

# The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard\*

Eric Budish<sup>†</sup> and Estelle Cantillon<sup>‡</sup>

November 8, 2009

## Abstract

This paper uses data consisting of agents' strategically reported preferences as well as their underlying true preferences to study strategic behavior in the course allocation mechanism used at Harvard Business School. We show that the mechanism is manipulable in theory, manipulated by students in practice, and that these manipulations cause meaningful welfare losses. However, we also find that ex-ante welfare is higher than under the random serial dictatorship (RSD), which is the only known mechanism that is anonymous, strategyproof and ex-post efficient. This discrepancy between ex-ante and ex-post performance of RSD is specific to the multi-unit assignment problem and can be traced to the callous behavior induced by RSD. We draw lessons for the design of multi-unit assignment mechanisms and for market design more broadly.

Keywords: multi-unit assignment, market design, course allocation, random serial dictatorship, ex-ante efficiency, ex-post efficiency, strategyproofness, strategic behavior, field data.

Working draft. Comments especially welcome.

---

\*We are extremely grateful to Sheila Connelly, Assistant Director, MBA Registrar Services, for providing us with the data and sharing her extensive knowledge about the HBS course allocation mechanism with us. For suggestions, we thank Susan Athey, Drew Fudenberg, Fuhito Kojima, Jordi Massó, Paul Milgrom, David Parkes, Ariel Pakes, Parag Pathak, Al Roth, and seminar audiences at ASSA 2008, Dagstuhl 2007, Paris Sorbonne, Harvard, Toulouse School of Economics and Autònoma de Barcelona. Financial support from the European Research Council, the Belgian National Science Foundation (FNRS) and the Division of Research at Harvard Business School is gratefully acknowledged.

<sup>†</sup>University of Chicago Booth School of Business. Email: eric.budish@chicagobooth.edu

<sup>‡</sup>FNRS, Université Libre de Bruxelles (Solvay Brussels School of Economics and Management and ECARES), and CEPR. Email: Estelle.Cantillon@ulb.ac.be

# 1 Introduction

Nearly all educational institutions place limits on the number of students in a class. An extensive education literature has debated the relationship between class size and students' educational attainment.<sup>1</sup> We are interested in a market-design problem that arises from class-size limits: if it is not possible for all students to take their most desired schedule of courses, then how should seats in over-demanded courses be allocated? Press coverage and anecdotal evidence suggest that the scarcity problem is particularly acute in higher education, especially at professional schools.<sup>2</sup>

Economists have increasingly played an active role in designing solutions to real-world resource-allocation problems like course allocation. In the past, mechanisms suggested by the pure theory literature have often provided a useful starting point. One prominent example comes from the redesign of the institution that matches medical-school graduates to hospital residency positions, on which Roth (2002) remarks that “the simple theory [of Gale and Shapley (1962)] offered a surprisingly good guide to the design [of the Roth-Peranson (1999) algorithm].” Other important examples include the design of school-choice procedures, advertising markets on internet search engines, and combinatorial auctions.<sup>3</sup>

For course allocation, however, the extant theory literature consists mainly of negative results. Course allocation is an example of a *multi-unit assignment problem*, in which a set of indivisible objects (seats in courses) is to be allocated amongst a set of agents (students), the agents have multi-unit demand (for schedules of courses), and there are exogenous restrictions against the use of monetary transfers. One set of negative results shows that specific efficiency, fairness and incentives criteria that are compatible for the single-unit assignment problem are impossible to achieve for the multi-unit assignment problem (Sonmez, 1999; Konishi, Quint and Wako, 2001; Klaus and Miyagawa, 2001; Manea, 2007; Kojima, forthcoming; Che and Kojima, forthcoming). A second set of results shows essentially that the *only* mechanisms that are ex-post Pareto efficient and strategyproof are dictatorships, which intuitively are highly unfair: for any two students, one gets to choose *all* her courses before the other gets to choose *any*. (Papai, 2001; Ehlers and Klaus, 2003).<sup>4</sup>

---

<sup>1</sup>See for instance Angrist and Lavy (1999), Hoxby (2000), Lazear (2001), Krueger (2003), and Rivkin, Hanushek and Kain (2005).

<sup>2</sup>See Bartlett (2008), Guernsey (1999), Lehrer (2008), and Neil (2008).

<sup>3</sup>New school choice procedures in New York and Boston incorporated aspects of Gale and Shapley's mechanism for two-sided matching and Shapley and Scarf's (1974) mechanism for single-unit assignment (Abdulkadiroglu and Sonmez 2003; Abdulkadiroglu et al 2005a, b). Google cites the influence of Vickrey's (1961) “Nobel Prize-winning economic theory” in the design of its auction for advertising slots (Edelman, Ostrovsky and Schwarz, 2008; Varian, 2007). On the relationship between theory and practice for combinatorial auction design, see Cramton et al (2007) and Milgrom's (2004) aptly named text “Putting Auction Theory to Work.”

<sup>4</sup>Budish (2009) formalizes the sense in which dictatorships are unfair in multi-unit assignment. Hatfield (2007) shows that the dictatorship theorem holds even in highly stylized settings, in which we do not allow courses to be

Given the lack of positive results from theory, a *sensible starting point for design is to see what we can learn from mechanisms that are actually used in practice*. In this paper, we study the mechanism used at Harvard Business School (HBS) since the mid-1990s to allocate roughly 9,000 elective course seats to about 900 second-year MBA students every year. We choose this mechanism for two reasons. First, it is a *prima facie* sensible mechanism, satisfying attractive efficiency and fairness properties and differing from the dictatorship in an intuitively attractive way.<sup>5</sup> Rather than choosing courses all at once, which will lead to highly unequal allocations, students take turns choosing a single course at a time. (More accurately, a computer chooses courses for them based on their reported preferences; the choosing order is random in the first round, and is reversed each subsequent round.) Similar mechanisms have been used for a long time both by other educational institutions and in other multi-unit assignment contexts (Brams and Straffin, 1979; Brams and Taylor, 1999). In this sense, the mechanism passes a market endurance test.

Second, we have great data. In addition to students' actual (strategic) reports of their preferences, we have data on students' *underlying truthful preferences*, from a survey conducted by the HBS administration. (We describe in detail in Section 5 our argument that the survey data are indeed truthful). Whereas strategic reports are naturally recorded by market administrators and commonly made available to researchers, data on underlying preferences typically are only available to researchers in laboratory settings. The combination of truthful and stated preferences is powerful for two reasons. First, it means that we can directly observe students' strategic manipulations and quantify their effect on welfare. Second, we can use the truthful preferences to simulate equilibrium play of the strategyproof Random Serial Dictatorship (RSD) suggested by the extant theory,<sup>6</sup> and so can compare the two mechanisms.

The main results of our analysis can be summarized as follows. The HBS mechanism is simple to manipulate in theory (Section 3), is heavily manipulated by students in practice (Section 5), and these manipulations cause substantial inefficiency, assessed either ex-ante or ex-post (Section 6). Yet, ex-ante welfare is higher under the HBS mechanism than under the ex-post efficient RSD; we trace the poor ex-ante performance of RSD to a phenomenon we call "callousness" that is specific to multi-unit assignment (Section 7).

In the Conclusion, we discuss our contribution to two active debates in the broader literature on market design: (i) the importance of strategyproofness; (ii) how to analyze the efficiency of random mechanisms. We also discuss how our findings suggest "where to look" for new solutions to the

---

substitutes or complements for each other, and ignore real-life complications like scheduling constraints, curricular constraints, prerequisites, etc.

<sup>5</sup>By contrast, Sonmez and Unver (forthcoming) show that another widely used course-allocation mechanism, the Bidding Points Mechanism, has a serious conceptual flaw. See also Krishna and Unver (2008) and Budish (2009; Section 7.3).

<sup>6</sup>Amongst dictatorship mechanisms, what distinguishes the Random Serial Dictatorship is that it satisfies the procedural fairness property of anonymity, because the choosing order is uniform random.

multi-unit assignment problem.

Our analysis begins by studying the theoretical properties of the HBS mechanism. It satisfies several criteria of fairness and yields outcomes that are consistent with ex-post Pareto efficiency if students reveal their preferences truthfully. However, the HBS mechanism is simple to manipulate: students should overreport how much they like popular courses and underreport how much they like unpopular courses, so they do not waste early-round draft picks on courses they can get in later rounds (Theorem 1). We provide a partial characterization of equilibrium behavior and equilibrium run-out times for courses (Theorems 3, 4). We find that strategic behavior affects students' welfare through two distinct channels. First, strategic behavior leads to congestion (popular courses reach capacity faster) which hurts students whose preferred courses are popular amongst other students. Second, strategic behavior can also lead to ex-post inefficient outcomes if students miss one of their preferred courses because they placed a less preferred but more popular course ahead of it in their submitted preferences. We show through an example that all students can be worse off ex-ante under strategic play.

Next, we describe the preferences data we collected. We have student-level survey data on preferences over courses at two different points in time, as well as reported preferences during a trial run and the real run of the HBS mechanism. We argue that the survey data collected prior to the initial allocation and the preferences submitted for the real run of the mechanism correspond to students' truthful preferences and equilibrium play respectively. We do so at the aggregate level and at the individual level. At the aggregate level, we show that submitted preferences at these two points in time differ significantly and in a way that is consistent with the predictions of the theory, and inconsistent with alternative explanations of behavior such as social learning or new information. At the individual level, we test whether students' submitted preferences rearrange popular and unpopular courses in a manner that is consistent with our characterization of equilibrium best responses (Lemma 2). We find that most students (82.2%) submit preferences that are consistent with equilibrium behavior. Most of those who do not seem to have changed their preference for some single course, or made a strategic error based on plausible incorrect beliefs about a course's popularity.

In Section 6 we quantify the welfare consequences of strategic play, by comparing actual equilibrium play with non-equilibrium truthful behavior. In other words, our counterfactual exercise asks what would happen if the social planner knew students' preferences and used the HBS mechanism to allocate courses. We first document that strategic behavior causes ex-post inefficient allocations. On average our Pareto-improvement-seeking integer program is able to find beneficial trades involving 84% of students and 15% of course seats. Ex-ante welfare comparisons are more subtle: our data consist of ordinal preferences over individual courses, but welfare depends on von Neumann-Morgenstern preferences over bundles. We develop a novel computational method which allows us

to draw welfare comparisons for many students based on this limited ordinal information: nearly half of students' are unambiguously harmed by strategic play, whereas only 5% unambiguously benefit. To reach comparisons for the other students, or to evaluate welfare from the perspective of a utilitarian social planner, we need to put more structure on preferences, i.e., on the map between ordinal preferences over individual courses and vNM preferences over bundles. Two simple measures that the HBS administration emphasizes are the likelihood that students obtain their single favorite course and the average rank of the ten courses students receive. For both measures, the magnitudes are large. Strategic play reduces the likelihood of receiving one's favorite course from 83% to 60%, and increases the expected average rank from 7.76 to 8.35 (higher is worse, 5.50 is bliss). More formally, if students have what we call lexicographic or average-rank preferences, we can conclude that a utilitarian social planner regards strategic play as harmful to welfare.

Because strategic behavior hurts students it is natural to compare the HBS mechanism with its strategyproof alternative, RSD. Here, our counterfactual exercise asks what would happen *in equilibrium* if the HBS administration adopted RSD as its course-allocation mechanism. Ex-post, we cannot Pareto rank the allocations from the HBS mechanism versus those from RSD. Lucky students who get to choose early will prefer RSD because they get to make all their choices before other students make any; unlucky students will have the reverse preference. Ex-ante, however, HBS looks much more attractive than RSD. If students have either average-rank preferences or lexicographic preferences, then both the large majority of individual students and a utilitarian social planner prefer the HBS mechanism to RSD. The magnitudes are large: using RSD would reduce the likelihood of receiving one's favorite course from 60% to 47%, and increase the average rank of the courses that a student receives from 8.35 to 9.84.

This is surprising at first because RSD is ex-post Pareto efficient whereas HBS is not. Our theoretical explanation for RSD's poor ex-ante performance is simple. Under RSD, fortunate students with good random draws make their second, third, ..., last choices independently of whether these courses would be some unfortunate students' *first* choices (Example 4, Theorem 6). That is, the lucky "callously disregard" the unlucky. In any one random draw, there is no way to improve the allocation of the unlucky without harming the lucky, so outcomes are ex-post Pareto efficient. But in expectation the harm to the unlucky exceeds the benefit to the lucky, so RSD does very poorly on measures of ex-ante efficiency. Note that RSD's unattractiveness does not depend on risk aversion; even risk-neutral students overwhelmingly dislike RSD ex-ante.

## 2 The course allocation problem

### 2.1 Environment

**Courses.** There is a finite set of  $C$  courses,  $\mathcal{C}$ .<sup>7</sup> Courses have capacities  $\mathbf{q} = (q_1, \dots, q_C)$ .

**Students and preferences.** There is a continuum of students described by the interval  $[0, 1]$  and endowed with the Lebesgue measure.<sup>8</sup> Each student  $s$  is endowed with a von Neumann-Morgenstern utility function  $u_s$  that indicates her utility from each bundle of courses, including singletons. Associated with each utility function  $u_s$  is an ordinal preference relation defined over permissible bundles of courses,  $P_s \in \mathcal{P}$ . We assume that the mapping from the set of students to the set of preference relations,  $\mathcal{P}$ , is measurable. We further assume that the utility functions are such that students' ordinal preferences over individual courses are strict, and that their ordinal preferences over bundles are responsive to their preferences for individual courses.<sup>9</sup>

**Demand.** Students are allowed to consume any bundle that consists of 0 or 1 seat of each course, and at most  $m > 1$  courses in total. Let  $r_s(c) \in \mathbb{N}$  denote course  $c$ 's rank in student  $s$ 's preferences over individual courses. Thus  $r_s(c) < r_s(c')$  if and only if  $cP_s c'$ , with  $r_s(c) = 1$  if  $cP_s c'$  for all  $c' \neq c$ . This allows us to define the demand for individual courses:

**Definition 1 (Demand for Courses).** The demand for course  $c$  is defined as  $D_c(\rho) = \frac{\int_0^1 1_{\{r_s(c) \leq \rho\}} ds}{q_c}$ ,  $\rho = 1, \dots, C$ .

The allocation problem is non-trivial if at least one capacity constraint binds. Thus, in the rest of our analysis we assume that there exists at least one course  $c$  such that  $D_c(m) > 1$ .

**Feasible Allocations.** An allocation in this environment is an assignment of courses to students. We denote by  $a_s \subset \mathcal{C}$  student  $s$ ' allocation of courses. An allocation is feasible if  $|a_s| \leq m$  for all  $s$  and  $\int_0^1 1_{\{c \in a_s\}} ds \leq q_c$  for all  $c$ . We denote by  $\mathcal{A}$  the set of feasible allocations. A random assignment is a probability distribution over feasible allocations. We denote by  $L(\mathcal{A})$  the set of random assignments.

**Ex-Ante and Ex-Post Efficiency.** A random assignment is ex-ante Pareto efficient if there is no other random assignment that all students weakly prefer and a strictly positive measure of students

---

<sup>7</sup>We use the terms "students" and "courses" because of our application. We could equally use the generic terms "agents" and "objects".

<sup>8</sup>The use of a continuum of students is a technical, rather than substantive, assumption. It simplifies proofs and helps clarify the key forces behind the results.

<sup>9</sup>Preferences are responsive if, for any student  $s$ , courses  $c, c'$ , and bundle of courses  $X$  with  $c, c' \notin X$  and  $|X| < m$ ,  $cP_s c' \iff (X \cup c)P_s(X \cup c')$ . Also,  $cP_s \emptyset \iff (X \cup c)P_s(X \cup \emptyset)$  (Brams and Straffin, 1979, Roth, 1985). While restrictive, survey evidence suggests that responsiveness is a reasonable assumption in the case of Harvard Business School, and the HBS elective curriculum is explicitly designed to avoid overlap or interdependence amongst courses. The course-allocation mechanism proposed in Budish (2009) relaxes the responsiveness assumption.

strictly prefers. A feasible allocation is ex-post Pareto efficient if there is no other feasible allocation that all students weakly prefer and a strictly positive measure of students strictly prefers.

**Information.** We assume that the realization of students' preferences is common knowledge. Since we are working with a continuum, this is equivalent to assuming that students' preferences are private information but that the distribution over preferences is common knowledge.

## 2.2 Allocation Mechanisms

We focus attention on two specific course-allocation mechanisms, the HBS mechanism and the Random Serial Dictatorship (RSD).<sup>10</sup> Both mechanisms are ordinal in the sense that they take as inputs ordinal information about students' preferences (Bogomolnaia and Moulin, 2001). In both mechanisms, each student  $s$  reports a rank-order list (ROL)  $\widehat{P}_s$  indicating her ordinal preferences over individual courses. We write  $\widehat{P}_s : c_1, c_2, c_3, \dots$  to describe that student  $s$  puts course  $c_1$  ahead of  $c_2$ , course  $c_2$  ahead of  $c_3$ , and so on (with a slight abuse of notation, we will also write  $P_s : c_1, c_2, c_3, \dots$ , to describe her true preferences over individual courses). Then, the mechanism uniformly randomly selects a priority order over students, which is a bijection from the set of students onto itself. Let  $\lambda(t)$  denote the student who has priority  $t$ , and  $\lambda^{-1}(s)$  the priority of student  $s$ . The set of priority orders is  $\mathcal{L}$ . We assume that all elements of  $\mathcal{L}$  have a measurable inverse.

Under RSD, students are allocated their courses all-at-once in ascending priority order. The algorithm has a single round that takes place from time  $t = 0$  to time  $t = 1$ . At time  $t$ , student  $\lambda(t)$  is allocated a seat in her  $m$  most-preferred courses on  $\widehat{P}_s$  that still have remaining capacity.<sup>11</sup>

Under the HBS mechanism, students are allocated courses one-at-a-time over a series of  $m$  rounds. In odd rounds, which occur during time intervals  $[0, 1], [2, 3], \dots$ , students are allocated courses one-at-a-time in ascending priority order. In even rounds, which occur during time intervals  $[1, 2], [3, 4], \dots$ , students are allocated courses one-at-a-time in descending priority order. When it is student  $s$ ' turn in the algorithm to be allocated a course, she is allocated her most-preferred course on  $\widehat{P}_s$  that (i) she has not already been allocated in a previous round; and (ii) still has remaining capacity. Following the  $m$  rounds of the HBS mechanism, students have one additional opportunity to modify their schedule. Students can drop courses they obtained in the initial allocation and add courses that have excess capacity. This is conducted using a multi-pass algorithm that cycles over students (using a new random priority order) until no more add-drop requests can be satisfied. In particular, the algorithm satisfies a student's add-drop request only if the course that the student requests has spare capacity. It does not look for Pareto-improving trades amongst students. For modeling purposes, we model the add-drop phase as a random serial dictatorship where the only

<sup>10</sup>Sonmez and Unver (forthcoming) and Budish (2009) describe other course-allocation mechanisms used in practice.

