Strategic Behavior in Multi-unit Assignment Problems: Theory and Evidence from Course Allocation*

Eric Budish[†] and Estelle Cantillon[‡]

June 15, 2009

Abstract

This paper uses unusual data – consisting of agents' strategically reported preferences as well as their underlying 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.

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

.

*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, Paul Milgrom, David Parkes, Ariel Pakes, Parag Pathak, Al Roth, and seminar audiences at ASSA 2008, Dagstuhl 2007, Paris Sorbonne and Harvard. 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.

[†]Department of Economics, Harvard University and Harvard Business School. Email: ebudish@hbs.edu

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

1 Introduction

In recent years, economists have been called upon to provide advice on the practical design of allocation procedures in a broad range of settings. Applications have been as diverse as the spectrum auctions and the allocation of advertising slots on internet portals on the one hand, and the assignment of students to schools and doctors to hospitals, on the other hand. A common difficulty that economists have faced in all these settings is that there usually is no perfect solution: that is, there usually does not exist a mechanism that incorporates all of the constraints of the environment and, at the same time, satisfies all of the efficiency and fairness properties desired by market participants. Instead, the relevant choice set is typically comprised of second-best solutions, each of which makes different kinds of compromises. Theory provides some guidance on how to choose, but theory alone is usually insufficient as it fails to provide a sense of magnitudes (Roth, 2002)

This paper quantifies some of the trade-offs that market designers face by studying a *specific* market design problem, the multi-unit assignment problem, and a *specific* solution, the course allocation mechanism used at the Harvard Business School (HBS) for which we have unusual data. Specifically, in addition to students' actual (strategic) reports of their preferences under the HBS mechanism, we have survey data on their underlying truthful preferences, which we argue in the paper can be taken as the true preferences. The combination of truthful and stated preferences is powerful for two reasons. First, it allows us to directly evaluate actual equilibrium play of a non-strategyproof mechanism in terms of the truthful preferences, which is highly unusual for field data (unlike, for example, experimental data). Second, we can use the truthful preferences to simulate equilibrium outcomes in alternative mechanisms, with the advantage over numerical simulations that the counterfactual now incorporates realistic preference heterogeneity. Some of the lessons we draw are specific to the multi-unit assignment problem, others apply to market design more generally.

The multi-unit assignment problem consists in allocating objects among agents without using monetary transfers. Agents demand several objects. The course allocation problem is one example of multi-unit assignment problems: students need to fill their schedule with several courses, courses have limited capacities and, for this reason, it may not be possible to give every student his most desired set of courses. Other examples include the assignment of tasks within an organization, the allocation of shared scientific resources amongst its users, sport drafts, and the allocation of airport slots in many countries. In all these examples, money transfers are ruled out a priori or restricted, so that prices cannot be set to equalize supply and demand.

There is no perfect solution to the multi-unit assignment problem because of the intrinsic conflict between strategyproofness, efficiency and fairness. If we require ex-post efficiency and strategyproofness, we are left with some form of dictatorship over bundles of objects as the only candidate (variants of this result have been shown by Sönmez, 1999; Papai, 2001; Klaus and Miyagawa, 2001; Konishi et al, 2001; Ehlers and Klaus, 2003; Hatfield, 2005; Kojima, 2007). Obviously, dictatorship over bundles results in very unequal outcomes ex-post and is thus unacceptable in most applications. Budish (2008) relaxes strategyproofness and ex-post efficiency to strategyproofness in large markets and approximate ex-post efficiency. He proposes a solution, approximate competitive equilibrium from approximately equal incomes, that has attractive fairness properties. Other solutions have been used in practice. Broadly speaking, these solutions fall into two categories: draft mechanisms where agents choose one object at a time (see e.g. Brams and Straffin, 1979, for an an application to sports drafts) and bidding mechanisms where some kind of fake currency is used to express preferences (e.g. Krishna and Unver, 2008; Sönmez and Unver, 2010, for applications to course allocation). None of these mechanisms elicit truthful preference reports however.

We begin in sections 2 and 3 by studying the theoretical properties of the course allocation mechanism used at HBS. In the HBS mechanism, students submit their preferences over individual courses. The mechanism then allocates courses one-at-a-time to students on the basis of their submitted preferences and a random priority number. At each round, each student receives his most preferred course among those that are available and that he has not received yet. The random priority number determines a student's choosing order in each round. Students choose in increasing order in odd rounds and in decreasing order in even rounds.

We argue that the HBS mechanism is a reasonable candidate to allocate courses: it satisfies several criteria for fairness and it is consistent with ex-post pareto efficiency if students reveal their preferences truthfully. We also show that truthtelling is a Nash equilibrium when preferences are perfectly correlated or when they are independent (theorem 1). However, when preferences are partially correlated, truthful play is no longer a Nash equilibrium. In fact, we show that it is easy to find profitable deviations from truthful behavior in the HBS mechanism: students should underreport their preference for unpopular courses, and overreport their preference for popular courses (Theorem 2). This incentive is intrinsic to the HBS mechanism and does not vanish in large markets. We provide a partial characterization of equilibrium behavior and equilibrium run-out times for courses (Theorems 3, 4).

We then bring the theory to data in sections 4 and 5. We first provide support that our survey-based preference data can indeed be taken as the true preferences of the students. We do this in three ways. First, we argue that students had no incentive to misreport in the survey organized by the administration. Second, we compare our preference data with reported preferences in a second survey that we ran asking again students for their preferences over courses. Third, we show that the relationship between our preference data and actual behavior is consistent with all the predictions from equilibrium play.

