have full information, thus the set of equilibria of a specific RSDP procedure will be a subset of the set of strongly-stable allocations. However, example (REFERENCE) shows that an equilibrium may not exist, and the following example shows that this mechanism may lead to equilibrium allocations in which the pair structure is not stable.

Example 1. Let there be four students in the market and two rooms, with $v(r_1) > v(r_2)$. Figure 2 represents the placement of the students in the metric space, and the numbers on the edges connecting the students are the relevant distances. The unique stable pair structure is $\psi = \{\{s_1, s_2\}, \{s_3, s_4\}\}$, since otherwise $\{s_1, s_2\}$ would block. However, suppose s_3 receives lottery number $l = 2$ and s_4 receives lottery number $l = 1$. Then s_4 is

[13] At this point there is no need to specify the probability of receiving a specific number. Often it is convenient to assume that each student has probability $1/|2N|$ of receiving any number, but sometimes it may be useful to split the students into "classes" and give the different classes different probabilities.

guaranteed r_1, and thus s_1 is willing to be s_4's roommate if:

$$v(r_1) - 2 > v(r_2) - 1 \iff v(r_1) > v(r_2) + 1 \tag{1}$$

If inequality (1) is satisfied, then the equilibrium allocation which results from this realization of the *RSDP*-mechanism would be strongly-stable, even though the pair structure is not stable. Furthermore, this allocation has the same total welfare as a strongly-stable allocation produced from the stable pair structure.[14] Therefore if we changed the distance between two of the students, say s_1 and s_4, it would be possible for this mechanism to result in an equilibrium which is a strongly-stable allocation without a stable pair structure, and has either more or less total welfare than the strongly stable allocation with the stable pair structure.

Thus efficiency and stability of the pair structure are distinct notions. The possible tension between stability and efficiency is a common theme in the matching literature.[15] Here, since students treat rooms and roommates as substitutes, it should be expected that there may be situations in which students are willing to trade roommate quality in favor of room quality. Also, since the choice of one student to deviate from the stable pair structure may impose a negative externality on his roommate in the stable pair structure, it is not possible to say that one mechanism Pareto dominates another. It is easy to extend this logic to show that there is an ambiguous relationship between the expected welfare of the *RSDP*-mechanism and the expected welfare of the *H*-mechanism.

Proposition 1. *There exist markets in which the expected total welfare from the RSDP-mechanism is greater than the expected total welfare from the H-mechanism, and vice versa.*

Proof. Take the market in Example 1. Recall that maximizing total welfare is equivalent to minimizing the sum of the distances between roommates, since all students have the same preferences over rooms. The roommate welfare from the *H*-mechanism is: $-(d(s_1, s_2) + d(s_3, s_4)) = -4$. Assume that $v(r_1) > v(r_2) + 1$, so that if s_3 receives the best lottery number then he will pair with s_2, and if s_4 receives it then he will pair with s_1. If either s_1 or s_2 receives the best lottery number then they will pair with each other. Assume there are $|S| = 4$ distinct lottery numbers and the probability of receiving any

[14]The sum of the distances of the roommates is 4 in each of these strongly-stable allocations.

[15]See for example **?** and **?**.

lottery number is 1/4. Then the expected welfare of the $RSDP$-mechanism is:

$$\frac{1}{2}\left(-(d(s_1,s_2)+d(s_3,s_4))\right)+\frac{1}{2}\left(-(d(s_1,s_4)+d(s_2,s_3))\right)=-4$$

Now suppose the distance between s_1 and s_2 is $1+\varepsilon<2$. Then we get the welfare of the H-allocation is $-4-\varepsilon$, and the expected welfare of RSDP is $-4-\frac{1}{2}\varepsilon$. Similarly, if we decrease the distance between s_1 and s_2 to $1-\varepsilon>v(r_2)+2-v(r_1)$, then the relative welfares of the H-allocation and the $RSDP$-mechanism are $-4+\varepsilon$ and $-4+\frac{1}{2}\varepsilon$. □

Even for individual students the expected utility from the different mechanisms may be ambiguous. For example, in Example 1 we cannot even say if the expected benefit to s_4 is greater in the $RSDP$-mechanism or the H-mechanism: under the $RSDP$-mechanism s_4 can sometimes match with a more preferred roommate, however under $RSDP$-mechanism s_4 receives the best room less often than under the H-mechanism.[16] Since the efficiency effects of each mechanism are ambiguous, our results will focus on stability and fairness. The following section examines how rooms are distributed under the $RSDP$-mechanism when there are varying levels of student organization and collusion.

V Student Organizations

At most universities there are well defined groups which students belong to which may influence their roommate formation strategies. For example, students may be members of an athletic team, have different majors, different graduation years, and, possibly most relevant for housing choices, on many campuses students may be members of a fraternity or sorority. Referred to in this paper as "Greek Letter Organizations," sororities and fraternities are often well organized and can exert extreme influence on their members.[17] What's more, these groups often have some sort of outside option for housing, the so-called fraternity or sorority "house." Due to their high level of organization and ability to move members into housing outside of the centralized campus housing market, there is concern that Greek organizations may be able to place members in more desired rooms, crowding out students who are not members of Greek life. In this section we develop

[16] The expected utility for s_4 is greater under the H-mechanism relative to the $RSDP$-mechanism if $v(r_1)>v(r_2)+2$.