<sup>11</sup>A course has capacity remaining at time  $t$  if the measure of students allocated a seat in that course during time  $[0, t)$  is strictly less than the course's capacity.

courses that can be requested are those with spare capacity at the end of round  $m$  of the initial allocation. Students have the opportunity to modify their reported preferences, and a new random priority order is drawn. Thus, each student in turn creates the best possible schedule out of the courses they got in the initial allocation and those still with excess capacity.

### 2.3 Equilibrium

A Nash equilibrium in this setting is a measurable mapping from the set of students to the set of mixed strategies  $\Delta\mathcal{P}$ . There exists a pure strategy Nash equilibrium in both mechanisms:

1. RSD is dominant-strategy incentive compatible and we focus on this pure strategy equilibrium in the remainder.
2. The add-drop phase of the HBS mechanism is equivalent to RSD and thus we also restrict attention to the equilibrium where students report truthfully in this stage. Existence of a pure strategy Nash equilibrium is guaranteed in the initial allocation phase of the HBS mechanism because the action space is finite and students' expected utilities are continuous in the strategies of the other students and only depend on the fraction of students who report each preference profile (Schmeidler, 1973).

## 3 Properties of the HBS mechanism

The HBS mechanism satisfies several attractive efficiency and fairness properties if we ignore incentives and assume that students report their preferences truthfully. In terms of efficiency, it yields allocations that are ex-post Pareto possible (Brams et al, 2003). This means that there exist preferences over bundles of courses that are responsive to the reported preferences over individual courses, and for which the allocation is ex-post Pareto efficient. With respect to fairness, it is procedurally fair in the ex-ante sense of equal treatment of equals, and also in an interim sense, in that no students' set of choosing times dominates any others'. It also satisfies attractive criteria of outcome fairness, as described in Budish (2009).

These attractive properties explain the HBS administration's decision to adopt this mechanism, and may explain the widespread use of similar draft mechanisms in practice. However, as the following example illustrates, truthful play is not the expected outcome in the HBS mechanism.

**Example 1 (Over-reporting and Congestion)** Let  $m = 2$  and suppose there are 4 courses

with capacity of  $\frac{2}{3}$  seats each. Preferences are as follows:

| Proportion of Population | Type  | Preferences          |
|--------------------------|-------|----------------------|
| $\frac{1}{3}$            | $P_1$ | $c_1, c_2, c_3, c_4$ |
| $\frac{1}{3}$            | $P_2$ | $c_2, c_1, c_3, c_4$ |
| $\frac{1}{3}$            | $P_3$ | $c_1, c_3, c_4, c_2$ |

Truthful play is not a best response for the  $P_2$  types. Indeed, suppose that student  $s$  is of type  $P_2$ . If all other students play truthfully, student  $s$  gets  $\{c_2, c_3\}$ . If, instead, he submits preferences  $\widehat{P}_2 : c_1, c_2, c_3, c_4$ , then he gets his first choice bundle  $\{c_1, c_2\}$  for sure.

In fact, the  $P_1$  and  $P_3$  types reporting truthfully and the  $P_2$  types reporting  $\widehat{P}_2 : c_1, c_2, c_3, c_4$  is a Nash equilibrium, independently of risk attitudes or cardinal information about preferences. In this equilibrium, the  $P_1$  and  $P_2$  types get  $\{c_1, c_2\}$  with probability  $\frac{2}{3}$  and  $\{c_2, c_3\}$  with probability  $\frac{1}{3}$  and the  $P_3$  types get  $\{c_1, c_3\}$  with probability  $\frac{2}{3}$  and  $\{c_3, c_4\}$  with probability  $\frac{1}{3}$ . More students request  $c_1$  (the most popular course based on true demand) in the first round than under truthful play, making  $c_1$  fill up earlier in the round than under truthful play.

The story that Example 1 tells is that students in the HBS mechanism will have a tendency to overreport their preferences for popular courses, and that this causes those courses to reach capacity sooner. A  $P_2$  type should not waste his first-round choice on  $c_2$ , since he can get it for sure in the second round, and if he waits until round two to ask for  $c_1$  he is sure not to get it. Instead, he should attempt to obtain the popular  $c_1$  in the first round.

In this section, we explore the incentives in the HBS mechanism further and provide a partial characterization of equilibrium outcomes. In particular, we argue that detecting profitable deviations from truthful behavior is very easy and that there will indeed be a tendency in equilibrium to overreport preferences for popular courses. This leads to congestion and earlier run-out times for courses. In the last subsection, we explore the welfare properties of the Nash equilibrium in the HBS mechanism. One consequence of strategic behavior is redistributive: some students, especially those whose top choices are not very popular, are better off; others, especially those whose top choices are very popular, are worse off. We also show that strategic behavior can lead to ex-post inefficiency and pin down conditions when this is not the case.

### 3.1 Incentives in the HBS mechanism

The basic trade-off that students face in the HBS mechanism is that upgrading a course relative to their truthful preferences increases the chance of getting that course, at the risk of missing another, possibly preferred, course. This trade-off is intrinsic to the HBS mechanism and does not depend on market size. Indeed, Example 1 would work just as well with three students.

To be able to say something about incentives in the HBS mechanism without making any assumption on risk attitudes, we analyze best responses for a fixed strategy profile  $\widehat{\mathbf{P}} = \{\widehat{P}_s\}_{s \in [0,1]}$  and a fixed priority order  $\lambda$ . The outcome, then, is determinate (and, in particular, courses are characterized by their run-out times). A property of best responses will hold irrespective of any assumption on risk attitudes, if we can show that it holds for all  $\lambda$ .

Consider student  $s$  and relabel courses such that student  $s$ 's strategy reads  $\widehat{P}_s : c_1, c_2, \dots, c_C$ . Because of the continuum assumption, his strategy does not affect run-out times and thus does not affect the outcome and timing of requests by the other students. The only thing it does affect is whether student  $s$  gets a seat in the courses he requests. He will do so if he requests a particular course before that course runs out. He will not do so otherwise.

Our workhorse for most of the proofs in this section is the comparison between two strategies by student  $s$  that differ from one another by the position of a single course. We show that allocations for student  $s$  under those two strategies differ at most by one course.

For example, denote by  $\widehat{P}_s^{c_k \downarrow l}$ , the strategy that corresponds to  $\widehat{P}_s$  except that course  $c_k$  has been downgraded to position  $l$  in the ROL, i.e.  $\widehat{P}_s^{c_k \downarrow l} : c_1, \dots, c_{k-1}, c_{k+1}, \dots, c_l, c_k, c_{l+1}, \dots$ . Until the procedure reaches the  $k^{\text{th}}$  position in student  $s$ 's ROL, the timing of requests and the outcomes are identical under both strategy profiles. From then on,  $\widehat{P}_s^{c_k \downarrow l}$  asks for  $c_{k+1}$  one round earlier than  $\widehat{P}_s$ , and if either both strategies get  $c_{k+1}$  or both do not, then  $\widehat{P}_s^{c_k \downarrow l}$  asks for  $c_{k+2}$  one round earlier than does  $\widehat{P}_s$ , and so on. If, before  $\widehat{P}_s^{c_k \downarrow l}$  requests  $c_k$ , it gets a course that  $\widehat{P}_s$  does not, then the two strategies get back in synch, requesting (and getting) the same courses at the same times. They can get back out of synch if  $\widehat{P}_s^{c_k \downarrow l}$  turns out also to get  $c_k$ , only now  $\widehat{P}_s$  is making requests for  $c_{l+1}$  one round earlier than does  $\widehat{P}_s^{c_k \downarrow l}$ . Clearly, the exact details on the timing and outcome for student  $s$  depend on the availability of courses when he requests them (they are developed in the proofs). The key aspect to note here is that requests by student  $s$  under both strategies are either in synch or out of synch by a maximum of one round. When requests are in synch, he must get the same outcome under both strategies. When requests are not in synch, he may get a course under one strategy that he does not get under the other strategy. Because requests become in synch after such event, the difference in outcomes between the two strategies is one course at most.

Let  $a_s(\widehat{\mathbf{P}}, \lambda) \subset \mathcal{C}$  indicate student  $s$ 's final allocation, including what happens in the add-drop phase, when students use the strategy profile  $\widehat{\mathbf{P}}$  and  $\lambda$  is the realized priority order. Let  $a_s(\widehat{\mathbf{P}})$  refer to student  $s$ 's final (random) allocation under profile  $\widehat{\mathbf{P}}$ .

**Definition 2 (Popularity):** Course  $c$  is  $\widehat{\mathbf{P}}$ -popular if there exists a positive measure of students for whom  $\Pr(c \notin a_s(\widehat{\mathbf{P}}) \text{ and } c' \in a_s(\widehat{\mathbf{P}})) > 0$  for some  $c'$  such that  $c \widehat{P}_s c'$ . Course  $c$  is  $\widehat{\mathbf{P}}$ -unpopular otherwise.

In words: the popularity of a course is defined relative to a strategy profile. A  $\widehat{\mathbf{P}}$ -popular course runs out with positive probability under strategy profile  $\widehat{\mathbf{P}}$ . A  $\widehat{\mathbf{P}}$ -unpopular course never runs out.

Note that, by the continuum assumption,  $\widehat{\mathbf{P}}$ -popular courses are also  $(\widehat{P}'_s, \widehat{\mathbf{P}}_{-s})$ -popular for any  $\widehat{P}'_s$ , so that we can as well talk about  $\widehat{\mathbf{P}}_{-s}$ -popular courses.

The next result shows that it is easy for students to find profitable deviations from truthful play: the deviations that Theorem 1 identifies only require that students know which courses are likely to run out, and which aren't.

**Theorem 1 (Simple Manipulations):** Fix  $\widehat{\mathbf{P}}_{-s}$ . Form the strategy  $\widehat{P}_s^{\text{simple}}$  by taking the first  $m$  courses in  $P_s$  and rearranging them so that  $c\widehat{P}_s^{\text{simple}}c'$  whenever:

1.  $cP_sc'$  and both are  $\widehat{\mathbf{P}}_{-s}$ -popular or both are  $\widehat{\mathbf{P}}_{-s}$ -unpopular
2.  $c$  is  $\widehat{\mathbf{P}}_{-s}$ -popular and  $c'$  is  $\widehat{\mathbf{P}}_{-s}$ -unpopular

The strategy  $\widehat{P}_s^{\text{simple}}$  generates weakly greater utility than truthful play  $P_s$ .

Clearly, this simple strategy leads to overreporting of preferences for popular courses. Moreover, if everyone adopts this simple strategy starting from truthful play, these courses are likely to run out earlier than under truthful play. This is in line with Example 1. However, Theorem 1 is only indicative that this may be a general feature of equilibrium. Indeed, it is not yet an equilibrium characterization, nor even a characterization of best responses. We turn to this question next.

### 3.2 Equilibrium

We can identify two environments where truthful play in the HBS mechanism is always a Nash equilibrium.

**Theorem 2 (Truthful Play in Special Cases):** Consider the following two environments:

1. Identical preferences:  $P_s = P_{s'}$  for all  $s, s'$
2. Independent preferences: for any two  $\mathbf{P}$ -popular courses  $c, c'$ ,  $D_c(\rho) = D_{c'}(\rho)$  for  $\rho = 1, \dots, C$ , and all students rank  $\mathbf{P}$ -popular courses ahead of  $\mathbf{P}$ -unpopular courses

In either environment,  $\widehat{\mathbf{P}}^* = \mathbf{P}$  is a Nash equilibrium of the HBS mechanism.

The intuition is the following. When students have identical preferences, if a student prefers one course to another, then so do all other students and thus it is not possible to *overreport* one's preferences for more popular courses. Likewise, when all courses for which demand exceeds supply are equally popular, there is no basis for misreporting. Theorem 2 indicates that *partial* correlation of preferences is what drives strategic misreporting in the HBS mechanism.

We next provide a necessary condition for best responses in the HBS mechanism.

**Lemma 2 (Best-response Characterization):** Consider any equilibrium strategy profile  $\widehat{\mathbf{P}}$ . Suppose  $c$  is  $\widehat{\mathbf{P}}$ -popular, and that  $r_s(c) \leq m$ . Then, one of the three following properties must hold:

- (i)  $c$  appears before all  $\widehat{\mathbf{P}}$ -unpopular courses in  $\widehat{P}_s$ ,
- (ii)  $\Pr(c \in a_s(\widehat{\mathbf{P}})) = 1$ , or
- (iii)  $\Pr\left(c \in a_s\left(\widehat{P}_s^{c \uparrow k}, \widehat{\mathbf{P}}_{-s}\right)\right) = 0$  where  $k$  is the position of the first (highest)  $\widehat{\mathbf{P}}$ -unpopular course on  $\widehat{P}_s$ .

It is interesting to compare and contrast Lemma 2 and Theorem 1. Like Theorem 1, Lemma 2 suggests that popular courses will be ahead of unpopular courses in ROLs. There are two exceptions. A first exception is when student  $s$  is sure to get the popular course even if he requests it later. A second exception is when student  $s$  would not get the course even if he placed it before all unpopular courses. In that case, we cannot rule out that, at equilibrium, a student puts that course in another irrelevant position, that is, after some unpopular courses.

Another difference between Theorem 1 and Lemma 2 is that Lemma 2 is silent about the relative ordering of popular courses in students' ROLs. How students should optimally rank popular courses depends on the relative run-out times of these courses under  $\widehat{\mathbf{P}}$ , on the cardinal utilities they attach to these courses and on their attitudes towards risk. We will come back to this point later in the section. In the meantime, we state the properties of the equilibrium in the HBS mechanism that we can establish, just based on ordinal preferences.

**Theorem 3 (Equilibrium Characterization):** Suppose that  $\widehat{\mathbf{P}}$  is a Nash equilibrium, and that  $D_c(m) > 1$  for some  $c$ . Then,

1.  $c$  runs out with probability one
2. the supremum of run-out times for course  $c$  over the different realizations of  $\lambda, \bar{t}_c$ , is weakly less than the number of  $\widehat{\mathbf{P}}$ -popular courses.

Theorem 3 says that courses for which demand exceeds supply based on truthful reports will run out in any equilibrium. Moreover, such courses will reach capacity in the first  $k$  rounds, where  $k$  is the number of  $\widehat{\mathbf{P}}$ -popular courses.

However, Theorem 3 remains silent concerning the relative run-out times of popular courses. The reason is that equilibrium in the HBS mechanism has a coordination feature: if a course is expected to be ranked high by others then students will tend to rank it higher too. This yields equilibrium multiplicity as the next example illustrates.

**Example 2 (Multiple Equilibria)** Let  $m = 2$ . Courses have 0.4 capacity. Courses  $c_1, c_2, c_3$  have excess demand, all other courses do not. Students' preferences are as follows (where "other"

stands for courses other than  $c_1, c_2, c_3$ ) :

| Proportion of Population | Type  | Preferences        |
|--------------------------|-------|--------------------|
| .25                      | $P_1$ | $c_1, c_2$ , other |
| .25                      | $P_2$ | $c_2, c_1$ , other |
| .30                      | $P_3$ | $c_3$ , other      |
| .10                      | $P_4$ | $c_3, c_1$ , other |
| .10                      | $P_5$ | $c_3, c_2$ , other |

Truthful play is always an equilibrium. In round 1, the  $P_1$  and  $P_2$  types get their first choice and the  $P_3, P_4$  and  $P_5$  types get their first choice with probability 0.8, else they get their second choices. The remaining capacity for courses  $c_1$  and  $c_2$  at the beginning of round 2 depends on the priority order over students in round 1. In expectation, these courses have 0.13 remaining capacity ( $0.40 - 0.25 - 0.02$ ). The expected demand for these courses in round 2 is equal to 0.33 ( $0.25 + 0.08$ ) so a request is satisfied with probability  $0.13/0.33 \approx 0.4$ .<sup>12</sup> Thus, the  $P_1$  types get their best bundle  $\{c_1, c_2\}$  with probability 0.4 and otherwise they get their second best bundle  $\{c_1, \text{other}\}$  (the outcome for the  $P_2$  types is symmetric). The  $P_3$  types get their best bundle,  $\{c_3, \text{other}\}$ , with probability 0.8, and their second best bundle otherwise. The  $P_4$  types get their best bundle,  $\{c_3, c_1\}$ , with probability 0.32, their second best bundle,  $\{c_3, \text{other}\}$ , with probability 0.48 and  $\{c_1, \text{other}\}$  otherwise (the outcome for the  $P_5$  types is symmetric). Clearly, no type has a profitable deviation.

If the  $P_4$  and  $P_5$  types' intensity of preference for  $c_3$  versus  $c_1$  and  $c_2$ , respectively, is not too large, there exists another equilibrium in which types  $P_1, P_2$  and  $P_3$  play truthfully, the  $P_4$  types submit  $\hat{P}_4 : c_1, c_3, \text{other}$ , and the  $P_5$  types submit  $\hat{P}_5 : c_2, c_3, \text{other}$ . In this equilibrium, the  $P_4$  types get their best bundle,  $\{c_1, c_3\}$  with probability 0.5 and  $\{c_1, \text{other}\}$  otherwise. If they deviate from this equilibrium to play truthfully (the only deviation to consider given their preferences), they receive their best bundle with probability 0.2 ( $c_1$  is available in round 2 with probability  $\frac{4 - \frac{1 - .25}{.25}}$ ) and get  $\{c_3, \text{other}\}$  with probability 0.8. This is less preferable if the intensity of preference for  $c_3$  is not too large. The analysis for the  $P_5$  types is analogous.

In Example 2, course  $c_3$  runs out in equilibrium in round 1 or 2, i.e., at the same time or later than under truthful play, even though more students have  $c_3$  as their top choice relative to other courses. This is consistent with Theorem 3, which says that  $c_1, c_2$  and  $c_3$  should all run out with probability 1 during the first two rounds, but otherwise does not pin down the relative timing of these events.

<sup>12</sup>Recall that a property of expected utilities is that preferences over compound lotteries correspond to the preferences over the reduced lottery. Thus it is sufficient to look at expected probabilities of outcomes.

To go beyond Theorem 3, we make additional assumptions on the relationship between students' ordinal preferences over courses and their utility functions. In Example 2,  $c_3$  sells out at equilibrium in the first round if the  $P_4$  and  $P_5$  types students prefer the lottery  $[0.2 : \{c_1, c_3\}; 0.8 : \{c_3, \text{other}\}]$  to the lottery  $[0.5 : \{c_1, c_3\}; 0.5 : \{c_1, \text{other}\}]$ . This motivates the following restriction:

**Definition 3 (Lexicographic Preferences):** Consider two lotteries over final allocations,  $L_1$  and  $L_2 \in L(\mathcal{A})$ . Fix an arbitrary  $s$ , and label courses so that  $P_s : c_1, c_2, \dots, c_C$ . Let  $p_1(c)$ ,  $p_2(c)$  denote the probability of getting  $c$  under lottery  $L_1$  and  $L_2$  respectively. We say that student  $s$  has lexicographic preferences if he prefers  $L_1$  to  $L_2$  whenever there exists any  $k \in \mathbb{N}$  such that  $p_1(c_i) \geq p_2(c_i)$  for all  $i = 1, 2, \dots, k$  with at least one strict inequality.