In section 6, we quantify the effect of strategic behavior by comparing equilibrium play with non-equilibrium but truthful behavior in the HBS mechanism. Thus our counterfactual exercise is what would happen if the social planner knew students' preferences and used the HBS mechanism to allocate courses. Strategic behavior increases congestion, in the sense that popular courses reach capacity earlier. We also document that strategic behavior hurts students and that this effect is sizeable. Specifically, strategic behavior yields ex-post inefficient outcomes: on average, the allocation resulting from the HBS mechanism leaves at least 1.5 mutually beneficial trades per student on the table, involving 15% of allocated course seats. Strategic behavior also hurts students on a number of ex-ante measures of welfare. For example, 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. 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 (lower is better, and 5.50 is the best possible). Intuitively, there is an asymmetry between the benefits and the costs of strategic behavior. Holding other students' behavior fixed, strategic behavior by a student benefits that student because it increases his chance of getting a preferred bundle of courses. At the same time, the fact that other students act strategically causes the popular courses to reach capacity earlier which hurts the typical student. The empirical results show that the costs outweigh the benefits.

Because strategic behavior hurts students it is natural to compare the HBS mechanism with random serial dictatorship (RSD), which is an anonymous mechanism like the HBS mechanism, but is both ex-post efficient and strategyproof. We turn to this comparison in section 7. In RSD, students receive a random priority number that determines their order and they get to choose all their courses at once in that order. Thus, the first student in the order gets to choose all his preferred courses, but the last student gets to choose only from unpopular courses. Clearly, ex-post, the lucky students are likely to prefer their allocation under RSD than under the HBS mechanism; the unlucky ones will have the reverse preference. In other words, ex-post, we cannot pareto rank the allocations from the HBS mechanism and those from RSD. Ex-ante, however, we find that most students prefer the HBS mechanism to RSD. We also find that the HBS mechanism does better on different measures of expected social welfare. For example, using RSD instead of the HBS mechanism 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. Theoretically, we trace the origin of this poor performance to the fact that the lucky students in RSD make their last choices independently of whether these courses would be some unlucky student's first choice (Theorem 5). We use the term "callous" to describe the behavior induced by RSD. Callousness is specific to environments where agents require more than one good. In fact, in unit demand assignment problems such as school choice, Pathak (2006) and Che and Kojima (2009) show that RSD has attractive welfare properties.

An important empirical hurdle that we face in our analysis is that our data consist of preferences over individual courses whereas students' preferences are eventually driven by their preferences over distributions over bundles of courses. This problem is common in empirical settings where due to restrictions on actions by agents we only get partial information about their preferences.

We develop a series of comparison results that indicate under which conditions we can say with certainty that a particular student strictly prefers one mechanism or one strategy profile over another. 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. Specifically, consider two distributions over bundles and a preference order over these bundles. If one distribution first order stochastically dominates the other, we can say with certainty that this distribution is preferred to the other, irrespective of any information about preference intensities or risk attitudes. The notion of ordinal efficiency (Bogomolnaia and Moulin, 2001) is based on this first order stochastic dominance partial order. Suppose now that the two distributions cannot be ranked according to this f.o.s.d. criterion. There exist assumptions on preference intensities and risk attitudes such that one distribution is preferred to the other, and other assumptions for which the reverse ranking obtains. Our comparison results allow us to disentangle partially these three sources of discrepancy between ex-post and ex-ante welfare (distribution over outcomes, preference intensities, and risk attitude). They provide us with a sense of where the magnitudes are coming from.

Our findings are useful at three levels. At the level of the specific mechanism, our results suggest that the HBS mechanism is flawed but nevertheless a sensible choice relative to the extant alternatives. At the level of the problem, multi-unit assignment, our results suggest where to look for better solutions to the multi-unit assignment problem. One should seek a mechanism that, while not strategyproof, is likely to induce truthful reporting in realistic market environments, and that more resembles HBS than RSD in terms of its fairness and ex-ante efficiency characteristics. Such an approach is taken in Budish (2008). Finally, at the level of market design as a field, our study of a specific mechanism yields two generalizable lessons that contribute to recent active debates in the literature. First, we have identified a new reason (callousness) for the poor ex-ante performance of RSD in environments where agents require several objects and highlighted that, consequently, ex-post efficiency may not even be a proxy for ex-ante efficiency in such environments. So caution should be used when studying the former but hoping for the latter. Second, strategyproofness should not be an inflexible requirement when designing new markets. We elaborate on these themes in the conclusion.

2 Model

2.1 Environment

Courses. There is a finite set of C courses, $\mathcal{C}$.[1] Courses have capacities $\mathbf{q} = (q_1, ..., q_C)$. There are no other goods in the economy other than seats in courses. In particular, there is no divisible numeraire like money.

Students. There is a continuum of students described by the interval $[0, S]$. 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.

Preferences and Demand. Students are allowed to consume any bundle that consists of 0 or 1 seats of each course, and at most $m > 1$ courses in total.

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 P_s defined over permissible bundles of courses. We 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 (Brams and Straffin, 1979, Roth, 1985).[2]

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_sc', with $r_s(c) = 1$ if cP_sc' 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 $D_c(\rho) = \frac{\int_0^S 1_{\{r_s(c) \leq \rho\}}}{q_c}$, $\rho = 1, ..., C$.