[17] See "Greek Letters at a Price," *The New York Times*, Oct 28, 2014.

a model to see if and how sororities and fraternities can strategically place members in better housing.

Let α be a partition of S. We will use one subset, $\Omega \in \alpha$, as a reference "society" which represents a sorority or fraternity. The remainder of students, $S \setminus \Omega$, may or may not be members of different societies.[18] Let Ω be such that $|\Omega| = 2M$, $2 \leq M < N/2$. The society has a dictator who tells each member how to act under the room assignment mechanism, and whose objective is to maximize the total welfare of the society. In order isolate the effect of the society on the final room distribution, we assume that each member of Ω exists at the same point in the metric space, and that this is the defining characteristic of what it means to be in the society.[19] This means that, for $\omega, \omega' \in \Omega$, $d(\omega, \omega') = 0 < d(\omega, s)$ for any $s \in S \setminus \Omega$. This implies that the members of Ω would be matched to each other in the H-mechanism, and that the payoff function for the society's dictator is determined entirely by the values of the rooms assigned to the society. That is, for any allocation in which each member of Ω is paired with another member of Ω, the payoff of the dictator is $\sum_{\{C \in \psi | C \subset \Omega\}} v(M(C))$. Call this payoff the *room-welfare*, or simply *welfare*, of Ω for allocation A. Likewise the *room-welfare* for any set of student pairs $\mathcal{C}' \subset \psi$ is $\sum_{\{C \in \mathcal{C}'\}} v(M(C))$.

First we will look at two extreme cases for the behavior of students not in Ω: in one case we will assume complete naivety about the room-assignment mechanism, so that students outside of Ω always partition themselves into the stable pair structure, and in the other case we will assume that the students not in Ω in fact have their own dictator whose goal is to maximize the total room welfare of the group. That is, the two cases will be one in which $S \setminus \Omega$ does not take into account at all the value of rooms, and one in which $S \setminus \Omega$ is a society which only takes into account the value of rooms.

Our first result is that, regardless of the behavior of $S \setminus \Omega$, it is indeed more likely that Ω receives the best $k \leq M$ rooms under the $RSDP$-mechanism than under the H-mechanism.

Theorem 3. *The probability that Ω receives the first $k \leq M$ rooms under the RSDP-mechanism is greater than under the H-mechanism, regardless of the strategies of $S \setminus \Omega$.*

[18]We will make explicit assumptions about this set of students as required.

[19]This can be relaxed slightly, since we really need the maximum difference between any two total roommate welfares from Ω to be less than the difference between the value of any two rooms, but this assumption is simpler. This can be motivated by an assumption that all members of a society are similar in some way, or that the society has some exogenous enforcement mechanism.

Thus, perhaps unsurprisingly, the presence of societies can influence the distribution of rooms under the $RSDP$-mechanism. What's also interesting is that this result is independent of the strategies of the students who are not in the society. This result comes from the facts that Ω is allowed to act strategically under the $RSDP$-mechanism and that the set of lottery numbers is different under the two mechanisms, but in each mechanism Ω needs only the lowest k lottery numbers in order to obtain the k best rooms. This also shows that even in a perfectly symmetric setting in which the set of students is split into two equally sized societies, the strategic play of each society does not "cancel out" the other's actions. For example, as a corollary of Theorem 3 it is immediate that if $S \setminus \Omega$ is itself a society, then Ω will not only have a higher probability of receiving the best $k \leq M$ rooms under the $RSDP$-mechanism relative to the H-mechanism, but it will also have a higher probability of receiving the worst $k = M$ rooms.

Corollary 3. *If $S \setminus \Omega$ is a society which has a dictator who attempts to maximize room-welfare for the society, it is more likely that Ω receives the $k = M$ worst rooms under the $RSDP$-mechanism than under the H-mechanism.*

Thus, when the set of students is partitioned into two societies with dictators, each society has a higher probability of getting its most and least preferred set of rooms under the $RSDP$-mechanism relative to the H-mechanism. The increased number of lottery numbers in the $RSDP$-mechanism relative to the H-mechanism, and the ability of the societies to always use their best lottery numbers effectively, causes the distribution of rooms to be more likely at one extreme of the distribution or the other. This is because the order statistics of the lottery number draws are correlated: conditional on receiving the lowest lottery number, the expected second best lottery number held by a society is lower than the expected second best lottery number held by the other society, and so on.

Of course receiving these extreme best and worst sets may be low probability events, so we are also interested in the probability of receiving other sets of rooms. For the following results we will assume that $S \setminus \Omega$ is a single society.[20] Let R' and R'' be two sets of rooms of size M, each containing the same worst room, r_j, with $j < N$. That is, $|R'| = |R''| = M$, and $r_j = \max_i\{r_i \in R'\} = \max_i\{r_i \in R''\}$. Then the probability that the students in Ω occupy the set of rooms in R' under the $RSDP$-mechanism is equivalent to the probability that they occupy the set of rooms R''.

[20] As we will explain later, these results can be thought of a s a"worst case scenario" for Ω.

Theorem 4. *Under the RSDP-mechanism, if S is partitioned into two societies, then for any two sets of rooms* R' *and* R'' *such that* $|R'| = |R''| = M$, *and* $r_j = \max_i\{r_i \in R'\} = \max_i\{r_i \in R''\}$, $j < N$, *the probability that the students in* Ω *occupy the set of rooms in* R' *is equivalent to the probability that they occupy the set of rooms* R''.