In Example 2, if the  $P_4$  and  $P_5$  types have lexicographic preferences, they prefer any lottery in which they obtain  $c_3$  for sure to any lottery in which they don't, and so truthful play is the unique equilibrium. This yields sharper predictions about the structure of students' best responses in the HBS mechanism.

**Lemma 3 (Best-response Characterization with Lexicographic Preferences):** Suppose students have lexicographic preferences and consider any equilibrium strategy profile  $\widehat{\mathbf{P}}$ . Suppose  $c$  is  $\widehat{\mathbf{P}}$ -popular, and  $r_s(c) \leq m$ . Then, one of the three properties must hold:

- (i)  $c$  appears before all  $\widehat{\mathbf{P}}$ -unpopular courses in  $\widehat{P}_s$  and, for all  $c'$ ,  $cP_s c' \Rightarrow c\widehat{P}_s c'$
- (ii)  $\Pr(c \in a_s(\widehat{\mathbf{P}})) = 1$ , or
- (iii)  $\Pr\left(c \in a_s\left(\widehat{P}_s^{\text{simple}}, \widehat{\mathbf{P}}_{-s}\right)\right) = 0$ .

In words, Lemma 3 says that students only downgrade a course when they are sure to get it (or when they cannot hope to get it anyways). Theorem 4 provides a tighter characterization of equilibrium run-out times:

**Theorem 4 (Equilibrium with Lexicographic Preferences):** Suppose students have lexicographic preferences. Let  $\widehat{\mathbf{P}}$  be a Nash equilibrium. Then:

1. If  $D_c(m) > 1$  for some,  $\bar{t}_c \leq \rho_c = \inf\{\rho : D_c(\rho) > 1\}$
2.  $\min_r\{\widehat{D}_c(r) \geq 1\} \leq \min_r\{D_c(r) \geq 1\}$  where  $\widehat{D}_c(r)$  corresponds to the reported demand under  $\widehat{\mathbf{P}}$ .

Theorem 4 provides an upper bound on the times by which courses run out, based on their true demand, and a prediction on the relationship between truthful demand and reported demand at equilibrium. We revisit these predictions in section 5.

### 3.3 Welfare

Strategic behavior in the HBS mechanism has two effects. On the one hand, it helps students who, by overreporting their preferences for popular courses, increase their chances of getting them. On

the other hand, it creates congestion for these courses and hurts students who value these courses highly. This suggests that strategic behavior has, at the very least, redistributive consequences. This can be seen in Example 1 where the  $P_1$  and  $P_3$  types are better off ex-post and ex-ante from truthful play, and the  $P_2$  types are better off under strategic play.<sup>13</sup>

Strategic behavior also has efficiency consequences because of the intrinsic trade-off that students face between upgrading a less preferred but very popular course at the risk of potentially missing a seat in a preferred but less popular course. The next example illustrates this ex-post inefficiency of equilibrium in the HBS mechanism.

**Example 3 (Ex-post Inefficiency of Strategic Play)** Let  $m = 2$ . Courses  $c_1, c_2$  have excess demand with respective capacity 0.6 and 0.8. All other courses do not. Suppose preferences are as follows (where "other" stands for courses other than  $c_1$  or  $c_2$ )

| Proportion of Population | Type  | Preferences              |
|--------------------------|-------|--------------------------|
| 0.3                      | $P_1$ | $c_1, c_2, \text{other}$ |
| 0.4                      | $P_2$ | $c_2, c_1, \text{other}$ |
| 0.3                      | $P_3$ | $c_2, \text{other}$      |

Consider the strategy profile where all students of type  $P_1$  play  $\widehat{P}_1 : c_2, c_1, \text{other}$ , and the  $P_2$  and  $P_3$  types submit truthful ROLs. Under this strategy profile, all students request  $c_2$  in round 1, which means that only the first 0.8 are successful. Those who do not get their first choice get their second choice in round 1. At the beginning of round 2, 0.46 seats ( $0.6 - (0.7)(0.2)$ ) remain in  $c_1$  in expectation, whereas 0.56 students requests it. This means only 82% of these students are successful. Thus, the  $P_1$  and  $P_2$  types face the following lottery:

$$[0.66 : \{c_1, c_2\}; 0.20 : \{c_1, \text{other}\}; 0.14 : \{c_2, \text{other}\}] \quad (1)$$

To check whether this strategy profile is an equilibrium, we only need to look at the opportunity for a  $P_1$  student to deviate and submit his truthful preferences. If he does, he gets the deterministic outcome  $\{c_1, \text{other}\}$ . Thus  $\widehat{P}_1, P_2, P_3$  is an equilibrium if all  $P_1$  students prefer the lottery in (1) over the deterministic outcome  $\{c_1, \text{other}\}$ .

This equilibrium is ex-post inefficient because it is possible that a  $P_1$  student ends up with  $\{c_2, \text{other}\}$  and that a  $P_2$  student ends up with  $\{c_1, \text{other}\}$ . Those students would prefer to trade. (in expectation, 0.042  $P_1$  students get  $\{c_2, \text{other}\}$  and 0.08  $P_2$  students get  $\{c_1, \text{other}\}$ ).

---

<sup>13</sup>The  $P_1$  and  $P_3$  types got their best bundle under truthful play and now only get it with probability 1/3. The  $P_2$  types who never got their top bundle now it with probability 1/3, and otherwise they get what they used to get for sure under truthful play.

Because the ex-post inefficiency of the HBS mechanism is due to risk-taking behavior by students who deviate from truthful behavior, we can expect that no such inefficiency occurs in settings where, at equilibrium, students do not deviate from truthful behavior or when risk aversion is so high or preferences so extreme that they do not take any risk. Theorem 5 confirms this.

**Theorem 5 (Ex-post Efficiency in Special Cases)** (1) Nash equilibria in truthful strategies are always ex-post efficient possible, (2) All Nash equilibria are ex-post efficient possible when preferences are lexicographic.

By contrast, ex-post inefficiencies are likely when preferences are correlated, students are not very risk averse and cardinal utilities attached to each course are not very different.

Not surprisingly, the HBS mechanism is also ex-ante inefficient, independently of whether it is ex-post inefficient or not, for the very same reasons why RSD is inefficient when students require a single course (Bogomolnaia and Moulin, 2001). More interestingly, the next example shows that strategic behavior can hurt all students ex-ante, relative to truthful behavior. In other words, strategic behavior has consequences beyond redistribution. We will revisit this issue in section 6.

**Example 3 (cont'd) Strategic Behavior May Hurt All Students Ex-Ante** Consider again Example 3. The next table compares the lotteries that students face under truthful behavior and under the equilibrium identified in Example 3 above.

| Type  | Lottery under truthful play                           | Lottery under strategic behavior  |
|-------|---|---|
| $P_1$ | $[0.33 : \{c_1, c_2\}; 0.67 : \{c_1, \text{other}\}]$ | $[0.66 : \{c_1, c_2\}; 0.20 : \{c_1, \text{other}\}; 0.14 : \{c_2, \text{other}\}]$ |
| $P_2$ | $[0.75 : \{c_1, c_2\}; 0.25 : \{c_2, \text{other}\}]$ | $[0.66 : \{c_1, c_2\}; 0.14 : \{c_2, \text{other}\}; 0.20 : \{c_1, \text{other}\}]$ |
| $P_3$ | $[1 : \{c_2, \text{other}\}]$                         | $[0.80 : \{c_2, \text{other}\}; 0.20 : \{\text{other}, \text{other}\}]$             |

Clearly, the  $P_2$  and  $P_3$  types are worse off under strategic behavior, independently of cardinal information about their preferences. The  $P_1$  types are worse off if, for example  $u_s(\{c_1, c_2\}) = 4$ ,  $u_s(\{c_1, \text{other}\}) = 3.4$  and  $u_s(\{c_2, \text{other}\}) = 0.8$ . (These numbers also ensure that strategic behavior is indeed an equilibrium.)

## 4 Description of Data

Our dataset covers the allocation of second year courses at Harvard Business School during the 2005-2006 academic year.

### 4.1 Timing of actions and information

Figure 1 summarizes the timing of actions and the timing of the information received by students. Students are asked three times for their preferences as part of the initial allocation process: in

early May, in mid-May and at the end of July. Prior to this, students have information on course overenrollment in the previous year, and they have the official evaluations for the Winter and Fall 2004 courses, as well as unofficial course evaluations.

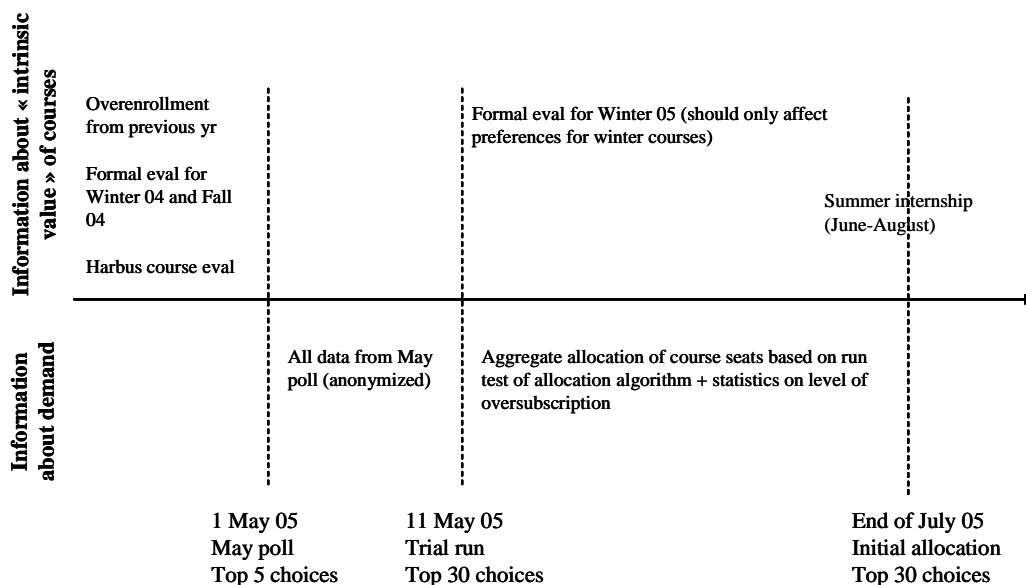


Figure 1: Timing of information and actions

In early May, they are asked to participate in a poll where they must rank their top 5 courses. Participation is voluntary. The results are used to aggregate information about demand and adjust some course capacities.<sup>14</sup> The students have access to the full results, except for the student identities which are removed.

The following week, students participate in a trial run of the allocation mechanism. Participation is compulsory and students must rank their top 30 courses (rankings can be section-specific for courses offered in different sections). The administration reports the resulting course enrollments based on one single run of the algorithm. For courses at capacity, students are told how many times a course was overenrolled based on the submitted preferences. In addition, the administration reports the 10 courses most often ranked first in the submitted rank order lists (ROLs) with the number of times each was ranked first. Students do not receive any feedback on their individual assignment of courses from the trial run.

Finally, end of July is the deadline for submitting ROLs for the real run of the mechanism. The ROLs submitted for the May trial run serve as the default ROLs in case a student does not submit

<sup>14</sup>The exact text is the following: "This poll has been set up to gauge current interest in 2005-06 courses. Be sure you enter your top 5 course selections for the coming year, with #1 the course you want most. Your selections will be anonymous to others. As a participant, you'll be able to view anonymous results on Friday, May 6."

new preferences. Students receive their allocation in early August. Some changes are possible at the beginning of each semester during the "add-drop" phase.

We have already mentioned the feedback students receive about aggregate demand after the May poll and the trial run. These are the most important new pieces of information they receive between the May poll and the July run. In addition, they receive the official course evaluations for the Winter 2005 courses after the trial run but, according to the HBS administration, this information was already available early May through the student newspaper, the Harbus. Very few changes were also made to courses. Specifically, two courses were added between the May poll and the trial run. One course was added, four courses were cancelled, one full semester course was changed into a half course between the trial run and the July run and several courses had their capacities increased or decreased slightly. Finally, note that students usually work as interns during the summer and this experience may impact their preferences over courses.

## 4.2 Course Characteristics

Our data contain all course characteristics, including section, capacities, term and scheduling information as they were available at the time of the May trial run and the July run of the algorithm. Seats in 71 courses and 21 half-courses (147 sections) were offered in May for a total capacity of 11,871 seats. Course capacities ranged from 12 to 404 students. The numbers for July were 67 courses and 22 half-courses (141 sections) for a total of 10,898 seats. The capacities range was the same. A total of 9,269 seats were allocated in the July 2005 run of the algorithm.<sup>15</sup>

## 4.3 Submitted Preferences

Our data contain the submitted preferences in the May poll, May trial run and the July run of the algorithm with student identifiers. In addition, we asked students in January 2006 to report their top 30 choices. The poll was conducted after the add-drop phase for the second semester but before courses started. In the poll, we explicitly asked students to rank the courses according to their true preferences, independently of whether they got the course or not. The stated objective of the poll was to collect data on preferences to investigate potential improvements to the HBS allocation mechanism.<sup>16</sup>

---

<sup>15</sup>The reason this sums to a bit more than 10 courses per student is the half courses.

<sup>16</sup>The exact text was the following: "Please use the following pull down menus to rank your top 30 most preferred EC courses for 2005-2006, irrespective of whether you were assigned the course or not. Courses are not section-specific. If you have fewer than 30 courses that you would like to rank, please select "Finished Ranking Courses" from the pull-down menu and move on to Question #2. It is critical that the ranking you submit completely reflects your preferences. In particular, do not feel the need to rank courses that fill up quickly first. Alternatively, do not ignore courses just because you perceived that they would be difficult to get. You should rank the courses according to how you actually feel about them." The interface was identical to the interface used for the May poll.

Table 1 summarizes the number of students and courses covered by the data on each occasion. Because participation was compulsory, the May trial run data and the July run data cover the entire population. The small discrepancy in numbers is due to students leaving for or returning from military duty, maternity leave or any other leave of absence. 163 students filled in our poll in January 2006 in a consistent manner.<sup>17</sup> The table also reports the number of courses ranked by the students. For the May trial run and the July run, the submitted rank order lists can be section-specific. When constructing rank order list over courses, we kept the first time a course appeared in the original rank order list.<sup>18</sup>

Table 1: Descriptive statistics – submitted preferences

|                                | May poll | May trial run | July run | Jan poll |
|--------------------------------|----------|---------------|----------|----------|
| # students                     | 460      | 922           | 916      | 163      |
| avg # courses per ROL          | 5        | 22.33         | 21.96    | 17.46    |
| std dev # courses per ROL      | 0        | 5.13          | 4.86     | 7.31     |
| # courses listed at least once | 85       | 92            | 88       | 92       |

## 5 Evidence of Strategic Behavior

In this section, we provide evidence that students understand the strategic incentives of the HBS mechanism and act accordingly. Our analysis provides support for the joint hypothesis that: (1) students’ May poll responses represent their truthful preferences, and (2) students’ July run preferences represent equilibrium behavior.

This joint hypothesis is natural given the context of our data. In the May poll, students were explicitly asked by the administration to state their true preferences, and we are not aware of any compelling reason for them to disobey this request. Submitted preferences in the July run are those used for the initial allocation of courses. We have argued in section 3 that profitable deviations from truthful behavior are easy to detect in the HBS mechanism: they only require the kind of information that the HBS administration provides to the students through the May trial run. In a high-stakes environment with sophisticated and informed players that have had some opportunity to learn, equilibrium play is a natural hypothesis.

Whether this joint hypothesis holds is an empirical question however. Specifically, we are worried about two issues. First, preferences may change over time, idiosyncratically, or as a response

<sup>17</sup>We dropped two students who ranked the same course more than once for a course appearing in their top 5 courses.

<sup>18</sup>This convention affects very few observations. Out of the 20,279 student-course observations in the July run, 14,296 observations are for courses that have multiple sections but for most of them the student listed the different sections of the course in consecutive order. Requests for different sections of the same course were non consecutive for only 282 student-course observations (2%). In our robustness checks we considered alternative conventions for treating those non-consecutive requests.

to new information or social learning, and May poll preferences may no longer represent preferences in July. Second, students may fail to play a best response despite the information and learning opportunities they have had.<sup>19</sup>

We address these questions by first showing that May poll preferences and July run ROLs differ systematically. We then provide evidence, both at the aggregate level and at the individual level, that our joint hypothesis can explain this systematic difference and that alternative accounts cannot.

### 5.1 Evidence based on aggregate data

We consider four distinct reasons why a student might submit different preferences for courses at different points in time: genuine preference change, new information, social learning or strategic consideration. Genuine preference changes not driven by new information or social learning are likely to be idiosyncratic. They should therefore not affect aggregate demand for individual courses over time.<sup>20</sup> This leaves three potential drivers for changes in the aggregate demand for individual courses over time:

1. **New information about courses** (whether good or bad) should lead to a correlated and persistent shock to preferences and thus a persistent shock to the aggregate demand for these courses. These shocks should not, however, be related to the popularity of these courses. We have argued in section that very little new information about courses was provided to students between May and July. However, we cannot rule out this explanation entirely.
2. **Social learning**, i.e. the updating of one's own valuation for a course, through word-of-mouth or the observation of its popularity among students (for example, through the feedback from the May poll or the trial run), should lead to correlated and persistent shocks to individual preferences and thus to aggregate demand. These shocks are likely to be correlated to the initial popularity of the course.
3. From section 3, we know that **strategic behavior** should lead to a correlated but temporary shock to submitted preferences and thus to aggregate demand during the initial allocation. This effect is related to the likely popularity of courses.

---

<sup>19</sup>If our joint hypothesis used the May trial run ROLs instead of the July actual run ROLs, the first issue would be less of a concern, because less time would have elapsed since the May poll, and the second issue would be more of a concern, since the May trial run was an opportunity for students to learn.

The results of this section would be essentially unchanged if we hypothesized that the May trial run preferences represent equilibrium behavior. We focus on the July actual run since that is what mattered for students' welfare, the focus of Sections 6 and 7.

<sup>20</sup>An additional benefit of focusing on aggregate demand is that it also removes noise coming from any small randomness in submitted preferences due to students' "carelessness" or near indifference.