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.

[1] We use the terms "students" and "courses" because of our application. We could equally use the generic terms "agents" and "objects".

[2] Preferences are responsive if, for any student s, courses c, c', and bundle of courses X with $c, c' \notin X$ and $|X| < m$, $cP_sc' \iff (X \cup c)P_s(X \cup c')$. Also, $cP_s\emptyset \iff (X \cup c)P_s(X \cup \emptyset)$.

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 (2008) relaxes the responsiveness assumption.

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 $\int_0^S 1_{\{c \in a_s\}} \leq q_c$ 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 at least one student strictly prefers. A feasible allocation is ex-post Pareto efficient if there is no other feasible allocation that all students weakly prefer and at least one student 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 Draft Mechanism and Random Serial Dictatorship.[3]

In each mechanism, each student s reports a rank-order list (ROL) $\widehat{P}_s$ indicating their ordinal preferences over individual courses.[4] Then, the mechanism uniform randomly selects a priority order over the S students. Specifically, a priority order is a bijection λ from the set of students onto itself. $\lambda(t)$ indicates which student has priority t, and $\lambda^{-1}(s)$ gives the priority of student s. The set of priority orders is $\mathcal{L}$.

Random Serial Dictatorship. Students are allocated their courses all-at-once in ascending priority order. Specifically, the algorithm has a single round that takes place from time $t = 0$ to time $t = S$, and 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.[5]

HBS Draft Mechanism (or HBS mechanism). Students are allocated their courses one-at-a-time over a series of m rounds. In odd rounds, which occur during time intervals $[0, S], [2S, 3S], \ldots$, students are allocated courses one-at-a-time in ascending priority order. In even rounds, which occur during time intervals $[S, 2S], [3S, 4S], \ldots$, students are allocated courses one-at-a-time in descending

[3] Several other course-allocation mechanisms are described in Budish (2008), with references provided therein.

[4] A rank order list is a list of individual courses in the order of stated preferences. We write $\widehat{P}_s : c_1, c_2, c_3, \ldots$ to describe that student s puts course c_1 ahead of c_2, and course c_2 ahead of c_3, and so on, in his rank order list. With a slight abuse of notation, we also write $P_s : c_1, c_2, c_3, \ldots$, to describe the true preferences of student s over individual courses.

[5] A course has capacity remaining at time t if the measure of the set of students allocated a seat in that course during time $[0, t]$ is strictly less than the course's capacity.

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 Draft 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. In practice, 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 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.

Equilibrium. The Random Serial Dictatorship is dominant-strategy incentive compatible. For the HBS Draft Mechanism, we focus attention on pure-strategy Nash Equilibria in undominated strategies. For the aftermarket we assume that students report their preferences truthfully, since it is an RSD. Existence of a pure-strategy Nash Equilibrium is guaranteed under Schmeidler (1973) because of the continuum, finite action space, and the fact that a students' payoff depends only on the distribution of other students' reports.

3 Equilibrium of the HBS Draft Mechanism

If we ignore incentives and assume that students report their preferences truthfully, then the HBS Draft Mechanism would satisfy several attractive efficiency and fairness properties. 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 (2008).

These attractive properties help explain the HBS administration's decision to adopt the Draft Mechanism, and may explain the widespread use of similar mechanisms in practice (Brams and Strafin, 1979; Brams and Taylor, 1999). However, as the following example illustrates, truthful play is not to be expected under this procedure.

3.1 A Motivating Example

Example 1 (Overreport Popular Courses) Let $m = 2$ and suppose there are 4 courses with capacity of $\frac{2}{3}S$ seats each. Preferences are as follows:

$\frac{S}{3}$ students have preferences $P_1 : c_1, c_2, c_3, c_4$

$\frac{S}{3}$ students have preferences $P_2 : c_2, c_1, c_3, c_4$

$\frac{S}{3}$ students have preferences $P_3 : c_1, c_3, c_4, c_2$

Truthful play is not a best response for the P_2 types. If they play truthfully they obtain $\{c_2, c_3\}$ with probability 1, whereas the P_1 types receive $\{c_1, c_2\}$ with probability one. If instead they play $\widehat{P}_2 : c_1, c_2, c_3, c_4$, they obtain $\{c_1, c_2\}$ with probability $\frac{2}{3}$ and $\{c_2, c_3\}$ with probability $\frac{1}{3}$. This outcome first order stochastically dominates the outcome under truthful play. In fact, it is easy to check that the strategy profile where type-1 and type-3 students report truthfully and type-2 submit $\widehat{P}_2$ is a Nash equilibrium. In this equilibrium, more students request c_1 in the first round than under truthful play, making c_1 fill up (stochastically) earlier in the round than under truthful play.

The basic story of Example 1 is that students in the HBS draft mechanism will have a tendency to overreport their preferences for popular courses, and that this causes the popular courses to reach capacity sooner. A P_2 type should not waste his first-round choice on c_2, since he can get it 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.

The example suggests that we can make empirical predictions about the relationship between students' truthful preferences and their strategic reports, and about the effects of this strategic reporting on equilibrium course run-out times. We begin (Theorem 1) by highlighting two environments in which we will *not* see strategic misreporting – identical preferences, and independent preferences. Then we show (Theorem 2) that in most other environments many students will have a simple profitable deviation from truthful play.