Corollary 4. *Under the RSDP-mechanism, if S is partitioned into two societies, then for any two sets of rooms* R' *and* R'' *such that* $|R'| = |R''| = M$, *and* $r_j = \max_i\{r_i \in R'\}$, $r_k = \max_i\{r_i \in R''\}$, $j < k$, *the probability that the students in* Ω *occupy the set of rooms in* R' *is greater than the probability that they occupy the set of rooms* R''.

In some sense these results can be viewed as the "worst case scenario" for Ω. It should be expected that competing against a single, well organized society is worse for Ω than competing against smaller and less organized groups. The following results formalize this intuition, but first we must define what we mean when we refer to smaller organizations. We assume each student can be a member in at most one society, thus the set of societies, and students not in any society, can be described as a partition of S. For set A, a partition α is *finer* than partition β, and equivalently β is *coarser* than α, if every element of α is a subset of some element in β.

Theorem 5. *Let* α *and* β *be distinct partitions of* $S \setminus \Omega$ *into societies. If* α *is finer than* β, *then under the RSDP-mechanism the expected room-welfare of* Ω *is strictly higher when competing against partition* α *relative to partition* β.

An immediate corollary of this result is that societies can improve their combined room-welfare by colluding with each other.

Corollary 5. *Let* α *and* β *be partitions of S such that there exist societies* $\Omega, \Omega' \in \alpha$ *and* $\Omega'' \in \beta$ *such that* $\Omega \cup \Omega' = \Omega'' \neq S$, *and* $\alpha \setminus \{\Omega, \Omega'\} \cup \{\Omega''\} = \beta$. *Then under the RSDP-mechanism the sum of the expected room-welfares of* Ω *and* Ω' *in* α *is strictly lower than the expected room-welfare of* Ω'' *in partition* β.

Therefore larger societies have a larger impact on the room distribution than do smaller societies. Similarly, if a society is less organized, less concerned about room-welfare, or more naive about the workings of the mechanism, then its expected room welfare will be lower. This is an immediate consequence of the fact that the total room-welfare is the same under any allocation, so any set of students that does not attempt to improve room-welfare will necessarily reduce its own room-welfare and increase other groups' room-welfare relative to a situation in which it tries to maximize room-welfare.

VI Description of Data and Details of Allocation Procedure

Empirically, we study the housing allocation mechanism at a medium-sized private university with undergraduate enrollment of roughly 1,000 students per year. On-campus university housing is available for undergraduate students in any year of study. The incoming first-year students (we use the year of student to denote their upcoming year when they choose housing) are assigned housing using a process that is distinct from the other students, and they are not included in any of the analyses in this paper. Students are required to live on campus for three years, but many choose to live on campus for all four. The second-, third- and fourth-year students (there are no fifth-year or older students in our data) who live on campus select housing from the centralized process that is the focus of this study. Housing is available for every group size from 1 to 11 with the most common sizes being two, four and six. We refer to all of these as suites in the remainder of the paper, even though some may be more properly called rooms.

About three weeks before students choose where they will live each student is assigned a lottery number. These lottery numbers correspond to the date and time at which they may log into the system to select housing. They are informed of their login time, but not the lottery number which has no extra significance. Between 20 and 100 lottery numbers are associated with each selection time, and the login times are spread over five days. The lottery numbers are partitioned into ranges corresponding to the number of semesters a student has lived on campus and are uniformly distributed within these ranges. For example, two students with four semesters of credit have the same probability of receiving any of the lottery numbers in the range allocated to four-semester students. All of the lottery numbers for n-semester students are better than any lottery number received by an m-semester student if $n > m$. The corresponding login times for the n-semester students are no later than the login times for the m-semester students.[21] Because two different login times might accommodate different numbers of students, the login times are not uniformly distributed within the semester groups, although it is still true that two students with the same number of semesters have the same chances of receiving each of the corresponding login times.

The data we have correspond to the process run in the spring of 2014 for housing in

[21] The five-semester students share the latest login time of the six-semester students, the three-semester students share latest login time of the four-semester students, and the one-semester students share the latest login time of the two-semester students. These are the only cases where students with differing numbers of semesters on campus share times.

the 2014–2015 academic year. The dataset includes every student that selected housing in that year. In addition to the lottery number of each student and their corresponding login time, we observe the following demographic information for each student: sex, number of semesters, ethnicity and race, indicators for athlete and foreign students, grade point average (GPA), and membership in a "Greek Letter Organization".

We also construct two new variables. First, we rank the selection times from earliest (1) to latest (57) and assign each lottery number in the data to the rank of the corresponding selection time. We refer to this number as the *lottery rank* in our analysis. Any two students will have different lottery numbers, but may share the same lottery rank, meaning that they can log in at the same time. We also observe the precise time at which housing was selected for a group of students. We count the number of students who pick ahead of each student to construct *pick order*. This variable takes the same value for all members of the same suite. The students have no control over their own lottery rank but may influence their pick order if they do not select housing at their earliest opportunity.