Our joint hypothesis implies that aggregate difference between the May poll and July run preferences is driven by strategic behavior. To distinguish strategic behavior from the effect of new information or social learning, we compare the aggregate demands for courses at three different points in time: May poll, July run and January poll. Consider course  $c$  and a sample of students. Course  $c$ 's distribution of ranks in that sample,  $D_c(r)$ , is the proportion of students in the sample that rank course  $c$  on or before  $r$ . To test for the equality or inequality of aggregate demand for a course across time, we use Gehan (1965)'s extension of the Wilcoxon rank-based test for discrete and censored data (censoring in our data arises from the fact that students only rank 5 courses in the May poll and that some students rank less than 30 courses in the July run). The idea behind this non parametric test is the following. Fix a course, say course  $c$ , and consider two independent samples of students of size  $n_1$  and  $n_2$ . An observation is a student's rank for course  $c$  or, if the student did not rank that course, the rank of the last course s/he ranked, which will be taken as the censoring point for that observation (in words, we do not know how that student ranks course  $c$  but we know that it must be below this censoring point). Pair every observation in the first sample with each observation in the second sample. This creates  $n_1 n_2$  pairs. To each pair, we assign a value of -1 if the observation in the first sample is definitely before the observation in sample 2 (this will be the case if the rank in sample 1 is smaller than the rank in sample 2 or if the observation in sample 2 is censored and the censoring point is higher than the observation in sample 1). Similarly, we assign a value of +1 if the observation in sample 2 is definitely before the observation in sample 1. We assign a value of zero to the pair otherwise. Gehan (1965) shows that the resulting sum over each pair is distributed according to a normal distribution which can be used to test the null hypothesis that  $D_{1c}(r) = D_{2c}(r)$  for  $r \leq R$  where  $D_{1c}$  and  $D_{2c}$  are the distributions of ranks in sample 1 and sample 2 respectively and  $R = \min\{\text{highest censoring point in sample 1, highest censoring point in sample 2}\}$ .

Table 2 reports the results of the Gehan test for the 82 courses that appear in both the May poll and the July run at the 5% significance level. Courses are categorized into low demand courses, medium demand courses and high demand courses, depending on the results from the trial run. Specifically, we categorize a course as high demand if demand in the May trial run was reported to be at least twice the available capacity.<sup>21</sup> A course is said to be low demand, if demand in the May trial run was less than 70% of the available capacity. All other courses are medium demand courses.

---

<sup>21</sup>We also considered alternative definitions of popularity based on the May poll demand rather than the May trial run. The results are similar. The advantage of using the feedback from the May trial as measures of a course's popularity is that these popularity measures are directly available to students (to get a sense of a course's popularity based on the May poll data, a student would have to aggregate the submitted rank order lists and extrapolate for courses at positions 6 or lower).

Table 2: Comparison between May poll demand and July run demand

|                       | N  | July demand lower | No difference | July demand higher |
|-----------------------|----|-------------------|---------------|--------------------|
| Low demand courses    | 25 | 17                | 8             | 0                  |
| Medium demand courses | 37 | 17                | 20            | 0                  |
| High demand courses   | 20 | 1                 | 12            | 7                  |

Table 2 shows that the null hypothesis of unchanged demand between the May poll and the July run is rejected for 42 out of the 82 courses (51%), i.e. submitted preferences in the May poll and in the July run differ significantly. For low demand courses and medium demand courses, rejection occurs because reported demand in July is lower than reported demand in May, whereas the reason for rejection for high demand courses is mostly because demand is higher in July. This is consistent with the prediction from equilibrium behavior that students will have a tendency to overreport popular courses and consequently underreport less popular courses. This is also consistent with social learning, however.

To further validate our interpretation that the discrepancy between the May poll and the July run is driven by strategic behavior and not social learning, we applied the same test on the 44 Winter courses that appear in both the May poll and the January poll. We focus on Winter courses because the experience of Fall courses may have affected students' preferences over them. Table 3 shows that we can reject the null hypothesis of unchanged demand between the May poll and the January poll for 13 courses out of 44 (29.5%). This is a lower rejection rate than in Table 2 despite the fact that more than 8 months separate these two polls whereas only 3 months separate the May poll and the July run.<sup>22</sup> Moreover, unlike in Table 2, there is no systematic pattern in the rejections: the direction of rejections is unrelated to whether a course is high or low demand. This suggests that most of the observed changes in submitted preferences in July are due to (short term) strategic considerations induced by the HBS mechanism, rather than long-lasting social learning.<sup>23</sup>

Table 3: Comparison between May poll demand and January poll demand

|                       | N  | January demand lower | No difference | January demand higher |
|-----------------------|----|----------------------|---------------|-----------------------|
| Low demand courses    | 13 | 1                    | 11            | 1                     |
| Medium demand courses | 23 | 5                    | 16            | 2                     |
| High demand courses   | 8  | 1                    | 4             | 3                     |

<sup>22</sup>Because the demand for a course is not independent of the demand for another course, it is difficult to interpret the levels of rejection beyond the fact that one is lower than the other.

<sup>23</sup>Course-level results confirm this. The null hypothesis of unchanged demand between the May poll and July was rejected for 19 Winter courses. For only 7 of those is the null hypothesis of unchanged demand between the May poll and January poll rejected. In other words, for the 12 other Winter courses, the data rule out persistent shocks to aggregate demand like those that would be generated by new information or social learning.

As another piece of evidence for strategic behavior by students, we check whether the predictions of theorems 3 and 4 are born out in our data. Theorem 3 predicts that in any equilibrium, those courses for which (true) demand exceeds supply will run out during the initial allocation. We use the May poll preferences to construct the set of such courses. Because only 456 students filled in the poll we scale course capacities accordingly. A conservative estimate is that any course whose demand restricted to the top 5 rank exceeds adjusted capacity should belong to the set of courses that run out at equilibrium. Six courses satisfy this definition, and they all run out during the initial allocation. As an alternative we considered that any course whose demand in the poll exceeded 70% of adjusted capacity satisfies the condition of theorem 3. Again, we found that these 13 courses all run out during the initial allocation.

Theorems 3 and 4 also provide predictions on equilibrium run-out times. Because  $m = 10$  and 43 courses run out at equilibrium, Theorem 3 is automatically satisfied in our data. Theorem 4 has more bite. For each course for which  $D_c(5) > 1$  based on the poll data and adjusted capacities, we check whether this course always runs out before the round at which true demand exceeds supply. All 6 such courses satisfied this stronger test.

## 5.2 Evidence based on individual data

We now turn to individual data. We first run the HBS mechanism using the July run submitted preferences and 10,000 randomly drawn priority orders over students to identify popular courses from unpopular courses using definition 2. We then focus on the 456 students who submitted their preferences for the May poll and in July. Out of the 2,280 courses that appear as their 5 most preferred courses according to the poll, 1,744 are popular.

For each of these courses, we check whether lemma 2's necessary conditions for a best response are satisfied in the July run preferences submitted by these students. Specifically, if a popular course is after an unpopular course on a student's ROL, we check whether this student gets it for sure (condition (ii)), and if not, whether moving it up to the position of the first unpopular course on his ROL would secure a positive probability of getting it (condition (iii)). Out of the 1,744 popular course entries, 106 (6.1%) violate the necessary conditions for a best response. These violations involve 90 students (19.7%).

Violations of lemma 2 mean that the relationship between the May poll preferences and the July run preferences does not correspond to the predicted relationship between a student's true preferences and his reported preferences in the equilibrium of the HBS mechanism. In our data, such discrepancy can be explained by genuine preference changes between May and July so that the submitted preferences in the May poll are actually no longer the true preferences of the students in July, or by strategic mistakes during the play of the HBS mechanism. A third possibility is that our test of lemma 2 is based on an incorrect classification between popular and unpopular courses.

We investigate each hypothesis in more detail.

We start with violations that can be explained by the way we coded our data and, in particular, by our ignorance of section-specific factors. Nine violations are caused by multiple-section courses which are recorded in our data as unpopular because they do not fill up all their seats, but which experienced excess demand for at least one of their sections based on the trial run. Thus, some students may have rightfully viewed these courses, or more specifically, a section of one of these courses, as popular, and ranked it highly. Yet, our test counts these as lemma 2 violations. We ignore these in the remainder. We are thus left with 97 violations of lemma 2 (5.6%), involving 81 students (17.8%).

For every popular course that violates lemma 2, we first look at its position in the student's preferences in May and in July. All lemma 2 violations are caused by courses that were strictly downgraded relative to the student's reported preferences in May. Out of the 97 violations, 59 correspond to a course that appears in a student's top 5 courses in May and no longer appears in the submitted preferences in July. These involve 49 students. A likely interpretation of these violations is that those students changed their preferences between May and July.

Another 20 violations can be explained by courses that appeared to be unpopular based on the trial run but happened to be popular in the real run of the mechanism.<sup>24</sup> Students slightly downgraded these courses, thinking they were "safe", and failed to get some of them as a result. Likewise, 10 violations can be explained by courses that seemed to be popular based on the trial run but ended up not being popular in July. Some students needlessly placed these courses ahead of other popular courses. In total, these (slight) strategic mistakes explain 30 violations. They involve 28 students.

Twenty two violations are not explained by either obvious preference changes or strategic mistakes of the kind just described, involving 22 distinct students. Ten of these mistakes involve the same popular course that is the outlier in Table 2; we suspect that students received information about this course that caused a meaningful number of them to downgrade their preference for it between May and July. There is no discernible pattern to the remaining lemma 2 violations.

To conclude, the analysis of the individual data confirms the overall picture from the aggregate data. 82.2% of students submit ROLs that are consistent with the joint hypothesis that May poll preferences correspond to their true preferences in July and that they play a best response in July; 94.4% of popular-course requests are consistent. About 60% of violations can be traced to likely preference changes and about a third can be traced to slight strategic mistakes due to incorrect beliefs about the popularity of courses.

---

<sup>24</sup>Recall that students receive information on course enrollment and oversubscriptions based on one single trial run of the algorithm in May. If we define courses as popular based on the trial run if all their sections reach their capacity (courses are unpopular otherwise), there are five courses that are popular based on the trial run but are no longer popular in the July run, and nine courses that are unpopular based on the trial run and are popular in July.

### 5.3 Constructed Truthful Preferences

We have argued that the May poll preferences can be taken as an approximation of truthful preferences in July. However, May poll preferences have two limitations for the analysis of the welfare properties of the HBS mechanism. First, they are restricted to students' top 5 courses. Second, they are available for only 456 students out of the 916 students who participated in the July run.

To address the first limitation, we construct an extension of students' truthful preferences as follows. We consider that students' truthful top 5 courses correspond to their top 5 courses in the May poll.<sup>25</sup> Other courses are moved down to position 6 and below in a way that preserves their relative ranking in the July ROLs. We call the result "constructed truthful preferences." To illustrate, suppose a student submitted the ROL  $c_1, c_2, c_3, c_4, c_5$  in the May poll but submitted  $c_4, c_3, c_6, c_1, c_2, c_7, c_8$  in the July run. His constructed truthful preference is given by  $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ . Our use of constructed truthful preferences will cause us to underestimate the extent of strategic behavior, because we assume that the relative ranking of courses not in the top 5 is truthful. In the other direction, idiosyncratic preference changes between May and July for May-poll-ranked courses will lead us to overestimate the extent of strategic behavior.

To address the second limitation, we restrict attention to the 456 students for whom we have May poll preferences and adjust course capacities accordingly. For each course, we used the Gehan test to compare the distribution of course ranks in July for the students who did answer the poll to the distribution for those who did not. We found no significant difference between the distributions across these two samples of students, suggesting similar submitted ROL patterns across the two groups. This is consistent with the two groups being replicas of one another, and the equilibrium in the May-poll-only economy being the same as in the original economy.

These constructed preferences and the scaled down course capacities are the main input in our welfare analysis in the next two sections.

## 6 Welfare Consequences of Strategic Play

The purpose of this section is to quantify the welfare consequences of students' strategic play in the HBS mechanism.

We begin by examining the extent to which strategic behavior leads to ex-post Pareto inefficient allocations. Our approach is to formulate an integer program that seeks to execute the maximum number of Pareto-improving trades. We find substantial ex-post inefficiencies: around 15% of course seats can be profitably reallocated in a typical run of the HBS mechanism, involving over three-quarters of students.

We then turn to an ex-ante evaluation (i.e., before priority orders are drawn). In a first step,

---

<sup>25</sup>Courses that were offered in the May poll but were no longer available in the July run are dropped.

we take the perspective of individual students. We compare the distribution of outcomes a student receives under the actual play of the HBS mechanism to his distribution under a non-equilibrium counterfactual in which all students report their preferences truthfully. The challenge we face is that our data consist of students' ordinal preferences over individual courses. Yet, their expected utility depends on lotteries over bundles of courses. We develop a series of comparison results that indicate conditions under which we can say that a particular student prefers her distribution over outcomes from truthful play to that from strategic play. These conditions place restrictions on the mapping from preferences over courses to preferences over bundles, and on the mapping between preferences over sure outcomes and preferences over lotteries. For all the cases we consider, more students are worse off under strategic play.

In a second step, we assess welfare from the perspective of a utilitarian social planner. We develop analogous comparison results that indicate conditions under which we can say with certainty that the social planner prefers one distribution of outcomes to another. For most of the cases we consider, the social planner strictly prefers truthful play, and the magnitudes appear meaningful.

For all results in this and the next section, our economy consists of the 456 students who filled in the May poll, with course capacities scaled accordingly (see section 5.3). We include an aftermarket phase for the HBS mechanism as described in Section 2.2, in which all students play truthfully.

## 6.1 Ex-Post Efficiency Consequences of Strategic Play

We assess the magnitude of ex-post inefficiency in the HBS mechanism as follows. First, we randomly draw a priority order  $\lambda$  and run the HBS mechanism using students' July-run preferences. Then we seek to execute as many Pareto-improving trades amongst the students as possible.

Because our data consist only of ordinal preferences over individual courses, there are some profitable trades that we will not be able to find. For instance, if a student's ROL is  $P_s : c_1, c_2, c_3, c_4$  and his allocation is  $\{c_1, c_4\}$  then we know, from responsiveness, that he is willing to trade  $c_4$  for  $c_2$  or  $c_3$ , but we do not know whether he is willing to trade the bundle  $\{c_1, c_4\}$  for the bundle  $\{c_2, c_3\}$ .

Subject to this caveat of data incompleteness, it is without loss of generality to restrict attention to trades in which each participant gives and receives a single course seat. Whatever many-to-many trades we are able to find can be found using multiple one-for-one trades. For instance, student  $s$  above would be willing to trade  $\{c_2, c_4\}$  for  $\{c_1, c_3\}$ , but this can be executed using two one-for-one trades of  $\{c_2\}$  for  $\{c_1\}$  and  $\{c_4\}$  for  $\{c_3\}$ . We therefore formulate the following binary integer program:

$$\max \sum x_{sc'} \text{ s.t.} \tag{2}$$

$$\sum_s \sum_{c'} x_{sc'} - x_{sc} = 0, \forall c$$

$$\sum_{c'} x_{sc'} + x_{sc} \leq 1, \forall s, c \tag{3}$$

$$\begin{aligned}
x_{sc'c} = 1 &\Rightarrow c \in a_s(\widehat{\mathbf{P}}, \lambda), c' \notin a_s(\widehat{\mathbf{P}}, \lambda), c' P_s c \\
x_{sc'c} &\in \{0, 1\}, \forall s, c, c'
\end{aligned} \tag{4}$$

Variable  $x_{sc'c}$  indicates whether we execute the one-for-one trade in which student  $s$  gives  $c$  and gets  $c'$ . For this trade to be feasible and desirable it must be that student  $s$ 's original allocation includes  $c$ , does not include  $c'$ , and that he prefers  $c'$  to  $c$ ; see (4). The constraints (2) capture the adding-up condition that each course must be given as often as it is received. The constraints (3) prevent a student from trading the same course twice, both to ensure feasibility and to avoid double-counting.

We repeat this exercise for 100 priority orders. The results are as follows:

Table 5. Ex-Post Pareto Improving Trades

|                                    | Mean  | Std. Dev. |
|------------------------------------|-------|-----------|
| # of Executed Trades per Student   | 1.54  | (0.04)    |
| % of Allocated Course Seats Traded | 15.4% | (0.31%)   |
| % of Students Executing            |       |           |
| 0 Trades                           | 16.4% | (1.1%)    |
| 1 Trade                            | 35.4  | (1.7)     |
| 2 Trades                           | 30.5  | (1.6)     |
| 3+ Trades                          | 17.8  | (1.3)     |

Table 5 suggests that the level of ex-post inefficiency under the HBS mechanism is substantial: on average 15% of course seats can be profitably reallocated, involving 84% of students. We were somewhat surprised by the magnitude of this inefficiency and so ran several robustness checks, each of which places additional restrictions on what constitutes a desirable trade (i.e., on (4)). First, we eliminated from consideration any trades in which either course in the trade is requested in a way that violates our characterization of equilibrium best responses (Lemma 2). This reduces the volume of Pareto-improving trade to roughly 12% of all course seats. Next, we restricted attention to trades that are true to the spirit of the equilibrium ex-post inefficiency of Example 3. Specifically, we restrict attention to trades in which the course received is either overreported by the student or is a course that he downgrades but still receives with strictly positive probability (see type  $P_1$  in Example 3). This reduces the volume of Pareto-improving trade to roughly 7% of all course seats. Combining the two tests reduces the level to 6%. This magnitude remains economically meaningful.

Evidence from an informal survey conducted by two HBS students is consistent with our finding that the HBS mechanism is meaningfully inefficient ex-post.<sup>26</sup> By contrast if students played

<sup>26</sup>In Spring 2005 two HBS students surveyed 160 of their classmates for a class project related to the HBS course-

truthfully there would be zero one-for-one Pareto improving trades available.

## 6.2 Ex-Ante Comparisons at the Individual Level

We now turn to the ex-ante comparison between truthful and strategic play of the HBS mechanism. The assumption that preferences are responsive yields the following comparison criterion:

**Comparison Result 1 (Responsive Preferences)** Suppose that student  $s$  has responsive preferences. Then student  $s$  strictly prefers play  $\widehat{\mathbf{P}}$  to play  $\widehat{\mathbf{P}}'$  if the lottery over bundles she receives under  $\widehat{\mathbf{P}}$  first-order stochastically dominates that under  $\widehat{\mathbf{P}}'$  based on the responsiveness partial order.