An example (Example 2) illustrates that it is not possible to reach a complete characterization of equilibrium. The basic issues are multiple equilibria and data incompleteness - we know students' ordinal preferences over individual courses, but equilibrium behavior depends on their cardinal preferences over bundles. We reach a weak characterization of equilibrium (Lemma 2 and Theorem 3) using just ROLs, and then we are able to reach a stronger characterization (Lemma 3 and Theorem 4) by imposing additional structure on the relationship between ROLs and utility functions. All proofs are in an appendix.

Before proceeding, we point out two features of Example 1 that suggest that it is realistic to expect strategic misreporting to emerge in practice. First, the information that P_2 types need to possess in order to find their profitable manipulation is reasonably simple. All they need know is that c_1 is likely to reach capacity in the first round, whereas c_2 is likely to be available in the second round. They need not know anything more specific about which agents report which preference profiles, as is required to find profitable manipulations in other contexts such as two-sided matching markets. Second, the incentive to misreport one's preferences is independent of market size: the example works equally well for a finite economy with $S = 3$ students as it does for a continuum economy.

3.2 Popularity and Simple Manipulations

We begin by defining popularity. Given a priority order λ and students' strategies $\widehat{\mathbf{P}} = (\widehat{P}_s)_{s\in[0,S]}$, let $a_s(\widehat{\mathbf{P}}, \lambda)$ indicate student s's final allocation, including what happens in the aftermarket. Similarly, $a_s(\widehat{\mathbf{P}})$ refers to student s's final (random) allocation under strategies $\widehat{\mathbf{P}}$.

Definition 2 (Popularity): Course c is $\widehat{\mathbf{P}}$-popular if there exists $\lambda \in \mathcal{L}$ and a positive-measure set of students $\mathcal{S}' \subset \mathcal{S}$ such that, for each $s \in \mathcal{S}'$: (i) $c \notin a_s(\widehat{\mathbf{P}}, \lambda)$; and (ii) $c' \in a_s(\widehat{\mathbf{P}}, \lambda)$ for some c' such that $c\widehat{P}_s c'$. Course c is $\widehat{\mathbf{P}}$-unpopular otherwise.

Given $\widehat{\mathbf{P}}$, any course with excess demand during the initial allocation is $\widehat{\mathbf{P}}$-popular. A course that does not run out during the initial allocation but for which more requests than available seats are submitted in the aftermarket is also $\widehat{\mathbf{P}}$-popular. Note that, by the continuum assumption, $\widehat{\mathbf{P}}$-popular courses are also $\widehat{\mathbf{P}}_{-s}$-popular in an economy without student s.

We begin by highlighting two environments in which truthful play constitutes an equilibrium.

Theorem 1 (Truthful Play in Equilibrium): Consider the two following environments:

1. Identical preferences: $P_s = P_{s'}$ for all s, s' or
2. "Independent" preferences: for any two $\mathbf{P}$-popular courses c, c', $D_c(\rho) = D_{c'}(\rho)$ for $\rho = 1, ..., C$.

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

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 *over*report 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 1 indicates that *partial* correlation of preferences is what drives strategic misreporting in the HBS draft mechanism.

Our next result pins down the effect of a change in reported preferences by one student on his allocation. It provides the workhorse for the analysis of strategic incentives in the HBS mechanism.

Lemma 1 (Downgrade Lemma): Label courses such that $\widehat{P}_s : c_1, c_2, ..., c_k, c_{k+1}, ...$ Let $\widehat{P}_s^{c_k \downarrow l}$: $c_1, c_2, ..., c_{k-1}, c_{k+1}, ..., c_l, c_k, c_{l+1}, ...$ (c_k is "downgraded" to position l). Consider two strategy profiles $\widehat{\mathbf{P}}$ and $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$. Limit attention to the initial allocation. For all priority orders λ:

(i) if s does not obtain c_k under $\widehat{\mathbf{P}}$ then he obtains the same bundle under $\widehat{\mathbf{P}}$ as under $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$.

(ii) if s obtains c_k under $\widehat{\mathbf{P}}$ but not under $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$ then he obtains exactly one course under $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$ that he does not obtain under $\widehat{\mathbf{P}}$. This course is from the set $\{c_{k+1}, c_{k+2}, ..., c_l, c_{l+1}, ...\}$. Otherwise, the two allocations are identical.

(iii) if s obtains c_k under both $\widehat{\mathbf{P}}$ and $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$, then either (a) he obtains the same bundle under both profiles, or, (b) he obtains exactly one course under $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$ from the set $\{c_{k+1}, c_{k+2}, ..., c_l\}$ that he does not obtain under $\widehat{\mathbf{P}}$, and exactly one course under $\widehat{\mathbf{P}}$ from the set $\{c_{l+1}, c_{l+2}, ...\}$ that he does not obtain under $(\widehat{P}_s^{c_k \downarrow l}, \widehat{\mathbf{P}}_{-s})$. Otherwise, the two allocations are identical.

The proof works as follows. Fix an ordering λ. Assume c_k is available at the time $\widehat{P}_s$ asks for it, because otherwise the two strategies will obtain exactly the same courses at exactly the same times. By downgrading c_k, $\widehat{P}_s^{c_k \downarrow l}$ (the downgrade strategy) asks for c_{k+1} one round earlier than does $\widehat{P}_s$ (the original strategy), 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 $\widehat{P}_s$, etc. If, before he asks for c_k, $\widehat{P}_s^{c_k \downarrow l}$ gets a course that $\widehat{P}_s$ does not, then the two strategies get back "in synch", requesting 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 $\widehat{P}_s^{c_k \downarrow l}$, etc. This is the path that leads to case (iii)(b) of the Lemma. The other cases are similar.