Upon logging into the selection system, a student can select any of the available suites. The selection occurs in real-time and is effectively final,[22] so anyone who has the ability to log in but has not yet made a selection can remove an available suite from the pool. Groups select suites by giving permission to one person in the group to register all of the group members for a suite. This means that the group can choose a suite as soon as any one of its members can, which then implies that the only lottery rank that impacts the room selection is the lowest (best) one. Groups do not reveal to the university that they want to live together until a suite is selected for them, so while in principle groups may change up until the time they select housing, we only observe the final groupings into suites.

There are three other ways that upper-class students may select a room on-campus. The first is theme housing; students may form a group corresponding in size to one of the theme houses available as part of on-campus housing and propose a theme for their living area. If chosen, these students do not enter the lottery as described above. Second, a number of the "Greek Letter Organization" on campus have set blocks of rooms within on-campus housing that they may utilize for their members. These organizations are required to indicate which of their members will live in these blocks prior to the release of the lottery times and numbers, and therefore these students are also not in-

[22]We were told that there are rare exceptions to this rule, but we have no way of measuring how rare they are in the data we have.

volved in the lottery process. Lastly, there are a small number of students that receive accommodations in the selection process for certain health and physical disability issues. These students do pick their housing during the lottery process but are given priority, and some restrictions on what housing is available. As none of these three groups picks housing in the typical way just described, we omit these students from our analysis.

To obtain further information on the preferences of the students that would be involved in the lottery, the authors conducted a survey prior to the beginning of the selection process in the spring of 2014. The survey was released to all students eligible for the lottery at a point when they were aware of their possible selection times, but had not yet selected a suite. Students were asked to report what size suites they were interested in, and to rank the buildings with available suites of that size. The survey respondents were also asked to report how many hours they had spent preparing for the housing lottery, as well as their membership in on-campus activities such as clubs and other organizations and how much time they spend weekly with these organizations. 442 students responded and these data were matched to the housing data provided by the university.

VII Housing Selection Outcomes

In the data we observe membership in "Greek Letter Organizations" (GLOs) and participation on an athletic team for every student in the housing selection process. Because GLOs play a large role in the daily lives of their members, they are likely to be the organizations most capable of enticing their members to collude.

Figure 1 in the introduction shows that members of certain GLOs were able to use better lottery numbers than non-members from the same year.[23] The groups used in this figure are a selection of sororities and fraternities that perform much better in terms of the average used lottery numbers than both other GLOs and students who are not members of any GLO.

Table 1 shows the results of six OLS regressions, three for the women and three for the men. Each regression includes dummy variables for each of semesters one through five with the omitted group being six-semester students. The odd number semester coefficients are not shown to save space and because they represent a small fraction of the total students.[24] We do not include dummy variables for organizations with less

[23] About half of the students in our data are members of some GLO.
[24] Less than 1% of our observations in each case.

than five members, and the reported standard errors for (W2), (W3), (M2) and (M3) are clustered at the suite level to account for the fact that the dependent variable is constant across the observations within a suite.

The lottery rank regressions, (W1) and (M1), treat the lottery rank received as the dependent variable. These regressions serve as a test that the lottery ranks were distributed to individual students independent of their affiliation with an organization. In (W2) and (M2), we regress the lottery rank that each individual used in the housing selection process on the same dummy variables. For two sororities (S01 and S02) and three fraternities (F01, F02 and F03), the average lottery rank used is significantly better than the average non-sorority woman or non-fraternity man after controlling for the number of semesters.

The lottery rank differences indicate improvements in the login time of the students, but the number of students are not evenly distributed across login times so the importance of these numbers can be difficult to interpret. The last set of regressions uses Pick Order as a dependent variable, which measures the number of students that selected before a student made their choice with all students in the same suite receiving the same value of Pick Order. These regressions, (W3) and (M3), show the average improvement for each of these organizations. On average members of S02, for example, picked about 11 time slots ahead of non-sorority women, which translates to being ahead of 170 students in the selection process.[25] Comparing this figure to the coefficients on the two-semester and four-semester dummy variables, this benefit exceeds the benefit of being in the next class year (a two-semester increase).

The sororities and fraternities are arranged by listing the ones we are tentatively identifying as collusive (marked with a (c)) first and the non-collusive ones second. Organizations with less than five members are included in the regression as non-GLO students. S01 and S02 represent about 30% of the sorority members involved in the process, while the fraction of fraternity members represented by F01, F02 and F03 is about 32%.

These results show that members of several organizations are able to pick housing significantly earlier than the other students. Given the design of the mechanism this increases the number of choices they have and hence (weakly) improves their choice set relative to other organizations. It should therefore yield more desirable choices of

[25]Pick Order may reflect delays or mistakes made by the students in slecting rooms. A student may have logged in at their assigned time but failed to pick a room right away, falling in the Pick Order.