We implement this criterion for the comparison between truthful play and strategic play of the HBS mechanism as follows. First, we run the HBS algorithm for both truthful play and strategic play for 100 randomly drawn priority orders,  $\lambda_1, \dots, \lambda_{100}$ . Second, for each student  $s$ , we form a bipartite graph where one set of nodes is her outcomes from truthful play,  $a_s(\mathbf{P}, \lambda_1), \dots, a_s(\mathbf{P}, \lambda_{100})$ , and the other set of nodes is her outcomes from strategic play,  $a_s(\widehat{\mathbf{P}}, \lambda_1), \dots, a_s(\widehat{\mathbf{P}}, \lambda_{100})$ . Third, we draw an edge from node  $a_s(\mathbf{P}, \lambda_j)$  to node  $a_s(\widehat{\mathbf{P}}, \lambda_k)$  if we know from  $s$ 's ordinal preferences over courses,  $P_s$ , and the assumption of responsiveness that she weakly prefers  $a_s(\mathbf{P}, \lambda_j)$  to  $a_s(\widehat{\mathbf{P}}, \lambda_k)$ . Fourth, we check if the resulting bipartite graph has a perfect matching, i.e., a subset of edges such that each node in the truthful-play node set is connected to exactly one node in the strategic-play edge set. If there is a perfect match, this means that the set of outcomes under  $\mathbf{P}$  first-order stochastically dominates that under  $\widehat{\mathbf{P}}$  based on the responsiveness partial order. To check if  $\widehat{\mathbf{P}}$  dominates  $\mathbf{P}$  reverse the way the edges are drawn in Step 3.

We find that 45% of students are unambiguously harmed by strategic play, versus just 5.5% that unambiguously benefit. 1% of students are indifferent; for the remaining 48.5% the comparison is indeterminate. Two elements drive the indeterminacy. First, responsiveness yields only a partial order over bundles; mechanically, this means that fewer edges are drawn in the bipartite graph than would be the case if we knew students' complete preferences over bundles. Second, even if we knew students' ordinal preferences over bundles, we still would need to know their preferences over lotteries to always reach a comparison; first-order stochastic dominance is a demanding order. In the rest of this section, we impose additional assumptions on preferences that fill in this information gap and pin down the indeterminate cases.

---

allocation procedure. One of the survey questions was "Did you know of a trade with another student that could have made you both better off?" 58.1% responded yes. This is suggestive of both the magnitude and students' awareness of ex-post Pareto inefficiency. This figure likely includes some double counting; e.g., if many students want to trade A for B, and only one wants to trade B for A, then many more students might know of a Pareto-improving trade than could actually execute them. Of course, it also likely excludes lots of trades that students aren't aware of, including multi-way trades. For many random priority orders we are able to find 43-way trades involving one seat in each of the 43 popular courses.

We say that student  $s$  has *additive* preferences if there exist numbers  $v_s(c)$  for all courses in  $C$ , such that  $u_s(a_s) > u_s(a'_s) \iff \sum_{c \in a_s} v_s(c) > \sum_{c \in a'_s} v_s(c)$  where  $u_s$  is student  $s$ 's vNM utility function and  $a_s$  and  $a'_s$  are allocations. Additive preferences are a special case of responsive preferences. By itself, the additivity assumption does not yield new results, but it provides a structure onto which we can layer additional assumptions. Specifically, if, in addition, student  $s$  is risk neutral, then his expected utility under strategy profile  $\widehat{\mathbf{P}}$  can be expressed as  $\sum_{\lambda} \sum_{c \in a_s(\widehat{\mathbf{P}}, \lambda)} v_s(c)$ . This yields our second comparison result:

**Comparison Result 2 (Additive Preferences).** Suppose that student  $s$  is risk neutral and has additive preferences. Student  $s$  prefers play  $\widehat{\mathbf{P}}$  to play  $\widehat{\mathbf{P}}'$  if, for any  $j$ , the expected number of top- $j$  courses he gets under  $\widehat{\mathbf{P}}$  exceeds that under  $\widehat{\mathbf{P}}'$ . He prefers  $\widehat{\mathbf{P}}'$  if the reverse relationship holds. He is indifferent if he gets each course with equal probability under both strategy profiles.

Note the difference versus Comparison Result 1. The combination of additivity and risk neutrality allows us to aggregate a distribution over  $\binom{C}{m}$  bundles into a distribution over just  $C$  courses.

A special case of additive preferences that allows us to compress this distributional information even further is when the difference in utilities derived from the 1<sup>st</sup> and 2<sup>nd</sup> favorite courses is the same as the difference in utilities derived from the  $n^{\text{th}}$  top course and the  $n - 1^{\text{th}}$  top course, for any  $n$ . Equivalently, a student that has those preferences cares about the *average rank* of the courses in her allocation. Average rank is a measure of mechanism performance emphasized by the HBS administration. Combined with different assumptions on risk attitudes, this yields the following comparison result:

**Comparison Result 3 (Average-rank Preferences).** Assume student  $s$  has average-rank preferences and let  $\bar{r}_s(\widehat{\mathbf{P}}, \lambda)$  denote the average rank of the courses that student  $s$  get under strategy profile  $\widehat{\mathbf{P}}$  for the priority order  $\lambda$ :

- (i) Independently of his attitude towards risk, student  $s$  prefers strategy profile  $\widehat{\mathbf{P}}$  to strategy profile  $\widehat{\mathbf{P}}'$  if  $-\bar{r}_s(\widehat{\mathbf{P}}, \cdot)$  first-order stochastically dominates  $-\bar{r}_s(\widehat{\mathbf{P}}', \cdot)$ .
- (ii) If student  $s$  is risk averse, he prefers strategy profile  $\widehat{\mathbf{P}}$  to strategy profile  $\widehat{\mathbf{P}}'$  if  $-\bar{r}_s(\widehat{\mathbf{P}}, \cdot)$  second-order stochastically dominates  $-\bar{r}_s(\widehat{\mathbf{P}}', \cdot)$ .<sup>27</sup>
- (iii) If student  $s$  is risk neutral, he prefers strategy profile  $\widehat{\mathbf{P}}$  to strategy profile  $\widehat{\mathbf{P}}'$  if  $\sum_{\lambda} \bar{r}_s(\widehat{\mathbf{P}}, \lambda) < \sum_{\lambda} \bar{r}_s(\widehat{\mathbf{P}}', \lambda)$ .

Another special case of additive preferences is lexicographic preferences (defined in Section 3) which puts a high premium on getting top courses. Lexicographic preferences can be seen as the other extreme from average-rank preferences. The HBS administration implicitly assumes

<sup>27</sup>For two cumulative distributions of average ranks, say  $F$  and  $G$ , with ranks distributed on  $[\underline{\rho}, \bar{\rho}]$ ,  $F$  second-order stochastically dominates  $G$  iff  $\int_x^{\bar{\rho}} (1 - F(x)) dx \leq \int_x^{\bar{\rho}} (1 - G(x)) dx$  for all  $x \in [\underline{\rho}, \bar{\rho}]$ . The difference versus the usual formula is due to the fact that lower is better. (Gollier 2001, 3.2)

lexicographic preferences when they evaluate the performance of the mechanism by the number of students who get their top course. Lexicographic preferences also generate a complete order over bundles over courses and we have the following comparison result.

**Comparison Result 4 (Lexicographic Preferences).** Assume student  $s$  has lexicographic preferences. He prefers strategy profile  $\hat{\mathbf{P}}$  to strategy profile  $\hat{\mathbf{P}}'$  if he gets his first choice course more often under  $\hat{\mathbf{P}}$  than under  $\hat{\mathbf{P}}'$  or if he gets each of his  $n$  top choice courses as often under both profiles but gets his  $n + 1^{\text{th}}$  top course more often under  $\hat{\mathbf{P}}'$ , for some  $n$ .

Table 6 uses Comparison Results 2-4 to compare truthful and strategic play of the HBS mechanism.

Table 6. Individual preferences over play of the HBS Mechanism using CR2-4

|                                | Assumption on Preferences |              |        |         |               |
|--------------------------------|---------------------------|--------------|--------|---------|---------------|
|                                | Additive                  | Average-Rank |        |         | Lexicographic |
|                                | Risk                      | Any Risk     | Risk   | Risk    | Risk          |
|                                | Neutral                   | Attitude     | Averse | Neutral | Neutral       |
|                                | (1)                       | (2)          | (3)    | (4)     | (5)           |
| Outcome                        |                           |              |        |         |               |
| Strictly Prefers HBS Truthful  | 46%                       | 56%          | 68%    | 73%     | 90%           |
| Strictly Prefers HBS Strategic | 5%                        | 13%          | 17%    | 26%     | 9%            |
| Indifferent                    | 1%                        | 1%           | 1%     | 1%      | 1%            |
| Indeterminate                  | 47%                       | 30%          | 14%    | 0%      | 0%            |

By each comparison criterion, strategic play harms more students than it benefits.

To understand the role of preference intensity in students' ex-ante evaluations, compare columns (4) and (5). If students have lexicographic preferences, 90% of students are harmed by strategic play, ten times as many as benefit. By contrast, if students have risk-neutral average-rank preferences, only three times as many are harmed as benefit. This contrast is due to a basic asymmetry between the benefits and costs of strategic play. The cost of strategic play is congestion; it is harder to obtain popular courses, holding the rank of a request fixed. The benefit of strategic play is opportunism; students can overreport their preference for popular courses. However, it is impossible to overreport one's *favorite* course. So the benefits of strategic play will be especially small for students with lexicographic preferences.

To understand the role risk preferences play in students' ex-ante evaluations, compare columns (2), (3), and (4). As we put structure on students' risk preferences, we resolve indeterminacies, and these resolutions if anything disproportionately tend to favor strategic play. Certainly there does not seem to be a stark difference in how these two plays expose students to risk.

### 6.3 Ex-Ante Comparisons at the Social Level

We now evaluate social welfare. Clearly, we cannot Pareto rank truthful play and strategic play based on the assumption of responsive preferences alone; some students prefer strategic play and others prefer truthful play. So in this section, we impose additive preferences and assume a utilitarian social planner. An alternative interpretation is that we take the perspective of an individual student who does not know his preferences but knows the distribution of preferences in the population; that is, a student behind a veil of ignorance in the sense of Harsanyi (1953). The "social" analogues of Comparison Results 2-4 are as follows:

**Comparison Result 5 (Additive Preferences).** Assume that students are risk neutral and have additive preferences. Society prefers play  $\hat{\mathbf{P}}$  to play  $\hat{\mathbf{P}}'$ , if, for any  $j$ , the expected number of top- $j$  courses allocated to students under  $\hat{\mathbf{P}}$  exceeds that under  $\hat{\mathbf{P}}'$ . Society prefers  $\hat{\mathbf{P}}'$  if the reverse relationship holds.

**Comparison Result 6 (Average-rank Preferences).** Assume students have average-rank preferences:

- (i) Independently of students' attitude towards risk, society prefers play  $\hat{\mathbf{P}}$  to play  $\hat{\mathbf{P}}'$  if  $-\bar{r}(\hat{\mathbf{P}}, \cdot)$  first-order stochastically dominates  $-\bar{r}(\hat{\mathbf{P}}', \cdot)$ . (the notation  $\bar{r}(\hat{\mathbf{P}}, \cdot)$  indicates that the distribution is taken over priority orders and students)
- (ii) If students are risk averse, society prefers play  $\hat{\mathbf{P}}$  to play  $\hat{\mathbf{P}}'$  if  $-\bar{r}(\hat{\mathbf{P}}, \cdot)$  second-order stochastically dominates  $-\bar{r}(\hat{\mathbf{P}}', \cdot)$ .
- (iii) If students are risk neutral, society prefers play  $\hat{\mathbf{P}}$  to play  $\hat{\mathbf{P}}'$  if  $\sum_{\lambda} \sum_s \bar{r}_s(\hat{\mathbf{P}}, \lambda) < \sum_{\lambda} \sum_s \bar{r}_s(\hat{\mathbf{P}}', \lambda)$ .

**Comparison Result 7 (Lexicographic Preferences).** Assume students have lexicographic preferences. Society prefers strategy profile  $\hat{\mathbf{P}}$  to strategy profile  $\hat{\mathbf{P}}'$  if the expected number of students who get their first choice course is higher under  $\hat{\mathbf{P}}$  than under  $\hat{\mathbf{P}}'$  or if the expected number of students who get their 1<sup>st</sup>, ...,  $n$ <sup>th</sup> top choice courses is the same under both strategy profiles, for some  $n$ , but the expected number of students who get their  $n + 1$ <sup>th</sup> top course is higher under  $\hat{\mathbf{P}}$ .

Figure 1 shows the average number of courses that students get among their top  $n$  choices. There is a first order stochastic dominance relationship between the distribution of outcomes under truthful and strategic play: students get more of their top choices, more of their top two choices and so on under truthful play than under strategic play.<sup>28</sup> Thus CR5 obtains and by consequence

<sup>28</sup>The kink in the HBS Truthful line at rank 6 is a mechanical effect due to the way we construct truthful preferences (see Section 5.3). Students report their top-5 truthful preferences in the May Poll. Their 6th favorite course is the first course they rank in the strategic rank order list that they didn't rank in the May Poll. If this course is rated highly by many other students in the May Poll, then the student will never obtain it under Truthful play, but might obtain it under Strategic play if he ranks it highly enough.

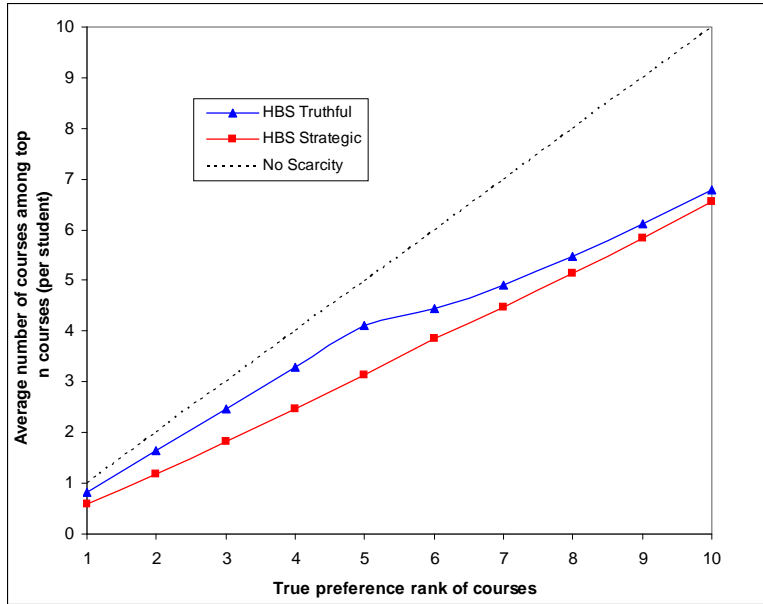


Figure 1: Aggregate outcomes by preference ranks: truthful versus strategic play of the HBS mechanism. The distribution under truthful play first-order stochastically dominates that under strategic play, so CR5 and CR7 obtain.

CR6(iii) and CR7 obtain as well since both are special cases of risk-neutral additive preferences. In other words, if students are risk neutral, a utilitarian social planner unambiguously prefers truthful play of the HBS mechanism. The difference is economically meaningful. 83% of students receive their favorite course under truthful play, and they receive 2.46 of their top three courses, versus 60% and 1.82 under strategic play. What is driving the result is that some of the most popular courses go to students for whom they are not the most preferred courses. For example, the two most popular courses in our data account for 50% of all truthful first choices, and 68% of all strategic first choices. These two courses reach capacity in the first round of strategic play, so, on average, around 26% of the seats in these courses go to students for whom it is not their true first choice.<sup>29</sup>

Next, we compare social welfare when students are not necessarily risk neutral. Figure 2 plots the distribution of the average rank of course allocations in the population over all 10,000 trials. There is a bit more mass at the very best outcomes under strategic play than under truthful play. This is due to the targeted opportunism of some fortunate students who mainly like unpopular courses. Truthful play, on the other hand, delivers more mass in the middle of the distribution. There is no first-order stochastic dominance relationship, but second-order stochastic dominance (CR6(ii)) does obtain. The mean average rank under truthful play is 7.76. The distribution under

<sup>29</sup>That is,  $(68\%-50\%)/68\%$ . These two courses alone account for around 100 fewer students (11% of the student body) obtaining their first-choice course under strategic play.

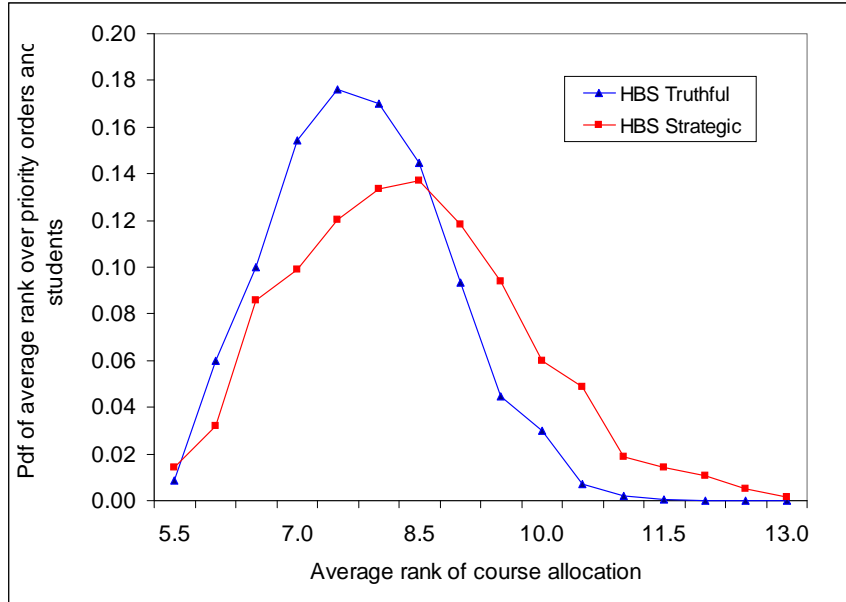


Figure 2: Distribution of the average rank received: truthful versus strategic play of the HBS mechanism. The distribution under truthful play second-order stochastically dominates that under strategic play, so CR6(ii) and CR6(iii) obtain.

strategic play has a higher (worse) mean of 8.35 and has thicker tails.

## 7 Comparison of the HBS Mechanism to a Strategyproof Alternative

In the previous section we showed that strategic play of the HBS mechanism harms efficiency, assessed either ex-ante or ex-post. This section asks the logical next question: should HBS switch to a strategyproof mechanism? To answer this, we perform a welfare comparison between actual play of the HBS mechanism and truthful play of its strategyproof alternative, Random Serial Dictatorship (RSD). We use the same methodology as in Section 6, though here the comparison is to an *equilibrium* counterfactual.

The first thing to note is that RSD is ex-post efficient, whereas we found in Section 6.1 that the the HBS mechanism is highly inefficient ex-post.

In order to assess ex-ante efficiency, we will need to impose additional structure on preferences beyond responsiveness. Under RSD, students will often obtain their ideal bundle of courses, but will also often obtain a very poor bundle. The responsiveness assumption does not rule out the possibility that a student only places value on obtaining his ideal bundle, nor does it rule out that the student only cares about maximizing the minimum bundle he obtains. So comparisons versus

the less-extreme HBS mechanism using CR1 are entirely indeterminate.