The Downgrade Lemma implies that many students will have a natural and simple deviation from truthful play: downgrade unpopular courses.

Theorem 2: (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 for all λ.

Theorem 2 formalizes our assertion that students are likely to strategically misreport their preferences in realistic environments. The simple deviations of Theorem 2 lead directly to overreporting preferences for popular courses, and increased congestion. However, Theorem 2 is not yet an equilibrium characterization.

3.3 Equilibrium

We begin with the following partial characterization of best responses and equilibrium.

Lemma 2 (Best-response Characterization): Consider any candidate equilibrium $\widehat{\mathbf{P}}$. Suppose c is $\widehat{\mathbf{P}}-$popular, and that c is amongst student s's top-m favorite courses (i.e., $r_s(c) \leq m$.). Then it is not a best response for student s to submit a ROL $\widehat{P}_s$ where (i) c appears after a $\widehat{\mathbf{P}}-$unpopular course; and (ii) $\Pr(c \in a_s|\widehat{\mathbf{P}}) \in (0, 1)$.

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

(i) c runs out with probability one during either the initial allocation or the aftermarket.

(ii) the supremum of run-out times for course c over the different realizations of λ, $\bar{t}_c$, is weakly less than the number of $\widehat{\mathbf{P}}-$popular courses.

Lemma 2 formalizes that students underreport their preferences for unpopular courses in equilibrium. Theorem 3 indicates that all courses for which demand, based on students' true top m choices, exceeds supply 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.

Lemma 2 and Theorem 3 do not yet fully capture the intuition of Example 1. The following example illustrates the difficulties.

Example 2 (Multiple Equilibria) S students require $m = 2$ courses. Courses have $.4S$ seats each. 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
$.25S$	P_1	c_1, c_2, other
$.25S$	P_2	c_2, c_1, other
$.30S$	P_3	c_3, other
$.1S$	P_4	c_3, c_1, other
$.1S$	P_5	c_3, c_2, other

Truthful play is always an equilibrium. 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 $\widehat{P}_4 : c_1, c_3, other$, and the P_5 types submit $\widehat{P}_5 : c_2, c_3, other$. In this equilibrium, the P_4 types receive c_1 for sure and receive c_3 with probability 0.5. If they deviate from this equilibrium to play truthfully (the only deviation to consider given their preferences), they receive c_3 for sure, and receive c_1 with probability $\frac{.4S-.1S-.25S}{.25S} = 0.2$, which may be less preferable if the intensity of preference for c_3 is not too large.

While the truthful play equilibrium seems natural, the equilibrium in which c_3 does not reach capacity until round two is surprising.[6] This latter equilibrium suggests that the order in which courses reach capacity might reverse between truthful play and strategic play, and suggests that some students' best responses may involve downgrading a popular course, even at the risk of not getting it.

It is possible to enhance Example 2 to create a third equilibrium in which c_3 reaches capacity earlier than under truthful play. Let fraction $\varepsilon > 0$ be type $P_1' : c_1, c_3, c_2, other$ instead of P_1, and similarly let fraction ε be type $P_2' : c_2, c_3, c_1, other$ instead of P_2. The two equilibria identified above exist so long as the P_1' types' intensity of preference for c_1 versus c_3 is sufficiently large (and similarly for the P_2' types), and ε is sufficiently small. In addition there can be a third equilibrium in which the P_1' and P_2' types report $\widehat{P}_1' : c_3, c_1, c_2, other$ and $\widehat{P}_2' : c_3, c_2, c_1, other$, and all other types report truthfully. This is an equilibrium so long as the P_1' and P_2' types have a high enough value for c_3 relative to c_1 and c_2, respectively. So, in the enhanced example, c_3 can reach capacity earlier, later, or at the same time as under truthful play.

The fundamental difficulty is that the HBS Draft Mechanism induces a coordination problem: students want to ask early for courses that other students ask for early. This coordination problem is especially difficult for two reasons that stem from the fact that students report ROLs. First, whereas truthful play depends only on one's ordinal preferences over singletons, strategic best responses depend on one's cardinal (vNM) preferences over bundles. So the relationship between truthful play and strategic best responses depends on information not contained in the truthful play. Second, the round at which a request for a course is actually made is stochastically related to its position on the student's reported ROL. A course at position five could be considered an any of the first five rounds, depending on which earlier requests are successful.

To go further than Theorem 3, we make additional assumptions on the relationship between a students' ordinal preferences over courses and their utility function. In Example 2, c_3 sells out at equilibrium in the first round (as it would under truthful play) if the P_4 and P_5 students have a

[6] Observe that both equilibria are consistent with Theorem 3. While course c_3 might run out later under strategic play than we would expect given the truthful preferences, it still is sure to reach capacity.

sufficiently higher value for c_3 than they do for their second choice. 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 $P_s : c_1, c_2, ..., c_k, ...$ 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 $i = 1, 2, ..., k$ 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 sharp 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 candidate equilibrium $\widehat{\mathbf{P}}$. Suppose c is $\widehat{\mathbf{P}}-$popular. Then it is not a best response for student s to submit a ROL $\widehat{P}_s$ where (i) c appears after any course c' such that cP_sc'; and (ii) $\Pr(c \in a_s|\widehat{\mathbf{P}}) \in (0, 1)$.