	(W1) Lottery Rank	(W2) Used Rank	(W3) Pick Order
S1(c)	0.261	-8.256***	-123.9***
	[0.821]	[1.879]	[27.62]
S2(c)	-0.0867	-10.90***	-170.0***
	[1.083]	[2.452]	[44.07]
S3	0.643	-2.502	-20.33
	[0.742]	[1.615]	[28.81]
S4	-1.488	0.237	-27.59
	[1.001]	[1.949]	[44.96]
S5	0.298	-0.180	7.016
	[1.420]	[1.862]	[52.42]
S6	1.286	-3.016*	-45.00
	[1.308]	[1.625]	[37.78]
S7	0.905	-1.870	-19.00
	[1.258]	[1.638]	[32.72]
S8	1.170	1.662	63.67*
	[1.858]	[2.442]	[35.97]
S9	-1.030	1.509	58.70
	[2.606]	[2.653]	[53.91]
S10	-1.090	-1.317	55.66***
	[5.804]	[0.876]	[17.36]
2 Sem	37.81***	22.90***	295.0***
	[0.610]	[1.716]	[26.20]
4 Sem	16.93***	8.455***	133.8***
	[0.615]	[0.961]	[18.74]
Const	9.156***	9.862***	285.6***
	[0.541]	[0.884]	[15.69]
N	657	657	657
R^2	0.874	0.507	0.376

Standard errors in brackets
6 Sem is omitted. 1, 3 and 5 Sem not shown
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

(a) Women

	(M1) Lottery Rank	(M2) Used Rank	(M3) Pick Order
F01(c)	1.912	-5.826**	-67.50
	[1.209]	[2.483]	[50.43]
F02(c)	0.0158	-9.375***	-145.5***
	[1.240]	[2.197]	[27.05]
F03(c)	-3.086*	-7.667***	-92.53**
	[1.654]	[1.358]	[35.95]
F04	0.941	-0.571	12.63
	[0.977]	[5.576]	[122.5]
F05	0.952	3.686**	65.10**
	[1.090]	[1.744]	[30.64]
F06	1.601	-2.556	-25.79
	[1.129]	[2.303]	[41.36]
F07	0.270	5.977***	127.3***
	[1.578]	[1.256]	[18.65]
F08	0.403	1.214	34.62
	[1.531]	[2.289]	[42.63]
F09	2.230	2.618	48.43
	[1.664]	[6.299]	[67.90]
F10	2.387	3.515**	98.90***
	[1.852]	[1.711]	[27.42]
F11	0.736	4.297***	147.0***
	[2.136]	[0.824]	[16.92]
2 Sem	37.62***	24.62***	360.2***
	[0.553]	[1.497]	[24.10]
4 Sem	16.21***	9.740***	184.5***
	[0.560]	[0.900]	[18.62]
Const	8.554***	6.963***	181.5***
	[0.456]	[0.634]	[13.00]
N	706	706	706
R^2	0.886	0.521	0.445

(b) Men

Table 1: Outcomes

	Collusive	Non-Coll.	Non-Coll. GLO	Non-GLO
Sophomore	0.288	0.405***	0.272	0.454***
Junior	0.353	0.389	0.498***	0.349
Senior	0.360	0.203***	0.229***	0.194***
Hispanic	0.014	0.042	0.040	0.042
Black	0.000	0.079***	0.046**	0.091***
Asian	0.050	0.102*	0.061	0.117**
Foreign	0.007	0.062***	0.012	0.080***
Athlete	0.007	0.110***	0.015	0.145***
GPA	3.402	3.213***	3.225***	3.209***
N	139	1,224	327	897

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2: Descriptive Statistics

housing. Note also that since the coefficients measure effects relative to the omitted group it is possible that some collusion is occurring among other subsets of students, so that the identified organizations are only "better colluders".

Table 2 compares the means of the demographic variables we have in the data across four groups. The stars reflect the results of t-tests when these means are compared to the ones in the collusive column. What stands out as a characteristic of the collusive groups is that they tend to have more seniors who select suites through the lottery. This does not mean that there are more seniors in these GLOs because all seniors have the option of living off-campus. There are at least two potential explanations for this. It could be that seniors in these groups simply have a stronger preference for on-campus housing, but it also might be that these groups are convincing the seniors to stay on campus so that the group can take advantage of the better lottery numbers. In the analysis that follows, we attempt to control for the first possibility to show that our evidence favors the second explanation, which would indicate collusive behavior.

The distribution of students across class years is clearly important for our current argument, but the above table also shows other demographic differences between the collusive groups and others. Perhaps the most interesting of these is GPA (grade point average measured on a four-point scale), which one might interpret as picking up more intellectual sophistication. The explanation is tempting, but our later results do not suggest a role for GPA or the other demographic characteristics as explanatory variables.

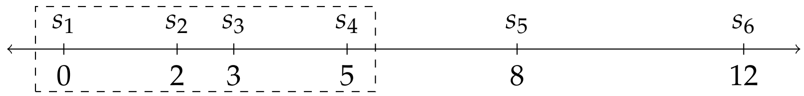

Figure 3: Roommate Preferences Example

VIII Is this collusion?

We next explore alternatives to labeling the identified organizations as collusive. Section V considered the consequences for the allocation, when collusive organizations each arrange their members into k suites to use the best k lottery ranks that the received. This would allow an organization to pick k suites at the earliest possible time. We do not see this extreme level of organization with respect to lottery ranks in the data, and as the last section showed there seems to be quite a bit of heterogeneity between organizations in terms of there ability to arrange members in a way that allows for early room selection.

The model guaranteed that organizations did not need to worry about differences in their members' preferences over living with other members, since it was assumed that these members were close enough in terms of preference distance to not be negatively affected by the rearrangement. We do not expect this to be the case in the data. Based on observable characteristics the members within each organization are similar, but it seems unreasonable to expect that they are equally as happy to live with any other member of an organization. To explain the trade-offs an organization might have to make, consider a simple example of the model in which each student is place on the real number line as shown in Figure 3.