As soon as we put additional structure on preferences we find that the HBS mechanism is more attractive ex-ante than RSD. RSD’s ex-ante unattractiveness may be surprising since ex-post it is efficient. Furthermore, RSD’s unattractiveness does not depend on risk aversion; even risk neutral students tend to dislike RSD ex ante. We provide a novel theoretical explanation of RSD’s poor performance at the end of this section.

## 7.1 Ex-Ante Comparisons at the Individual Level

We repeat the methodology of Section 6.3. The following table compares HBS to RSD under additive, average rank, and lexicographic preferences using Comparison Results 2-4:

Table 7. Individual preferences between HBS and RSD: CR2–4

|                                | Assumption on Preferences |              |        |         |               |
|--------------------------------|---------------------------|--------------|--------|---------|---------------|
|                                | Additive                  | Average-Rank |        |         | Lexicographic |
|                                | Risk                      | Any Risk     | Risk   | Risk    | Risk          |
|                                | Neutral                   | Attitude     | Averse | Neutral | Neutral       |
|                                | (1)                       | (2)          | (3)    | (4)     | (5)           |
| Outcome                        |                           |              |        |         |               |
| Strictly Prefers RSD           | 0%                        | 0%           | 0%     | 19%     | 25%           |
| Strictly Prefers HBS Strategic | 26%                       | 2%           | 81%    | 81%     | 75%           |
| Indifferent                    | 0%                        | 0%           | 0%     | 0%      | 0%            |
| Indeterminate                  | 74%                       | 98%          | 19%    | –       | –             |

Begin by examining columns (4) and (5). For either risk-neutral average-rank or lexicographic preferences, the large majority of students prefer the HBS mechanism to RSD. Unlike in the comparison in Table 6, the ratio does not vary severely between the two columns. This suggests that preference intensity is not what drives students’ ex-ante preference for HBS over RSD.

By contrast, consider columns (2), (3), and (4). Without any structure on students’ risk preferences, the comparison is almost entirely indeterminate. This is because RSD induces such extreme outcomes. As soon we put structure on risk preferences, we see that the large majority of indeterminacies are resolved in favor of the HBS mechanism. There is a fundamental difference in riskiness between the two mechanisms.

## 7.2 Ex-Ante Comparisons at the Social Level

We repeat the methodology of Section 6.4. Figure 3 compares the aggregate rank distributions of HBS and RSD. The distribution under strategic play of the HBS mechanism first-order stochasti-

cally dominates that under truthful play of RSD. Thus, a utilitarian social planner prefers HBS to RSD when students are risk neutral and have any additive preferences (CR5, CR6(iii), CR7).

This is surprising. It is obvious that RSD exposes students to more risk than HBS, but here we find that even a risk-neutral society prefers the ex-post inefficient HBS to the ex-post efficient RSD. This suggests that RSD’s ex-post efficiency is not a good proxy for social welfare.

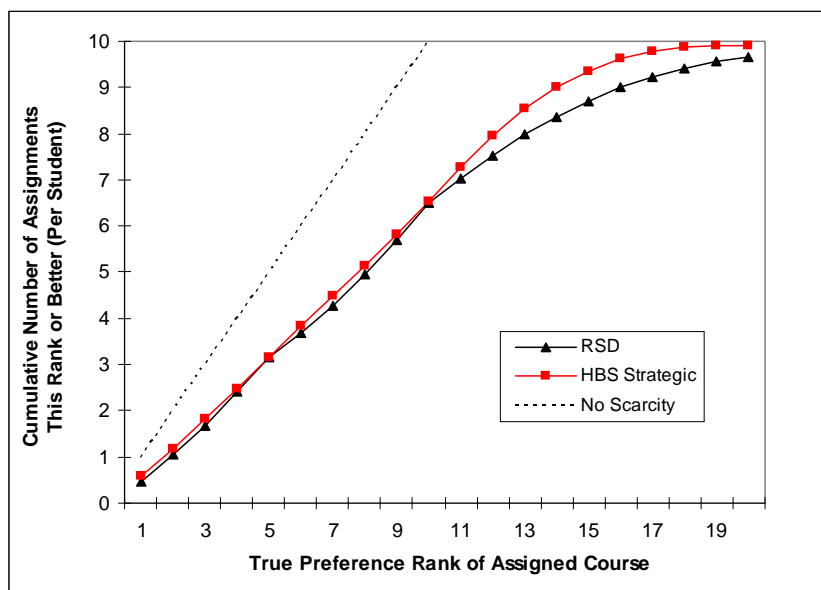


Figure 3: Aggregate outcomes by preference ranks: RSD versus strategic play of the HBS mechanism. The distribution under HBS first-order stochastically dominates that under RSD, so CR5 and CR7 obtain.

The magnitudes are of the most economic importance in the tails. Students receive their favorite course with 60% probability under HBS, but with only 47% probability under RSD.<sup>30</sup> Students actually receive slightly more of their 2nd-10th favorite courses under RSD (6.02) than under HBS (5.95). This is because students with lucky draws in RSD get all 10 of their favorite courses. The cost is that students receive twice as many courses they like less than 15th (1.30) under RSD than under HBS (0.65). As a result the average average rank under RSD is 9.84, versus 8.35 under HBS. This is an economically meaningful difference, and around twice the average rank difference between truthful and strategic play of the HBS mechanism.

We can get a better understanding of the risk to which RSD exposes students by examining the distribution of average ranks; see Figure 4. RSD puts much more weight on the tails of the

<sup>30</sup>To give a sense of the magnitude of this difference, we note Pathak’s (2006) findings in the context of single-unit assignment. He finds that students receive their first-choice object 60.6% of the time under RSD, versus 60.8% in the counterfactual of interest (Bogomolnaia and Moulin’s Probabilistic Serial mechanism; 2001).

distribution, and indeed is second-order stochastically dominated by HBS (CR6 (ii), CR6 (iii)). So a utilitarian social planner prefers HBS to RSD if students are weakly risk averse and have average-rank preferences.

Under RSD, around 29% of students obtain their "bliss bundle" consisting of their 10 favorite courses, versus around 1% under HBS. But over 17% of students obtain a bundle with average rank worse than 12, versus just 1% under HBS.

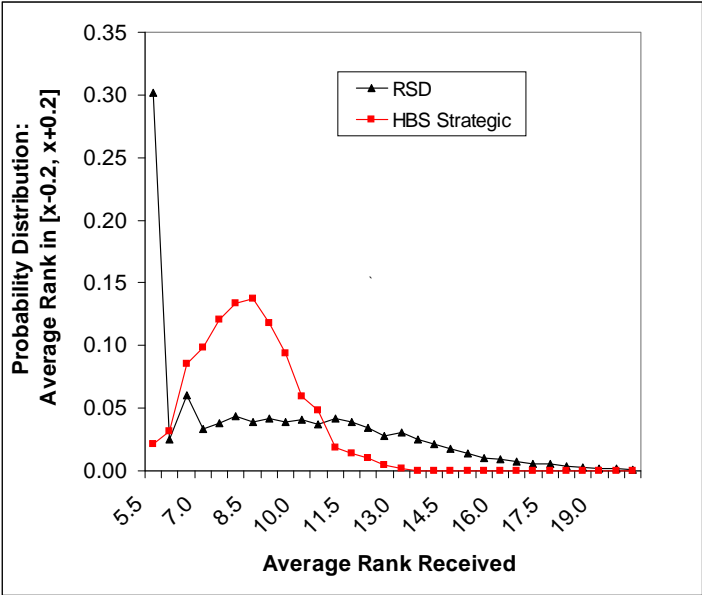


Figure 4: Distribution of the average rank received: RSD vs. strategic play of the HBS mechanism. The distribution under HBS second-order stochastically dominates that under RSD, so CR6(ii) and CR6(iii) obtain.

### 7.3 Explanation: Callousness

Our intuition for RSD’s poor ex-ante performance is simple. Under RSD, fortunate students with good random draws make their *last* choices independently of whether these courses would be some unfortunate students’ *first* choices; students "callously disregard" the preferences of those who choose after them. The reason this callous behavior matters for welfare is that the ex-post utility benefit to the fortunate students from these last choices is small relative to the ex-post harm these choices cause to the unfortunate students. Thus, RSD is unattractive when evaluated ex-ante. Notice that the unattractiveness of RSD does not depend on students being risk averse; even risk neutral agents regard a "win a little, lose a lot" lottery as unappealing.

We formalize this intuition with a simple example and a simple theorem.

**Example 4 (Callousness of RSD).** There are two students,  $A$  and  $B$ , and four courses each in unit supply. Students' ordinal preferences over singletons are drawn uniformly i.i.d., and they report their preferences truthfully. Without loss, assume that student  $A$  has a higher priority than student  $B$ .

Consider first the RSD choosing order  $AABB$ .  $A$  always gets his 1st and 2nd favorite objects, while  $B$  gets either his 1st/2nd, 1st/3rd, 1st/4th, 2nd/3rd, 2nd/4th, or 3rd/4th favorite objects, each with equal probability.  $A$ 's average rank is 1.5 and  $B$ 's is 2.5, so the societal mean is 2.0.  $A$  always gets his favorite course whereas  $B$  gets it with probability 0.5, so the societal mean is 0.75.

Now consider the HBS choosing order  $ABBA$ .  $A$  always gets his 1st favorite object. Then,  $B$  gets his 1st and 2nd favorites with probability one-half, and otherwise gets either his 1st/3rd or 2nd/3rd, each with probability one-quarter. Last,  $A$  gets either his 2nd, 3rd, or 4th, each with equal probability.  $A$ 's average rank is 2.0 and  $B$ 's is 1.875, for a societal mean of 1.9375.  $A$  always gets his favorite course, whereas  $B$  gets it with probability 0.75, so the societal mean is 0.875.

The following table summarizes.

Table 8. Results of Example 4

|     | Average-Rank Received |       |               | P(Get Favorite Course) |     |               |
|-----|-----------------------|-------|---------------|------------------------|-----|---------------|
|     | $A$                   | $B$   | Societal Mean | $A$                    | $B$ | Societal Mean |
| RSD | 1.5                   | 2.5   | 2.0           | 1                      | .5  | .75           |
| HBS | 2.0                   | 1.875 | 1.9375        | 1                      | .75 | .875          |

In this simple example, ex-ante welfare is higher under HBS than RSD for risk-neutral students with either average-rank or lexicographic preferences. The driving force behind both results is that it is harmful, in terms of these measures of welfare, to give  $A$  his second choice before  $B$  has made his first choice.

Simulations suggest that the average-rank finding in Example 4 generalizes to larger economies. For instance, in an HBS-sized version of Example 4 with 1000 students, 100 courses, 100 seats per course, and  $m = 10$ , the average rank under HBS is 5.72 versus 6.45 under RSD.<sup>31</sup>

The following simple theorem shows that the first-choice-course finding in Example 4 generalizes.

**Theorem 6 (Callousness of RSD):** Suppose there are  $S$  students, each of whom requires  $m$  courses, and  $mS$  courses each in unit supply. Students' ordinal preferences over courses are drawn uniformly i.i.d., and they report their preferences truthfully. Then the expected proportion of students who obtain their first-choice object is  $1 - \frac{(S-1)}{2Sm}$  under HBS which is strictly greater than the proportion  $1 - \frac{(S-1)}{2S}$  under RSD whenever  $m > 1$ . As  $S \rightarrow \infty$  the proportion converges to

<sup>31</sup>Further simulation results are available from the authors. We also are able to show theoretically that Example 4 generalizes to any number of students  $S$ , with  $m = 2$  and  $Sm$  courses each in unit supply. The proof is somewhat involved and is available upon request.

$1 - \frac{1}{2m}$  under HBS versus  $\frac{1}{2}$  under RSD.

First, note that Callousness is specific to multi-unit assignment. If  $m = 1$  then the two mechanisms are equivalent. Second, note that the HBS-sized simulation evidence and Theorem 6 illustrate that the Callousness of RSD in multi-unit assignment persists in large markets. This helps to illustrate that Callousness is distinct from Bogomolnaia and Moulin's (2001) critique of RSD in the single-unit assignment setting, since the magnitude of the inefficiency BM address goes to zero as the market grows large (Che and Kojima, 2009).

## 8 Conclusion

This paper uses theory and unusual data to study two specific multi-unit assignment mechanisms: one used in practice to allocate courses to students at Harvard Business School, and the other extensively studied in theory. We show that the HBS mechanism is importantly flawed: it is simple to manipulate in theory and heavily manipulated in practice, with meaningful welfare consequences. However, we also show that it is attractive relative to the strategyproof Random Serial Dictatorship, due to an aspect of RSD that we call "callousness". In this conclusion, we discuss the relevance of our analysis of these two specific mechanisms to the broader literature on market design.

The first contribution concerns the role of strategyproofness. Strategyproofness has traditionally been seen as a desideratum in market design for at least three reasons. First, strategyproof mechanisms are the ultimate robust mechanisms in the sense of Wilson (1987). A second and related reason is that strategyproof mechanisms make it easy to advise market participants and help to level the playing field between sophisticated and naïve players (Abdulkadiroglu et al (2009), Pathak and Sonmez (2008)). A third reason to favor strategyproof mechanisms is that they generate preference information that can be used for ex-post policy evaluation and public decisions. These arguments have been especially prevalent in the context of assignment and matching problems (Roth, 2008), and each is compelling in the context of course allocation.

Our data allow us to not only *directly document strategic behavior* but also to *quantify its effects on welfare*. We show that strategic behavior harms most individual students and has a sizable negative effect on overall social welfare. This finding is consistent with the emphasis in prior literature on strategyproofness as a desideratum. However, our data also allow us to simulate a strategyproof counterfactual, and we find that RSD is even worse. That is, the costs of manipulability are large, but the costs of requiring strict strategyproofness are larger. Our paper sounds a cautionary note against imposing strategyproofness as a strict design requirement.

The second contribution concerns efficiency analysis for random allocation mechanisms. Researchers have long acknowledged that ex-ante efficiency is the most compelling efficiency criterion for evaluating random allocation mechanisms, and it is strictly stronger than ex-post efficiency in

the sense that a necessary but not sufficient condition for a lottery over allocations to be ex-ante Pareto efficient is that all realizations of the lottery are ex-post Pareto efficient. But, the impossibility theorems for ex-ante efficiency are even more severe than they are for ex-post (Zhou, 1990), and empirical researchers often have only ordinal preference information, which makes it difficult to analyze welfare ex-ante.<sup>32</sup> As a result, the literature on random allocation mechanisms has largely focused on ex-post efficiency.<sup>33</sup>

We develop a series of comparison results that enable us to compare HBS and RSD ex-ante, from the perspective of both individual students and a utilitarian social planner. These ex-ante comparison results place restrictions on the mapping from preferences over individual objects to preferences over bundles, and on the mapping from preferences over sure outcomes to preferences over lotteries.

Our comparison results suggest that the ex-post inefficient HBS mechanism is more attractive ex-ante than the ex-post efficient RSD. That is, *ex-post efficiency need not even proxy for ex-ante efficiency*. We then identify a new theoretical reason for the ex-ante unattractiveness of RSD when agents require multiple objects, unrelated to risk preferences or to the source of inefficiency discussed in Bogomolnaia and Moulin (2001). In RSD, lucky students make their last choices independently of whether these courses would be some unlucky students' first choices. This "callousness" induces a big wedge between RSD's ex-post efficiency and its ex-ante performance.<sup>34</sup> Overall, our analysis sounds a cautionary note against advocating for a mechanism on the basis of its ex-post efficiency properties alone,<sup>35</sup> if ex-ante efficiency is indeed what we care about.

A final contribution of our paper is that it suggests aspects of "where to look" for more attractive solutions to the multi-unit assignment problem. First, one should seek a mechanism that is not strategyproof like a dictatorship, but that is not simple to manipulate like the HBS mechanism. The mechanism should aim to induce truthful reporting in realistic market environments. Second, one should seek a mechanism that more resembles HBS than RSD in its ex-post fairness and ex-ante

---

<sup>32</sup>For example, hospitals in the medical match report only ordinal preferences over individual doctors, but their welfare depends in the end on the team they assemble. See Roth and Peranson (1999). An exception is unit-demand assignment problems. See Pathak (2006), Abdulkadiroglu, Che and Yasuda (2009), Miralles (2009) and Featherstone and Niederle (2008).

<sup>33</sup>Again, the exception is unit-demand assignment problems. Hylland and Zeckhauser (1979) and Bogomolnaia and Moulin (2001) propose single-unit assignment mechanisms that are, respectively, ex-ante efficient and ordinally efficient.

<sup>34</sup>By contrast, in single unit environments, Che and Kojima (forthcoming) have shown that, when there is a continuum of agents as in our model, RSD is ordinally efficient. This is the best that can be achieved by an ordinal mechanism.

<sup>35</sup>For example, consider Ehlers and Klaus's (2003) argument that dictatorships are an attractive solution to the multi-unit assignment problem: "[Dictatorships] are efficient, strategyproof, and satisfy other appealing properties discussed below. They can be considered to be 'fair' if the ordering of the agents is fairly determined; for instance by queuing, seniority, or randomization."

efficiency characteristics. Participants' realized resources (here, their choosing times) should not be highly unequal as in RSD, but rather roughly equal as in HBS.

One such new mechanism, Budish (2009), is directly inspired by the present analysis. He proposes a mechanism that is strategyproof in a certain limit economy (unlike HBS), satisfies two new criteria of outcome fairness (like HBS and unlike RSD), and exposes participants to an arbitrarily small amount of resource variance (unlike RSD).

## 9 Appendix

### 9.1 Proof of Theorem 1 (simple manipulations)

The strategy of the proof is to show that a sequence of deviations from  $P_s$ , that consist in downgrading the  $\widehat{\mathbf{P}}_{-s}$ -unpopular courses to the bottom half of the top  $m$  courses in student  $s$ 's ROL while preserving the relative ordering of the  $\widehat{\mathbf{P}}_{-s}$ -popular and  $\widehat{\mathbf{P}}_{-s}$ -unpopular courses, leaves student  $s$  weakly better off for all  $\lambda$ . For ease of reference, relabel courses such that  $P_s : c_1, c_2, c_3, \dots, c_C$ . To save on notations, in the remainder, we simply refer to popular and unpopular courses for  $\widehat{\mathbf{P}}_{-s}$ -popular and  $\widehat{\mathbf{P}}_{-s}$ -unpopular courses respectively.

**Claim 1:** Suppose  $c_k$  is unpopular. Then, for all  $\lambda$ ,  $\widehat{P}_s^{c_k \downarrow l}$  gets exactly the same courses as  $\widehat{P}_s$  or exactly one more course in  $\{c_{k+1}, \dots, c_l\}$  than  $\widehat{P}_s$ , at the cost of a course in  $\{c_{l+1}, \dots, c_C\}$ .