In words, Lemma 3 says that students only downgrade a course when they are sure to get it. Theorem 4 provides a tighter characterization of when courses run out at equilibrium:

Theorem 4 (Equilibrium Characterization with Lexicographic Preferences): Suppose students have lexicographic preferences. Suppose that $\widehat{\mathbf{P}}$ is a Nash equilibrium $\widehat{\mathbf{P}}$, and that $D_c(m) > 1$. Then

(i) c runs out with probability one during the initial allocation

(ii) $\bar{t}_c \leq \rho_c = \inf\{\rho : D_c(\rho) > 1\}$

(iii) $\Pr(c \in a_s|\widehat{\mathbf{P}}) \in (0, 1) \Rightarrow \widehat{r}_s(c) \leq r_s(c)$ (where $\widehat{r}_s(c)$ denotes the rank of course c in student s's submitted preferences)

Theorem 4 provides an upper bound on the times by which courses run out, based on the true demand for these courses. Because run-out times under truthful play must also satisfy part (ii) of Theorem 4, and the relationship between $\inf\{\rho' : D_c(\rho') > 1\}$ and the round at which c reaches capacity under truthful play depends in a complicated way on what else students requesting c ask for in earlier rounds, it does not provide a direct comparison of run-out times under truthful and strategic play. However, part (iii) of the theorem is suggestive that popular courses are likely to run out earlier at equilibrium: students never underreport their preference for popular courses, but they might overreport.

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 Data

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.

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.[7] 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 at number one in the submitted rank order lists (ROLs) with the number of times each was ranked first. Students do not receive any feedback on their

[7] 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."

individual assignment of courses from the trial run.

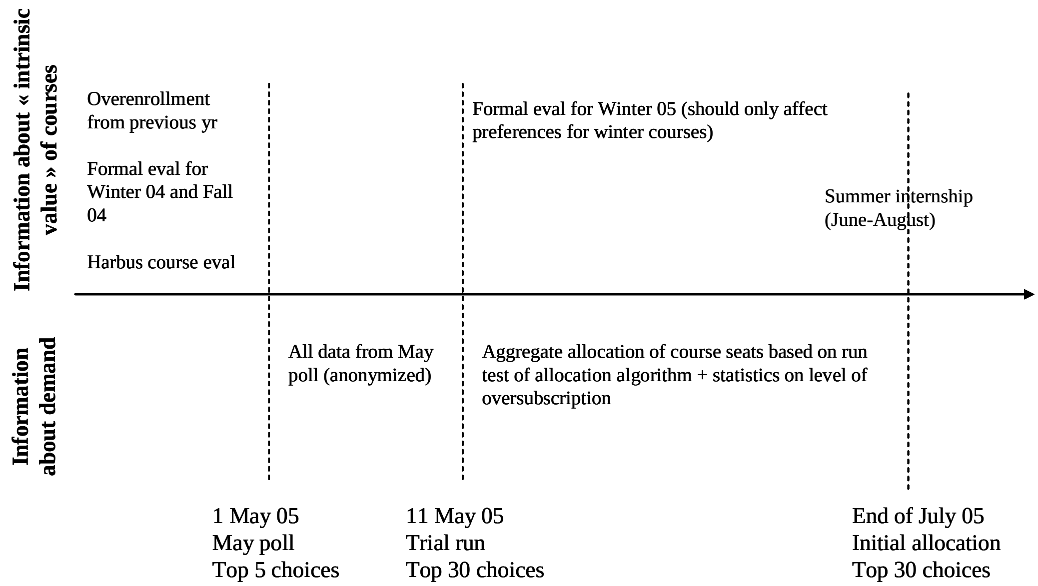

Figure 1: Timing of information and actions

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

Between any of these three rounds of preference elicitation, students receive new information about courses and aggregate demand, and some changes are made to courses. Specifically, students learned about the poll results and two courses were added before they submitted their ROLs for 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 capacity increased or decreased slightly. In addition, students received the official course evaluations for the Winter 2005 courses. Finally, 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.[8]

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.[9]

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

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.[10] 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

[8] The reason this sums to a bit more than 10 courses per student is the half courses.

[9] 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.

[10] We dropped two students who ranked the same course more than once for a course appearing in their top 5 courses.

appeared in the original rank order list.[11]

5 Evidence of Strategic Behavior

In this section, we provide evidence that students understand the strategic incentives of the HBS mechanism and we quantify the impact of their strategic behavior. We focus on the May poll preferences and the July run submitted preferences because incentives are clearest on these two occasions. Students in the May poll were explicitly asked by the administration to state their true preferences, and we do not see a particular incentive for them to disobey this request. Submitted preferences in the July run are those used for the initial allocation of courses.

5.1 Evidence that Students Overreport Popular Courses

Our analysis rests on the assumption that the course rankings elicited in the May poll represent students' truthful preferences and that the difference between the May poll preferences and the submitted preferences in the July run of the algorithm are mainly driven by strategic considerations. In this section, we provide suppport for the assumption that strategic behavior rather than something else drives behavior in July and, at the same time, we document the kind of overreporting predicted by the theory.