Suppose that the students in the dashed box are members of an organization, that s_2 and s_3 receive the best two lottery numbers while s_1 and s_4 receive the worst two, and that all students agree on how they rank the available rooms. In this case, s_2 and s_3 should live together in a stable matching because they each must place the other at the top of their preference list, but this leaves s_1 and s_4 with potentially poor choices for rooms and roommates. It could be for example that s_4 and s_5 prefer to live with each other over anyone else, leaving s_1 with s_6. By convincing s_1 and s_2 to live together and s_3 and s_4 to live together the organization improves its total payoff due to both a better choice of rooms (they get the first two choices) and a better arrangement of roommates (the total distance between roommates shrinks to 4). However, this requires that the organization is capable of convincing s_1 and s_2 to do this.

Our main concern in testing for collusive behavior is whether or not we are able to effectively control for unobserved differences in students' preferences over roommates.

VIII.1 Test 1: Organization by class year

This test estimates the likelihood that a sophomore or freshman lives with a more senior student based on whether or not they are a member of a collusive group. We use a logit specification in two regressions, in which we exclude all students with six semesters in the data. In the first regression, the dependent variable is a dummy representing the outcome that a senior lives in the same suite, while in the second, the dependent variable represents the outcome that a student with more seniority (an upperclassmen) lives in the same suite. We control for observable characteristics, including sex, ethnicity, grade point average (GPA), participation in athletics and the number of semesters.

Given the differences in the distribution across class years shown in Table 2, we control for the distribution across class years of the student's GLO. For students who are not members of a GLO, we treat their "organization" as being all students of the same sex outside of a GLO. If the collusive groups are keeping seniors on campus expressly to get better rooms, this would be clear evidence of collusion. In this test we start from the alternative hypothesis that the proportion of seniors living on campus reflects preferences and not strategic behavior. If this were the case, then we might see that the likelihood of living with a senior (or upperclassmen) increases with the fraction of seniors on campus, but does not depend on whether the organization is labeled collusive.

Table 3 shows the results of these two logit regressions in terms of the marginal effects associated with each variable computed using the mean observation in the data. By including a Collusive dummy and Non-Coll GLO dummy, we test for a significant difference in the probability of living with a senior or upperclassmen for either members of collusive organizations or members of non-collusive organizations, when compared to non-GLO members. The estimates show a strong tendency for younger members of collusive organizations to live with seniors and upperclassmen, even after controlling for the distribution of members across class years.

This is perhaps the simplest strategy that a collusive organization might use to receive better rooms, since the seniors automatically receive the best lottery ranks and by placing younger members with these seniors all of the members can benefit. This empirical cannot rule out the possibility that seniors in these organizations simply like living with

	(1) Use Senior	(2) Use Upperclass.
Collusive	0.373***	0.436***
	[0.0691]	[0.0576]
Non-Coll GLO	0.00774	0.0847*
	[0.0397]	[0.0461]
Female	-0.00414	-0.0619*
	[0.0292]	[0.0343]
Athlete	-0.0255	0.109**
	[0.0417]	[0.0523]
GPA	-0.0154	-0.0284
	[0.0288]	[0.0337]
2 Sem	-0.114***	0.0949***
	[0.0289]	[0.0343]
3 Sem	-0.153**	0.475***
	[0.0618]	[0.131]
Pct Jun in Org	-0.135	-0.273**
	[0.118]	[0.136]
Pct Sen in Org	0.239	0.378*
	[0.158]	[0.195]
N	1,018	1,018

Not shown: Black, Asian, Hispanic and Foreign.
4 Sem is omitted. Standard errors are in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3: Marginal Effects on Living with a More Senior Student

younger members irrespective of the consequences for room selection, but even with more data it seems impossible to completely rule out such a possibility and the net result in either case is an unequal distribution of rooms for these organizations.

VIII.2 Test 2: Optimal Use of Lottery Ranks

The idea behind the next test is to look at the suites that each member of an organization ended up in and ask if it was possible to rearrange the members between those suites so as to provide those suites with better choices of rooms. If the organizations have complete freedom to arrange their members to maximize the choices members have over rooms, they should never "waste" a lottery rank by allowing one or more of their members to select at a time that is later than available but unused time within the organization.

To explain the test, consider the members of one organization who are provisionally have places in suites. Suppose that each member is assigned a room within their respective suite, and that we take the members and permute them among those rooms, recalculating each suites minimum lottery rank. If we cannot permute the members and obtain a better average lottery rank for all the members, then we call that assignment of members to those suites optimal.[26]

In Table 4, we report the optimal permutation of members for each organization starting from the final assignment from the data. We exclude organizations whose members were only in one suite, since there is no way a permutation of members would affect their used ranks. We report the results for suite sizes four, six and eight because these are the most competitive of the suite sizes on campus. The results for all suite sizes (excluding singles) are reported in the Appendix (Table 5). The fourth and last columns report the fraction of random permutations that were no better than the one used.[27] If the used permutation is optimal, this fraction must be one. The final three columns compute the same statistics after excluding suites that were chosen using sophomore numbers. This was done in case the organizations paid less attention to these choices, which invariably end up in the least desirable housing. If an organization has more seniors (and hence more good lottery ranks to choose from), this is reflected in the low optimal average lottery ranks and is likely to be reflected in the Used column as well (depending on how

[26]This ignores possibilities such as the ability for members to form new suites or remove non-members from the suites in the provisional assignment.