**Proof of claim 1:** Fix an arbitrary  $\lambda$ . Because  $\widehat{P}_s$  and  $\widehat{P}_s^{c_k \downarrow l}$  only differ from position  $k$  onwards, the game proceeds identically until the time at which  $\widehat{P}_s$  requests  $c_k$  (and  $\widehat{P}_s^{c_k \downarrow l}$  requests course  $c_{k+1}$ ). Let  $r_k$  be the round at which this happens. By construction,  $c_k$  is available when  $\widehat{P}_s$  requests it in round  $r_k$ . Because student  $s$  has zero mass, the fact that his outcome in round  $r_k$  is different across the two strategies does not affect course seat availabilities and thus, a fortiori, the allocation and requests of other students in any given round.

From round  $r_k + 1$  onwards, student  $s$  requests courses one round earlier under strategy  $\widehat{P}_s^{c_k \downarrow l}$  than under strategy  $\widehat{P}_s$ , until we either reach a course, say  $c_{k'}$ , in  $\{c_{k+1}, \dots, c_l\}$  that student  $s$  gets under  $\widehat{P}_s^{c_k \downarrow l}$  but not under  $\widehat{P}_s$ , or reach position  $l$  in student  $s$ 's ROL. We consider each case in turn:

1. There exists  $c_{k'}$  in  $\{c_{k+1}, \dots, c_l\}$  that student  $s$  gets under  $\widehat{P}_s^{c_k \downarrow l}$  but not under  $\widehat{P}_s$ .

Let  $r_{k'}$  be the round at which student  $s$  requests but does not get this course under  $\widehat{P}_s$ . From round  $r_{k'}$  onwards, student  $s$ 's requests are in synch under both strategies and thus he gets the same outcome until the algorithm reaches position  $l$  in his ROL.

When the algorithm reaches the request in position  $l$ , student  $s$  requests (and gets) course  $c_k$  under  $\widehat{P}_s^{c_k \downarrow l}$ . From then on, student  $s$  requests courses one round earlier under  $\widehat{P}_s$ . This has two possible consequences: either there exists a course that he gets under  $\widehat{P}_s$  but not under  $\widehat{P}_s^{c_k \downarrow l}$  (after which his requests are in synch and there is no more discrepancy between the two outcomes), or the algorithm reaches round  $m$  (and thus the course that the student requests in round  $m$  under  $\widehat{P}_s$  is never requested by  $\widehat{P}_s^{c_k \downarrow l}$ ). In both cases, there is a single course in  $\{c_{l+1}, \dots, c_C\}$  that student  $s$  gets under  $\widehat{P}_s$  instead of  $c_{k'}$  that he does not get under  $\widehat{P}_s^{c_k \downarrow l}$ .

2. The algorithm reaches position  $l$  in student  $s$ 's ROL without any difference in allocations between the two strategies

At that round,  $\widehat{P}_s^{c_k \downarrow}$  requests  $c_k$  and student  $s$ 's requests become in synch again. There is thus no more difference in outcomes

**Claim 2:** Let  $c_k$  be the lowest-ranked unpopular course among the top  $m$  courses in  $P_s$ . Let  $\widehat{P}_s^1 = P_s^{c_k \downarrow m}$ . Student  $s$  is weakly better off using  $\widehat{P}_s^1$  than  $P_s$  for all  $\lambda$ .

**Proof of claim 2:** By claim 1,  $\widehat{P}_s^1$  gets exactly the same courses or exactly one additional course in  $\{c_{k+1}, \dots, c_m\}$  than  $P_s$ , at the cost of a course in  $\{c_{m+1}, \dots, c_C\}$ . Because all courses in  $\{c_{k+1}, \dots, c_m\}$  are strictly preferred to courses in  $\{c_{m+1}, \dots, c_C\}$ , student  $s$  is either indifferent or strictly better off using  $\widehat{P}_s^1$  (here we are using the fact that preferences are responsive and that students have vNM preferences over lotteries).

**Claim 3:** Let  $c_j$  be the  $n^{\text{th}}$  lowest unpopular courses among the top  $m$  courses in  $P_s$ . Let  $\widehat{P}_s^n = \widehat{P}_s^{n-1} c_j \downarrow^{m-n+1}$  (student  $s$  downgrades course  $c_j$  just above all the other less preferred unpopular courses that he has already downgraded). Student  $s$  is weakly better off using  $\widehat{P}_s^n$  than  $\widehat{P}_s^{n-1}$  for all  $\lambda$ .

**Proof of claim 3:** By claim 1,  $\widehat{P}_s^n$  gets either exactly the same courses or exactly one additional course among the popular courses that were between  $c_j$  and position  $m-n+1$  in  $\widehat{P}_s^{n-1}$ . This comes at the expense of a course in  $\{c_{m+1}, \dots, c_C\}$ . Given that preferences are responsive and take the vNM form, student  $s$  is weakly better off using  $\widehat{P}_s^n$  over  $\widehat{P}_s^{n-1}$ .

We continue until there is no further unpopular course to downgrade. At each deviation, student  $s$  is weakly better off for all  $\lambda$ . The claim then follows by transitivity. QED

## 9.2 Proof of Theorem 2 (truthful play in equilibrium)

**Proof:** Identical preferences: Let  $P_s = P_{s'} : c_1, c_2, c_3, \dots$ . Under truthful play, course  $c_1$  runs out earlier than  $c_2$ , which itself runs out earlier than  $c_3$  and so on. Also note that  $c_1$  runs out with probability 1 in round 1 for all strategy profiles  $(\widehat{P}_s, \mathbf{P}_{-s})$  for any  $\widehat{P}_s$ .

Towards a contradiction, suppose  $\widehat{P}_s \neq P_s$  constitutes a profitable deviation for student  $s$  when the other students play  $\mathbf{P}_{-s}$ . Let  $\widehat{P}_s^{c \uparrow}$  equal  $\widehat{P}_s$  except that  $c$  is moved to the first position. Similarly, let  $\widehat{P}_s^{c' \uparrow}$  equal  $\widehat{P}_s$ , except that  $c$  is moved to the first position and  $c'$  is moved to the second position, and so on for  $\widehat{P}_s^{c'c'' \uparrow}, \widehat{P}_s^{c'c''c''' \uparrow}, \dots$

We show that the sequence  $\widehat{P}_s^{c_1 \uparrow}, \widehat{P}_s^{c_1 c_2 \uparrow}, \dots, \widehat{P}_s^{c_1 \dots c_{C-1} \uparrow} = P_s$  constitutes a chain of profitable deviations so that student  $s$  is at least as well off under  $P_s$  as under  $\widehat{P}_s$ . This contradicts the hypothesis that  $\widehat{P}_s$  was a profitable deviation from  $P_s$ .

Consider first  $\widehat{P}_s^{c_1 \uparrow}$  and  $\widehat{P}_s$  and suppose  $c_1$  is not first in  $\widehat{P}_s$  (otherwise  $\widehat{P}_s^{c_1 \uparrow} = \widehat{P}_s$  and we are done).

**Claim 1:** Student  $s$  gets either exactly the same courses under  $(\widehat{P}_s^{c_1 \uparrow}, \mathbf{P}_{-s})$  and  $(\widehat{P}_s, \mathbf{P}_{-s})$  or his two allocations differ by exactly one course: he gets  $c_1$  under  $(\widehat{P}_s^{c_1 \uparrow}, \mathbf{P}_{-s})$  which he does not get under  $(\widehat{P}_s, \mathbf{P}_{-s})$ , in exchange for getting a course under  $(\widehat{P}_s, \mathbf{P}_{-s})$  that he does not get under  $(\widehat{P}_s^{c_1 \uparrow}, \mathbf{P}_{-s})$ .

**Proof of claim 1:** We compare how the game plays out under the two strategies. Partition the set of priority orders  $\mathcal{L}$  into  $\mathcal{L}_1$  and  $\mathcal{L}_0$  according to whether student  $s$  gets  $c_1$  in the first round when playing  $\widehat{P}_s^{c_1\uparrow}$ . Under all priority orders in  $\mathcal{L}_0$  the two games play out exactly in the same fashion (since student  $s$  never gets  $c_1$  under  $\widehat{P}_s$ ), so we focus on priority orders in  $\mathcal{L}_1$ .

Fix  $\lambda \in \mathcal{L}_1$ . Under  $(\widehat{P}_s^{c_1\uparrow}, \mathbf{P}_{-s})$ , student  $s$  gets  $c_1$  which he does not get under the original strategy. From round 1 onwards until we reach a course that student  $s$  gets under one strategy but not under the other, student  $s$  requests each specific course exactly one round later under  $\widehat{P}_s^{c_1\uparrow}$ . Because of the continuum assumption, other students' requests and outcomes are otherwise not affected and courses run out at the same time under both strategy profiles. Thus if there is a course that student  $s$  gets under one strategy but not under the other it is a course that he does not get under  $(\widehat{P}_s^{c_1\uparrow}, \mathbf{P}_{-s})$ . Call this course  $c_l$  and let  $r$  be the round at which this happens. From round  $r$ , student  $s$ 's requests are in synch again and so are other students' requests. This implies there are no additional discrepancies between the two allocations.

Claim 1, together with responsiveness  $(c_1 P_s c_l)$  and vNM preferences over uncertain outcomes, implies that student  $s$  is strictly better off playing  $\widehat{P}_s^{c_1\uparrow}$  than  $\widehat{P}_s$ . We next show that  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$  is preferred to  $\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}$ .

**Claim 2:** Student  $s$  gets either exactly the same courses under  $(\widehat{P}_s^{c_1 \dots c_k \uparrow}, \mathbf{P}_{-s})$  and  $(\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}, \mathbf{P}_{-s})$  or his two allocations differ by exactly one course: he gets  $c_k$  under  $(\widehat{P}_s^{c_1 \dots c_k \uparrow}, \mathbf{P}_{-s})$  which he does not get under  $(\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}, \mathbf{P}_{-s})$ , in exchange for getting  $c_l$ ,  $l > k$  under  $(\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}, \mathbf{P}_{-s})$  that he does not get under  $(\widehat{P}_s^{c_1 \dots c_k \uparrow}, \mathbf{P}_{-s})$  for  $k \geq 2$

**Proof of claim 2:** The proof proceeds along similar lines as the proof of claim 1. Without loss of generality, assume that  $\widehat{P}_s^{c_1 \dots c_k \uparrow} \neq \widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}$ . Until student  $s$  requests  $c_k$  under  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$ , the two games proceed identically. Partition the set of priority orders into  $\mathcal{L}_{11}$  ( $s$  gets  $c_k$  under both strategies),  $\mathcal{L}_{10}$  ( $s$  gets  $c_k$  only under  $(\widehat{P}_s^{c_1 \dots c_k \uparrow}, \mathbf{P}_{-s})$ ) and  $\mathcal{L}_{00}$  ( $s$  does not get  $c_k$  under either strategies). Clearly, for priority orders in  $\mathcal{L}_{00}$ , the two games proceed identically and  $s$  gets the same final allocation.

We claim that  $s$  gets also the same final allocation for priority orders  $\mathcal{L}_{11}$ . To show this, fix  $\lambda$  and let  $r$  be the round at which student  $s$  requests  $c_k$  under  $\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}$  and  $r' < r$  the round at which he requests  $c_k$  under  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$ . Because  $c_k$  fills up earlier than  $c_l$  for  $l > k$ , it means that all courses requested by student  $s$  between  $c_{k-1}$  and  $c_k$  under  $\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}$  are still available at the time of student  $s$ 's turn in round  $r$  under the alternative strategy  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$ . Thus, by round  $r$  student  $s$  has the same allocation under both strategies. Because requests are identical across the two games from then on, so are allocations.

Finally, we argue that, under priority orders in  $\mathcal{L}_{10}$ , student  $s$  gets  $c_k$  under  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$  at the cost of  $c_l$  for some  $l > k$ . The argument here is identical to the argument in the proof of claim 1. There exists a course  $c_l$  that student  $s$  does not get under  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$ . From the time of this unsuccessful

request, student  $s$ 's requests are identical across the two strategies. Thus so are his outcomes.

Claim 2, responsiveness and the assumption of vNM preferences over uncertain outcomes implies that student  $s$  prefers  $\widehat{P}_s^{c_1 \dots c_k \uparrow}$  to  $\widehat{P}_s^{c_1 \dots c_{k-1} \uparrow}$ . Theorem 2(2) then follows from transitivity.

Independent preferences: Let  $\bar{r}$  be such that  $D_c(\bar{r} - 1) < 1$  and  $D_c(\bar{r}) \geq 1$  for course  $\mathbf{P}$ -popular course  $c$ . Under truthful play, all  $\mathbf{P}$ -popular courses run out exactly in round  $\bar{r}$ . This also holds for all strategy profiles  $(\widehat{P}_s, \mathbf{P}_{-s})$  for any  $\widehat{P}_s$ . Truthful play guarantees that each student gets his top  $\bar{r} - 1$  courses. Moreover, it maximizes the chance that he gets  $\bar{r}$   $\mathbf{P}$ -popular courses, and conditional on getting  $\bar{r}$   $\mathbf{P}$ -popular courses, the probability distribution it generates on those  $\bar{r}$ -course bundles first order stochastically dominates the outcome from any alternative (here we use the assumption of responsiveness and the fact that all  $\bar{r}$ -course bundles differ by a single course, the one requested in round  $\bar{r}$ , to generate an order over them). Finally, the fact that  $\mathbf{P}$ -unpopular courses are listed in order of preferences ensures that he gets his  $m - \bar{r}$  (or  $m + 1 - \bar{r}$ ) most preferred courses among them, with no need to resort to the add-drop phase. The claim then follows from responsiveness and the assumption of vNM preferences over uncertain outcomes. QED

### 9.3 Proof of Lemma 2 (best response characterization)

Towards a contradiction, suppose that student  $s$ 's best response  $\widehat{P}_s$  involves ranking  $\widehat{\mathbf{P}}$ -popular course  $c$  lower than a  $\widehat{\mathbf{P}}$ -unpopular course despite the fact that (i)  $r_s(c) \leq m$ , (ii)  $\Pr(c \in a_s(\widehat{\mathbf{P}})) < 1$  and (iii)  $\Pr(c \in a_s(\widehat{\mathbf{P}}^{c \uparrow k})) > 0$  where  $k$  is the position of the first  $\widehat{\mathbf{P}}$ -unpopular course on  $\widehat{P}_s$ .

Let  $\tilde{c}$  denote the last  $\widehat{\mathbf{P}}$ -unpopular course to appear before  $c$  on  $\widehat{P}_s$ . Moreover, let  $c'$  denote the lowest ranked  $\widehat{\mathbf{P}}$ -popular course in  $\widehat{P}_s$  ( $c'$  could be  $c$ ) and let  $l$  and  $l'$  be the position of  $c$  and  $c'$  in  $\widehat{P}_s$ .

We construct an alternative strategy,  $\widetilde{P}_s$ , by making two changes relative to  $\widehat{P}_s$ . First, switch the positions of  $c$  and  $\tilde{c}$ . Second, downgrade  $\tilde{c}$  further down to position  $l'$  if there is a popular course below  $c$  in  $\widehat{P}_s$ . Thus, the two strategies can be written as:

$$\begin{array}{ccccccc}
 & & & l & & l' & \\
 \hline
 \text{If } c \neq c': & & & & & & \\
 \widehat{P}_s : & \dots & \tilde{c} & \dots & c & \dots & c' \dots \\
 \widetilde{P}_s : & \dots & c & \dots & \dots & c' & \tilde{c} \dots \\
 \hline
 \text{If } c = c': & & & & & & \\
 \widehat{P}_s : & \dots & \tilde{c} & \dots & c & \dots & \\
 \widetilde{P}_s : & \dots & c & \dots & \tilde{c} & \dots & 
 \end{array} \tag{5}$$

where the dots denote courses that do not change relative positions between  $\widehat{P}_s$  and  $\widetilde{P}_s$ . Partition

the set of priority orders into three:

$$\begin{aligned}\mathcal{L}_1 &= \{\lambda | c \in a_s(\widehat{\mathbf{P}}, \lambda) \text{ and } c \in a_s((\widetilde{P}_s, \widehat{\mathbf{P}}_{-s}), \lambda)\} \\ \mathcal{L}_2 &= \{\lambda | c \notin a_s(\widehat{\mathbf{P}}, \lambda) \text{ and } c \notin a_s((\widetilde{P}_s, \widehat{\mathbf{P}}_{-s}), \lambda)\} \\ \mathcal{L}_3 &= \{\lambda | c \notin a_s(\widehat{\mathbf{P}}, \lambda) \text{ and } c \in a_s((\widetilde{P}_s, \widehat{\mathbf{P}}_{-s}), \lambda)\}\end{aligned}$$

Note that we do not need to consider the case where  $c \in a_s(\widehat{\mathbf{P}}, \lambda)$  and  $c \notin a_s((\widetilde{P}_s, \widehat{\mathbf{P}}_{-s}), \lambda)$  because this outcome is impossible: for all  $\lambda$ ,  $\widetilde{P}_s$  requests  $c$  strictly earlier than the original strategy.

**Claim 1:** Student  $s$  is weakly better off using  $\widetilde{P}_s$  if  $\lambda \in \mathcal{L}_1$ .

**Proof of claim 1:** Because  $\widetilde{P}_s$  gets  $c$  when it requests it, and  $\widehat{P}_s$  gets  $\tilde{c}$  in that round, the two strategies are in synch and obtain identical outcomes until we reach position  $l$  in student  $s$ 's ROL. If  $c \neq \tilde{c}$   $\widehat{P}_s$  requests  $c$  at that round and, by hypothesis, gets it, while the alternative strategy requests the next course on the list. In other words,  $\widetilde{P}_s$  requests courses early from that round onwards until we reach position  $l'$  on the ROL or a course that  $\widetilde{P}_s$  gets but that  $\widehat{P}_s$  does not get. If we reach position  $l'$  first,  $\widetilde{P}_s$  gets  $\tilde{c}$  and requests are again back in synch. In that case, there is no difference in outcomes between the two strategies. Otherwise,  $\widetilde{P}_s$  gets a course that  $\widehat{P}_s$  does not get (it could be a popular course or the unpopular course requested by  $\widetilde{P}_s$  in round  $m$  and never requested under the original strategy because we never reach position  $l'$ ). If  $c = \tilde{c}$ ,  $\widehat{P}_s$  requests and gets  $c$  and  $\widetilde{P}_s$  requests and gets  $\tilde{c}$  when we reach position  $l$  in the ROL. Requests and outcomes are identical from then on.

To summarize, during the initial allocation,  $\widetilde{P}_s$  gets every popular course that  $\widehat{P}_s$  receives, plus possibly one additional popular course ranked lower than  $c$  on  $\widehat{P}_s$  (at the expense of an unpopular course). Because unpopular courses are available with probability one during the add-drop phase, the final schedule that  $s$  is able to form using  $\widetilde{P}_s$  is at least weakly preferred to that from using  $\widehat{P}_s$ .