Conceptually, the identification issue is the following. In principle, submitted preferences may change over time because of three distinct reasons: genuine preference change, new information, or strategic consideration. Genuine preference changes not driven by new information are likely to be idiosyncratic and should therefore not affect aggregate demand for individual courses. For this reason, our comparisons will focus on the distributions of course ranks in the student population.[12]

Consider course j and a sample of students. Define course j's distribution of ranks in that sample, $F_j(r)$, as the proportion of students in the sample that rank course j on or before r. Absent any new information or strategic considerations, rank distributions should be equal in two different samples. Theorems 2 and 4 suggest that - absent any new information - demand for popular courses should be higher in July than in the May poll, at the expense of courses that are not capacity constrained. New information could also shift demand but those shifts should not

[11] 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.

[12] Comparing aggregate preferences also eliminates any small randomness in the May poll preferences due to students' "carelessness" in filling in the poll.

be systematically related to how popular a course is.[13] This - the way demand changes with the popularity of courses - is at the heart of our identification strategy.

To compare demands 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 the test is the following. Fix a course, say course j, and consider two independent samples of students of size n_1 and n_2. An observation is a student's rank for course j 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 don't know how that student ranks course j 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 $F_{j1}(r) = F_{j2}(r)$ for $r \leq R$ where F_{1j} and F_{2j} are the distributions of ranks in sample 1 and sample 2 respectively and $R = \min\{$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.[14] A course is said to be low demand, if demand in the May trial run was less than 70% the available capacity. All other courses are medium demand courses.

[13] Moreover, very little new information about the courses was generated between the May poll and the July run. Students had plenty of information - both official and unofficial - about course qualities by the time of the May poll. The only new piece of information was the official evaluations for the Winter 05 courses released in July, but according to the HBS administration, students had already collected and posted unofficial evaluations for these courses before the May poll.

[14] 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 that reported demand is unchanged between the May poll and the July run is rejected for 42 out of the 82 courses (51%). 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.[15] This is consistent with Theorems 2 and 4 according to which students will overreport popular courses and consequently underreport less popular courses.

As a comparison, 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. We can reject the null hypothesis that reported demand is unchanged 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.[16] 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 strategic considerations, whereas most observed changes in submitted preferences in January are due to new information.[17]

[15] As a robustness check, we also implemented Gehan's test on the July run demand using the convention that whenever different sections of the same course appeared in a student's rank order list, the last time the course appeared was representative of his preference over that course. The results are similar, except that more courses are moved to the "no difference" category. The demand in July is lower for 14 low demand courses (instead of 17) and for 16 (instead of 17) medium demand. The demand in July is higher for 6 high demand courses (instead of 7).

[16] Because the demand for one course is not independent of the demand for another course, it is hard to interpret the levels of rejection beyond the fact that one is lower than the other.

[17] 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 rules out the hypothesis that differences in demand is driven by new information. This leaves the hypothesis that demand changes were driven by strategic considerations.

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

5.2 Effect of Strategic Behavior on Run-out Times and Congestion

Overreporting of popular courses moves up the times by which courses run out and increases demand for popular courses, creating a sense of congestion for those courses. 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(ii) has more bite. For each course for which $D_c(5) > 1$ based on the poll data and adjusted capacities, we checked whether $\bar{t}_c \leq \rho_c$. All 6 such courses satisfied this stronger test.

In the remainder of this subsection, we go one step further than the theory results and assess by how much strategic behavior increases congestion relative to truthful behavior. Theorem 2 and Theorem 4 suggest that run out times should move up because students overreport their preferences for popular courses. To investigate this question, we use the students who answered the May poll as representatives of the overall student population and construct a scaled down version of the HBS market on the basis of those students and adjusted course capacities. Our working assumption is that the submitted ROLs in July by those students correspond to an equilibrium of this scaled down market.[18]

Our May poll data contain the top 5 courses of 456 of the 916 students who participated in

[18] To test this working hypothesis we applied Gehan's test described above to the distributions of course ranks in July for the students who did answer the poll and 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.

the July run of the algorithm. We construct truthful preferences for each of these students as follows. We consider that their truthful top 5 courses correspond to their top 5 courses in the May poll.[19] Other courses are moved down to position 6 and below in a way that preserves their relative ranking in the July ROLs. 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 $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$. We call the result "constructed truthful preferences." Constructed truthful preferences provide a lower bound on the extent of strategic behavior because they assume that the relative ranking of courses not in the top 5 is truthful.

We ran the HBS algorithm for 10,000 random orderings over students using both the constructed truthful preferences and the July run preferences. Table 4 summarizes the results. We say that a course reaches capacity earlier under strategic play if the time it reaches capacity is earlier than the time it reaches capacity under truthful play in at least 99% of the simulations.

Table 4: Effect of Strategic Behavior on Congestion

		Course Reaches Capacity			
	N	Earlier	Later	Indet.	Never
Low demand courses	32	0	0	2	30
Medium demand courses	37	3	4	12	18
High demand courses	22	18	1	3	0

All high-demand courses but one reached capacity earlier on average under strategic play and 18 of them reached capacity earlier in more than 99% of the simulations (the one exception is the course that was also the exception in Table 2). More courses reach capacity under strategic behavior: 43 courses compared to 41 courses under truthful play. For the 37 courses that reach capacity under both truthful play and strategic behavior, the round by which they do so goes from 6.73 to 6.35 on average. If we focus on the 13 courses that reach capacity in the first 5 rounds under truthful play, the average time at which this happens moves from 3.48 to 2.56.

6 Welfare Consequences of Strategic Play

The purpose of this section is to assess 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

[19] Courses that were offered in the May poll but were no longer available in the July run are dropped.