[27]Since the total number of unique permutations grows rapidly, we chose to draw 50,000 random permuations when the total number of permuations was greater than 8!.

	Suite Sizes 4, 6 & 8			Excluding Soph. Selections		
GLO	Optimal	Used	Pr(Used $\leq P$)[a]	Optimal	Used	Pr(Used $\leq P$)[a]
S01(c)	4.55	7.35	0.46	4.80	7.35	0.46
S02(c)	6.32	7.88	0.83	6.32	7.88	0.83
S03	7.49	13.90	0.09	7.59	13.90	0.08
S04	4.89	7.89	0.29	4.89	7.89	0.30
S05	17.17	17.17	1.00	17.17	17.17	1.00
S06	8.00	13.60	0.07	8.00	13.60	0.07
S07	5.10	13.00	0.02	5.10	13.00	0.02
S08	14.67	19.00	0.33	11.00	11.00	1.00
S09	4.75	26.75	0.25			
F01(c)	6.89	9.67	0.22	6.89	9.67	0.22
F02(c)	4.00	4.00	1.00	4.00	4.00	1.00
F03(c)	11.89	13.00	0.45	11.89	13.00	0.44
F05	15.09	20.18	0.22	13.56	16.22	0.56
F06	9.45	10.35	0.95	9.45	10.35	0.95
F08	14.75	14.75	1.00	14.75	14.75	1.00
F10	19.14	19.86	0.71	19.14	19.86	0.71

[a] P is a random permutation of the organization's members.

Table 4: Optimal Use of Lottery Ranks

the seniors were arranged). In most cases members are not arranged to achieve the optimal pick time, and have room to improve their average lottery ranks. The exceptions are S05, F02, and F08. It is worth noting that there are likely decreasing returns to reducing the average used lottery rank as many of the suites within the same building are likely perceived to be identical.

The conclusion we draw from this test is that the low average lottery ranks for the collusive organizations are primarily due not to optimally arranging members across suites but to the availability of better lottery ranks among members on campus. In other words, it seems the organizations gain more of an advantage through keeping seniors on campus than they do by the way they arrange the on-campus members across suites. Some of the non-collusive groups, like S07, choose arrangements that are particularly bad given their available numbers, while others, like S05, are optimally using their ranks but have bad ones from which to choose. Our results on the arrangement of members are therefore mostly inconclusive.

IX Conclusion

In a model of student housing allocation mechanisms in which students have preferences over roommates and rooms, we show that in general there is no best "fair" mechanism, which means there is no mechanism which treats similar students equivalently and always results in a weakly more efficient allocation than all other fair mechanisms. We also show that, due to the trade-off between room quality and roommate quality, certain mechanism allow, and encourage, groups of students to collude.

In outcome data, we identify student organizations that select rooms significantly ahead of all other students in the mechanism. Our tests of collusive behavior suggest that this is the result of these organizations keeping more older students on campus in order to group them with younger students who (by the rules of the mechanism) must receive worse lottery numbers. Our test shows that underclassmen in collusive organizations are approximately 40% more likely to live with an older member of the organization after controlling for observable characteristics and the distribution of students across class year. The result of this collusive behavior is that the average member of a collusive organization chooses suites ahead of a full class year's number of students. In other words, the behavior has significant implications for the ex post fairness of this mechanism.

References

Abdulkadiroglu, A. and Sönmez, T. (1999). House allocation with existing tenants. *Journal of Economic Theory*, 88(2):233 – 260.

Arkin, E. M., Bae, S. W., Efrat, A., Okamoto, K., Mitchell, J. S., and Polishchuk, V. (2009). Geometric stable roommates. *Information Processing Letters*, 109(4):219 – 224.

Bartholdi, J. and Trick, M. A. (1986). Stable matching with preferences derived from a psychological model. *Operations Research Letters*, 5(4):165 – 169.

Chung, K.-S. (2000). On the existence of stable roommate matchings. *Games and Economic Behavior*, 33(2):206 – 230.

Dogan, O., Laffond, G., and Laine, J. (2011). The core of shapley-scarf markets with couples. *Journal of Mathematical Economics*, 47(1):60 – 67.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):pp. 9–15.

Graham, D. A. and Marshall, R. C. (1987). Collusive bidder behavior at single-object second-price and english auctions. *Journal of Political Economy*, 95(6):1217 – 1239.

Griffith, A. and Rask, K. (2013). The cost-quality tradeoff in the matriculation decisions of high-ability students. Working Paper.

Hylland, A. and Zeckhauser, R. (1979). The efficient allocation of individuals to positions. 87(2):293 – 314.

Jacob, B., McCall, B., and Stange, K. (2013). College as country club: Do colleges cater to students' preferences for consumption? NBER Working Paper No. 18745.

Kirp, D. L. (2005). This little student goes to market. In Hersch, R. H. and Merrow, J., editors, *Declining by degrees of mediocritiy: higher education at risk*. Palgrave Macmillan, New York.

McAfee, R. P. and McMillan, J. (1992). Bidding rings. *American Economic Review*, 82(3):579 – 599.

Pesendorfer, M. (2000). A study of collusion in first-price auctions. *Review of Economic Studies*, 67(3):381 – 411.

Shapley, L. and Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23 – 37.