**Claim 2:** Student  $s$  is weakly better off using  $\widetilde{P}_s$  if  $\lambda \in \mathcal{L}_2$

**Proof of claim 2:** By a similar argument as above, we can again show that  $\widetilde{P}_s$  gets every popular course that  $\widehat{P}_s$  receives, plus possibly one additional. This additional course might be any popular course ranked lower than  $\tilde{c}$  on  $\widehat{P}_s$ . Because of the add-drop phase,  $\widetilde{P}_s$  yields a weakly better outcome than does  $\widehat{P}_s$ .

**Claim 3:** Student  $s$  is strictly better off using  $\widetilde{P}_s$  if  $\lambda \in \mathcal{L}_3$ .

**Proof of claim 3:** Again we distinguish two cases depending on whether  $c = \tilde{c}$ . If  $c \neq \tilde{c}$ , the two strategies are essentially in synch until we reach their requests for  $\tilde{c}$ : indeed, except for the fact that  $\widehat{P}_s$  gets  $\tilde{c}$  when  $\widetilde{P}_s$  gets  $c$ , they request get the same courses in exactly the same round. Once we reach the (successful) request for  $\tilde{c}$  by  $\widetilde{P}_s$ , they become out of synch and  $\widehat{P}_s$  requests courses one round earlier. Because, by construction, all these courses are unpopular and thus available with probability one, the end result is that the allocations from the two strategies differ by exactly one course:  $\widehat{P}_s$  gets popular course  $c$  that  $\widetilde{P}_s$  does not get, at the cost of an unpopular course. Since

$r_s(c) \leq m$ , and unpopular courses are available with probability one in the add-drop phase, the final schedule that  $s$  is able to form using  $\tilde{P}_s$  is strictly preferred to that from using  $\hat{P}_s$ . If  $c = c'$ , the two strategies become out of synch when we reach position  $l$  in the ROL. By assumption,  $\hat{P}_s$  is unsuccessful at getting  $c$  and  $\tilde{P}_s$  is successful at getting  $\tilde{c}$ . From then on,  $\hat{P}_s$  requests courses one round earlier than  $\tilde{P}_s$ . Because, by construction, these are unpopular courses, the end result is that the allocations from the initial phase differ by a single course:  $\hat{P}_s$  gets  $c$  at the cost of an unpopular course. Since  $r_s(c) \leq m$ , and unpopular courses are available with probability one in the add-drop phase, the final schedule that  $s$  is able to form using  $\tilde{P}_s$  is strictly preferred to that from using  $\hat{P}_s$ .

To complete the argument and prove that student  $s$  is strictly better off using  $\tilde{P}_s$ , we need to argue that  $\Pr(\lambda \in \mathcal{L}_3) > 0$ . If  $\Pr(c \in a_s(\hat{\mathbf{P}})) \in (0, 1)$ , this follows from the fact that  $\tilde{P}_s$  requests  $c$  strictly earlier than does  $\hat{P}_s$ . If, instead,  $\Pr(c \in a_s|\hat{\mathbf{P}}) = 0$  but  $\Pr(c \in a_s(\hat{\mathbf{P}}^{c \uparrow k})) > 0$  where  $k$  is the position of the first  $\hat{\mathbf{P}}$ -unpopular course on  $\hat{P}_s$ , we may need to reiterate the argument using the next unpopular course ahead of  $c$  in  $\tilde{P}_s$ . Once no more unpopular courses are ahead of  $c$  then we can call upon the fact that  $\Pr(c \in a_s(\hat{\mathbf{P}}^{c \uparrow k})) > 0$  where  $k$  is the position of the first  $\hat{\mathbf{P}}$ -unpopular course on  $\hat{P}_s$ . QED

#### 9.4 Proof of Theorem 3 (equilibrium characterization)

(i) Fix an equilibrium  $\hat{\mathbf{P}}$ . For any priority order  $\lambda$ , every student for whom  $c$  belongs their top- $m$  favorite courses either requests  $c$  in the original allocation or requests it in the add-drop phase. Since  $D_c(m) > 1$  there exists a positive measure set of students whose requests are rejected.

(ii) Let  $k$  denote the number of  $\hat{\mathbf{P}}$ -popular courses. If  $k > m$  the claim follows trivially. Suppose  $k \leq m$  and  $\bar{t}_c > k$ . Then there exists a positive mass of students who (1) have  $c$  among their top- $m$  courses, but who place it in position  $k + 1$  or below in their submitted ROLs and (2) get course  $c$  with probability strictly less than 1. Consider one such student, say  $s$ .  $\hat{P}_s$  must contain at least one  $\hat{\mathbf{P}}$ -unpopular course,  $c^*$ , in the top  $k$  positions. This contradicts lemma 2. QED

#### 9.5 Proof of Lemma 3 (best responses with lexicographic preferences)

We prove Lemma 3 in two steps. The first step is identical to the proof of lemma 2.

##### Step 1: Move unpopular courses down

Suppose that none of (i)-(iii) hold for course  $c$ . Consider the deviation strategy  $\tilde{P}_s$ , described in (5), possibly repeated so that no unpopular course appears before  $c$  in  $\tilde{P}_s$ . Then, by the same arguments as claims 1-3 of the proof of Lemma 2, student  $s$  is weakly better off using  $\tilde{P}_s$  than using  $\hat{P}_s$ . He is strictly better off if  $\Pr(c \notin a_s(\hat{\mathbf{P}}, \lambda))$  and  $c \in a_s((\tilde{P}_s, \hat{\mathbf{P}}_{-s})) > 0$ , i.e. if by doing so student  $s$  gets  $c$  more often. Note, that unlike in the proof of Lemma 2, we cannot guarantee that  $\Pr(c \notin a_s(\hat{\mathbf{P}}, \lambda))$  and  $c \in a_s((\tilde{P}_s, \hat{\mathbf{P}}_{-s})) > 0$ .

If there exists another popular course that is preceded by an unpopular course and violates conditions (ii) and (iii) of the Lemma repeat the operation until we reach a point where the deviation strategy increases the probability of getting that popular course or there is no more popular course preceded by an unpopular course. In the first case, student  $s$  is strictly better off and we are done. In the second case, he is weakly better off.

**Step 2: Reorder popular courses**

Without loss of generality (given step 1), consider a strategy  $\widehat{P}_s$  where all unpopular courses appear after popular courses but where, for some popular course  $c$ , conditions (i)-(iii) are all violated. For ease of reference relabel courses such that  $P_s : c_1, c_2, c_3, \dots$

Consider first  $c_1$ . If conditions (i)-(iii) are violated for  $c_1$ , then move it up to position 1. Given that this strictly increases the probability of getting it and that student  $s$  has lexicographic preferences, student  $s$  is strictly better off using this strategy and so we are done.

Suppose next that conditions (i)-(iii) are not violated for  $c_1$  but are violated for  $c_2$ . Consider the deviation where the pair of courses  $\{c_1, c_2\}$  are moved up to the first and second positions of the ROL, *in the order in which they appear on  $\widehat{P}_s$* . This does not decrease the probability that  $c_1$  belongs to the final allocation but strictly increases the probability that  $c_2$  belongs to the final allocation. Given lexicographic preferences, student  $s$  is strictly better off, and so we are done.

We can reiterate this argument if conditions (i)-(iii) are satisfied for  $c_1, \dots, c_k$  but not  $c_{k+1}$ , and show that moving up these courses leaves student  $s$  strictly better off.

To finish the proof, we need to argue that, if the result of the deviations in part I satisfies conditions (i)-(iii), then these deviations must lead to student  $s$  being strictly better off. This follows directly given that conditions (ii)-(iii) were not satisfied by the original strategy. QED.

**9.6 Proof of Theorem 4 (equilibrium characterization with lexicographic preferences)**

1. Towards a contradiction assume that  $\bar{t}_c > \rho_c$ . Then there exist a positive mass of students for whom  $r_s(c) \leq \rho_c$  but who request the course later than round  $\rho_c$  with strictly positive probability and get rejected with positive probability. These students ranked  $c$  in position  $\rho_c + 1$  or lower in his rank order list. This contradicts Lemma 3.

2. From the first part of the Theorem, we know that all courses for which  $D_c(m) > 1$  are  $\widehat{\mathbf{P}}$ -popular in any equilibrium. Moreover, by Lemma 3, we know that the only time when these courses are moved down on students' ROLs is when (1) doing so does not preclude those students from getting them for sure, or (2) when putting them in their truthful position would not ensure that the students get them anyway. In both cases, this downgrading of course  $c$  does not affect the timing of its run-out time relative to truthful play. Because, on the other hand, unpopular courses are moved down, these high demand courses are requested weakly earlier. QED.

## 9.7 Proof of Theorem 5 (ex-post inefficiency of the HBS mechanism)

(1) An allocation is not ex-post efficient possible if there exist a chain of courses  $c_1, c_2, \dots, c_k$  and  $k$  students such that student  $s_1$  prefers  $c_1$  to  $c_2$  but got  $c_2$  and not  $c_1$ , student  $s_2$  prefers  $c_2$  to  $c_3$  but got  $c_3$  and not  $c_2, \dots$  etc, and student  $s_k$  prefers  $c_k$  to  $c_1$  but got  $c_1$  and not  $c_k$ . This means there exists a chain of one-for-one pareto improving trades among those students (given responsiveness, one-for-one trades are the only pareto improving trades we can detect). This implies that, for the particular priority order that generated this allocation,  $t_{c_1} < t_{c_2} < \dots < t_{c_k} < t_{c_1}$ . An impossibility.

(2) Suppose student  $s$  is part of a chain of one-for-one trades. Let  $c$  be the course he wants to give and  $c'$  the course he gets in return. Then either:

- (i) the time of the successful request for  $c <$  the time of the unsuccessful request for  $c'$ , or
- (ii) the time of the successful request for  $c >$  the time of the unsuccessful request for  $c'$ .

By Lemma 3, case (i) is impossible. If student  $s$  prefers  $c'$  to  $c$  then he only downgrades it if he is sure to get it. Thus everybody in the trade must be in the situation (ii). But this is impossible: it cannot be that everyone gets a course that sells out earlier than the course they give up. QED

## 9.8 Proof of Theorem 6

The probability that the  $j$ th student in the random priority order gets his first favorite course is  $\frac{Sm-(j-1)}{Sm}$  under HBS, as  $j - 1$  of the  $Sm$  objects have been selected by other students, and which objects were selected is random due to the uniform i.i.d. assumption. For RSD the figure is  $\frac{Sm-m(j-1)}{Sm}$ , as  $m(j - 1)$  objects have been randomly selected by the time of  $j$ 's turn. Taking the arithmetic average over all  $j$  yields the desired expressions. QED.

## References

- [1] Abdulkadiroğlu, Atila, Yeon-Koo Che and Yosuke Yasuda (2009). "Expanding 'Choice' in School Choice." Mimeo.
- [2] Abdulkadiroğlu, Atila, Parag Pathak, and Alvin E. Roth (2009). "Strategyproofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." Forthcoming, *American Economic Review*.
- [3] Abdulkadiroğlu, Atila and Tayfun Sönmez (2003). "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3), 729-47.
- [4] Abdulkadiroğlu, Atila, Parag Pathak, and Alvin E. Roth (2005a). "The Boston Public School Match" *American Economic Review*, Papers and Proceedings, 95(2), 368-71.

- [5] Abdulkadiroğlu, Atila, Parag Pathak, Alvin E. Roth, and Tayfun Sönmez (2005b). "The New York City High School Match" *American Economic Review, Papers and Proceedings*, 95(2), 364-67.
- [6] Abdulkadiroğlu, Atila and Tayfun Sönmez (1999). "House Allocation with Existing Tenants." *Journal of Economic Theory*, 88: 233-60.
- [7] Angrist, Joshua and V. Lavy (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, 114, 533-575.
- [8] Azevedo, Eduardo and Eric Budish (2009). "Strategyproofness and Manipulability in Large Economies." In preparation.
- [9] Bartlett, Thomas (2008). "Class Warfare: When Getting in is the Hardest Part." *Chronicle of Higher Education* 54(23), Feb 15, Page A1.
- [10] Bogomolnaia, Anna and Hervé Moulin (2001), A New Solution to the Random Assignment Problem, *Journal of Economic Theory*, 100, 295-328.
- [11] Brams, Steven J., Paul H. Edelman and Peter C. Fishburn (2003). "Fair Division of Indivisible Items." *Theory and Decision* 55: 147-80.
- [12] Brams, Steven J. and Philip D. Straffin, Jr. (1979). "Prisoners' Dilemma and Professional Sports Drafts." *The American Mathematical Monthly*, 86(2): 80-88.
- [13] Brams, Steven J., and Alan D. Taylor (1999). *The Win-Win Solution: Guaranteeing Fair Shares to Everybody*. W.W. Norton & Company.
- [14] Budish, Eric (2009). "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes." Mimeo.
- [15] Che, Yeon-Koo and Fuhito Kojima (forthcoming). "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms." *Econometrica*.
- [16] Edelman, Benjamin, Michael Ostrovsky and Michael Schwarz (2007). "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review*, 97(1), 242-59.
- [17] Ehlers, Lars and Bettina Klaus (2003). "Coalitional Strategy-proof and Resource-Monotonic Solutions for Multiple Assignment Problems." *Social Choice and Welfare*, 21: 265-80.
- [18] Featherstone, Clayton and Muriel Niederle (2008). "Ex-Ante Efficiency in School Choice Mechanisms: An Experimental Investigation." Mimeo.

- [19] Gale, David and Lloyd Shapley (1962). "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69(1), 9-15.
- [20] Gehan, Edmund A. (1965), "A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples." *Biometrika*, 52 (1 and 2), 203-223.
- [21] Gollier, Christian (2001). *The Economics of Risk and Time*. MIT Press.
- [22] Guernsey, Lisa (1999). "Business School Puts Courses in Hands of an On-Line Market." *New York Times*. Sept 9.
- [23] Harsanyi, John C. (1953). "Cardinal Utility in Welfare Economics and in the Theory of Risk-taking." *The Journal of Political Economy*, 61(5), 434-435.
- [24] Hatfield, John William (2007). "Strategy-proof and Nonbossy Quota Allocations." Mimeo.
- [25] Hoxby, Caroline (2000). "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *The Quarterly Journal of Economics*, 115, 1239-1285.
- [26] Hylland, Aanund and Richard Zeckhauser (1979). "The Efficient Allocation of Individuals to Positions." *Journal of Political Economy*, 87(2), 293-314.
- [27] Klaus, Bettina and Eiichi Miyagawa (2001). "Strategy-proofness, Solidarity, and Consistency for Multiple Assignment Problems." *International Journal of Game Theory*, 30: 421-435.
- [28] Konishi, Hideo, Thomas Quint and Jun Wako (2001). "On the Shapley-Scarf economy: the case of multiple types of indivisible goods." *Journal of Mathematical Economics*, 35: 1-15.
- [29] Kojima, Fuhito (forthcoming). "Random Assignment of Multiple Indivisible Objects." *Mathematical Social Sciences*.
- [30] Krishna, Aradhna and Utku Unver (2008). "Improving the Efficiency of Course Bidding at Business Schools: Field and Laboratory Studies." *Management Science*, 27(March/April), 262-82.
- [31] Krueger, Alan B. (2003). "Economic Considerations and Class Size." *The Economic Journal*, 113, 485, F34-F63.
- [32] Lazear, Edward P. (2001). "Educational Production." *Quarterly Journal of Economics*, 116(3), 777-803.
- [33] Lehrer, Jim (2008). "College Students Squeezed by Rising Costs, Less Aid: Straining the System." *NewsHour*, aired December 9, 2008. Transcript accessed at [http://www.pbs.org/newshour/bb/education/july-dec08/collegecosts\\_12-09.html](http://www.pbs.org/newshour/bb/education/july-dec08/collegecosts_12-09.html).

- [34] Manea, Mihai (2007). "Serial Dictatorship and Pareto Optimality." *Games and Economic Behavior*, 61(2): 316-330.
- [35] Mas-Colell, Andrew, Michael D. Whinston and Jerry R. Green (1995). *Microeconomic Theory*. Oxford University Press.
- [36] Miralles, Antonio (2009). "School Choice: The Case for the Boston Mechanism." Mimeo.
- [37] Neil, Martha (2008). "NYU Students Seek Coveted Law School Classes, Will Pay Cash." *ABA Journal*. Jul 28.
- [38] Pápai, Szilvia (2001). "Strategyproof and Nonbossy Multiple Assignments." *Journal of Public Economic Theory*, 3(3): 257-71.
- [39] Pathak, Parag (2006). "Lotteries in Student Assignment." Mimeo.
- [40] Pathak, Parag and Tayfun Sonmez (2008). "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review*, September.
- [41] Rivkin, Steven G., Eric A. Hanushek and John F. Kain (2005). "Teachers, Schools and Academic Achievement." *Econometrica*, 73(2), 417-58.
- [42] Roth, Alvin E. (1985). "The College Admissions Problem is not Equivalent to the Marriage Problem." *Journal of Economic Theory*, 36: 277-88.
- [43] Roth, Alvin E. and Elliot Peranson (1999). "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." *American Economic Review*, 89, 748-80.
- [44] Roth, Alvin E. (2002). "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica*, 70(4): 1341-1378.
- [45] Roth, Alvin E. (2008). "What Have We Learned from Market Design?" Hahn Lecture, *Economic Journal*, 118 (March), 285-310.
- [46] Roth, Alvin E. and Elliot Peranson (1999). "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." *American Economic Review*, 89, 748-80.
- [47] Schmeidler, David (1973). "Equilibrium Points of Nonatomic Games." *Journal of Statistical Physics*, 7(4), 295-300.
- [48] Shapley, Lloyd and Herbert Scarf (1974). "On Cores and Indivisibility." *Journal of Mathematical Economics*, 1(1) 23-37.

- [49] Sönmez, Tayfun (1999). "Strategy-Proofness and Essentially Single-Valued Cores." *Econometrica*, 67(3): 677-89.
- [50] Sönmez, Tayfun and Utku Unver (2010). "Course Bidding at Business Schools." Forthcoming, *International Economic Review*.
- [51] Varian, Hal R. (2007). "Position Auctions." *International Journal of Industrial Organization*, 25, 1163-78.
- [52] Vickrey, William (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, 16, 8-37.
- [53] Wilson, Robert (1987). "Game-Theoretic Analyses of Trading Processes," in *Advances in Economic Theory: Fifth World Congress*, ed. by Timothy Bewley. Cambridge, UK: Cambridge University Press, Chap. 2, 33-70.
- [54] Zhou, Lin (1990). "On a Conjecture by Gale About One-Sided Matching Problems." *Journal of Economic Theory*, 52(1), 123-135.