X Appendix

Proof. Theorem 3: The probability that Ω receives the first k rooms under either mechanism is the probability that Ω receives the first k lottery numbers, regardless of the strategies of $S \setminus \Omega$.[28] These probabilities are different because under the H-mechanism

[28]We assume that regardless of the strategies used by $S \setminus \Omega$, once pairs are formed the lower lottery number held by the pair is always the lottery number used.

there are fewer lottery numbers, since the numbers are assigned to pairs rather than individual students. The relevant probability under the H-mechanism is:

$$\text{Probability } \Omega \text{ receives best } k \text{ rooms under the } H\text{-mechanism} = \prod_{i=0}^{k-1} \frac{M-i}{N-i}$$

And the probability under $RSDP$ is:

$$\text{Probability } \Omega \text{ receives best } k \text{ rooms under the } RSDP\text{-mechanism} = \prod_{i=0}^{k-1} \frac{2M-i}{2N-i}$$

It is easy to show that the second probability is larger than the first as long as $N > M$. □

Proof. Corollary 3: By Theorem 3, the probability that $S \setminus \Omega$ receives the best $k' = N - M$ rooms under the $RSDP$-mechanism is higher than the probability it receives these rooms under the H-mechanism. □

Proof. Theorem 4: Theorem 3 shows that the probability Ω receives the best M rooms under the $RSDP$-mechanism is $\prod_{i=0}^{M-1} \frac{2M-i}{2N-i}$. In general, the probability that Ω receives a specific set of exactly M rooms in the range $r_1, \ldots, r_j$, including r_j, for $M < j < N$, is the probability that Ω receives a specific set of exactly M lottery numbers in that range, and $S \setminus \Omega$ receives the remaining numbers. This is given by:

$$\Pr(R_\Omega = R'\} = \Pr(R_\Omega = R''\} = \frac{\prod_{i=0}^{j-M-1} (2(N-M)-i) \prod_{i=0}^{M-1} (2M+i)}{\prod_{i=0}^{j-1} (2N-i)} \tag{2}$$

□

Proof. Corollary 4: Let $R_{\omega_{j+1}}$ be any set of M rooms such that the worst room in $R_{\omega_{j+1}}$is r_{j+1}. Then the probability that Ω occupies $R_{\Omega_{j+1}}$ under RSDP is given by:

$$\begin{aligned}\Pr(R_\Omega = R_{\Omega j+1}) &= \frac{\prod_{i=0}^{j-M} (2(N-M)-i) \prod_{i=0}^{M-1} (2M+i)}{\prod_{i=0}^{j} (2N-i)} \\ &= \Pr(R_\Omega = R') \underbrace{\left(\frac{2N-M-j}{2N-j}\right)}_{<1}\end{aligned}$$

Thus as k increases, as long as $k < N$, the probability that Ω receives any specific set of rooms R'', with r_k the worst room in the set, decreases. □

Proof. Theorem 5: The proof will be by induction on the number of societies Ω competes against. Let β be the partition of $S \setminus \Omega$ into a single society. The room assignment mechanism therefore represents a two player, constant sum game. Thus it is sufficient to show that when $S \setminus \Omega$ is partitioned into two societies, call this partition α, there is at least one distribution of lottery numbers in which Ω receives more room-welfare when playing against α than when playing against β. Let Ω' and Ω'' be the two societies in α, and without loss of generality assume $|\Omega'| \leq |\Omega''|$. Let $\mathcal{L}$ be a permutation of lottery numbers in which all of the students in Ω' received the first Ω' lottery numbers, and the students in Ω received the next M lottery numbers. Playing against β, $S \setminus \Omega$ would receive the first $|\Omega'|$ rooms, and Ω would receive the following M rooms. Playing against α, Ω' would receive the first $|\Omega'|/2$ rooms, and Ω would receive the following M rooms. The same logic applies to the inductive step.

□

	Suites Size 2–11			Excluding Soph. Selections		
GLO	Optimal	Used	$\Pr(\text{Used} \leq P)^a$	Optimal	Used	$\Pr(\text{Used} \leq P)^a$
S01(c)	6.2	9.5	0.28	5.5	7.8	0.56
S02(c)	7.0	11.4	0.50	6.4	9.7	0.54
S03	9.7	17.3	0.00	7.8	14.3	0.00
S04	10.1	24.7	0.00	6.3	12.9	0.03
S05	12.9	16.9	0.37	12.9	16.9	0.37
S06	9.0	13.9	0.13	8.5	12.3	0.19
S07	6.7	10.7	0.10	6.6	10.7	0.10
S08	18.9	26.9	0.01	13.8	13.8	1.00
S09	4.8	26.8	0.25			
F01(c)	6.9	9.7	0.22	6.9	9.7	0.22
F02(c)	5.1	8.5	0.25	5.1	8.5	0.25
F03(c)	11.9	13.0	0.45	11.9	13.0	0.45
F04	12.6	20.8	0.03	10.1	11.2	0.84
F05	17.3	23.1	0.01	15.9	19.6	0.19
F06	10.2	12.3	0.84	10.2	12.3	0.84
F07	22.1	28.1	0.02	21.4	21.4	1.00
F08	15.8	16.8	0.78	15.8	16.8	0.78
F09	30.7	32.2	0.72			
F10	19.1	19.9	0.71	19.1	19.9	0.71

[a] P is a random permutation of the organization's members.

Table 5: Optimal Use of Lottery Ranks (All Suites)