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, 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 5.2). 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 λ and run the HBS mechanism using students' strategically submitted 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:

$$\max \sum x_{scc'} \text{ s.t.}$$

$$\sum_s \sum_{c'} x_{scc'} - x_{sc'c} = 0, \; \forall c \tag{1}$$

$$\sum_{c'} x_{scc'} + x_{sc'c} \leq 1, \; \forall s, c \tag{2}$$

$$x_{scc'} \in \{0, 1\}, \; \forall s, c, c' \tag{3}$$

$$x_{scc'} = 1 \Rightarrow c \in a_s(\widehat{\mathbf{P}}, \lambda), c' \in a_s(\widehat{\mathbf{P}}, \lambda), c' P_s c$$

Variable $x_{scc'}$ 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 (3). The constraints (1) capture the adding-up condition that each course must be given as often as it is received. The constraints (2) prevent a student from trading the same course twice, both to ensure feasibility and to avoid double-counting.

We repeat this exercise for 20 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	15.5%	(0.37%)
% of Students Executing		
0 Trades	16.3%	(1.1%)
1 Trade	35.5	(2.2)
2 Trades	29.7	(2.0)
3+ Trades	18.5	(1.4)

Table 5 suggests that the HBS mechanism is meaningfully inefficient ex-post, and that these inefficiencies harm the large majority of students.[20] By contrast if students played truthfully there would be zero one-for-one Pareto improving trades available.

[20] In Spring 2005 two Harvard Business School MBA students surveyed 160 of their classmates for a class project related to the HBS course-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, suggesting that students are aware of ex-post Pareto inefficiencies. Of course, our integer program will find more trades than students realistically can be aware of: for many random priority orders we are able to find 43-way trades involving one seat in each of the 43 popular courses.

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 immediately yields the following simple comparison criterion:

Comparison Result 1 (Responsive Preferences) A student strictly prefers play $\widehat{\mathbf{P}}$ to play $\widehat{\mathbf{P}}'$ if, for every realization of λ, that student gets more of her top course, more of her top two courses, more of her top three courses, ... etc under $\widehat{\mathbf{P}}$ than under $\widehat{\mathbf{P}}'$. If the reverse relationship holds, we say that she prefers $\widehat{\mathbf{P}}'$. If, for all λ, she gets the same outcome under both strategy profiles, she is indifferent. The comparison is indeterminate otherwise.[21]

We implement this criterion for the comparison between truthful play and strategic play of the HBS mechanism by running the HBS algorithm for both truthful play and strategic play, for each of 10,000 randomly drawn priority orders.

We find that 25% of students are unambiguously harmed by strategic play: for all 10,000 trials they prefer their outcomes under truthful play to that under strategic play. 4% of the students strictly prefer strategic play and 1% get the same allocation for all trials. For the majority of students however, the comparison is indeterminate. This is not surprising because a ranking requires that a student likes her allocation weakly better under one play for all 10,000 trials. Two elements drive the indeterminacy. First, in some cases, we do not have the information to determine whether the student prefers one or the other allocation. Second, a student can prefer her allocation under one play for some realizations of λ, and under another play for some other realizations of λ, and we do not have the information to aggregate her preferences over all λ. In the rest of this section, we impose additional assumptions on preferences that fill in this information gap and pin down the indeterminate cases.

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:

[21] In principle, we could get a stronger result without further assumptions by requiring first order stochastic dominance for the distributions over bundles, instead of requiring dominance for every realization of λ. However, there are over 10^{13} bundles of courses, and responsiveness provides only a very partial ordering over these bundles. Thus, this approach is both computationally difficult to implement and unlikely to yield much stronger results.

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 comparison here is across all λ and not for every λ. However, Comparison Result 2 will still leave some cases indeterminate because the restrictions on preferences do not entirely pin down the preferences of the student over bundles. A special case of additive preferences that pins down preferences over bundles is when the difference in utilities derived from the 1st and 2nd favorite courses is the same as the difference in utilities derived from the n^{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 $\overline{r}_s(\widehat{\mathbf{P}}, \lambda)$ denote the average rank of the courses that student s get under strategy profile $\widehat{\mathbf{P}}$ for the priority order λ :

(i) Independently of his attitude towards risk, student s prefers strategy profile $\widehat{\mathbf{P}}$ to strategy profile $\widehat{\mathbf{P}}'$ if $-\overline{r}_s(\widehat{\mathbf{P}}, \cdot)$ first-order stochastically dominates $-\overline{r}_s(\widehat{\mathbf{P}}', \cdot)$.
(ii) If student s is risk averse, he prefers strategy profile $\widehat{\mathbf{P}}$ to strategy profile $\widehat{\mathbf{P}}'$ if $-\overline{r}_s(\widehat{\mathbf{P}}, \cdot)$ second-order stochastically dominates $-\overline{r}_s(\widehat{\mathbf{P}}', \cdot)$.[22]
(iii) If student s is risk neutral, he prefers strategy profile $\widehat{\mathbf{P}}$ to strategy profile $\widehat{\mathbf{P}}'$ if $\sum\limits_{\lambda} \overline{r}_s(\widehat{\mathbf{P}}, \lambda) < \sum\limits_{\lambda} \overline{r}_s(\widehat{\mathbf{P}}, \lambda)$.

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 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 $\widehat{\mathbf{P}}$ to strategy profile $\widehat{\mathbf{P}}'$ if he gets his first choice course

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

more often under $\widehat{\mathbf{P}}$ than under $\widehat{\mathbf{P}}'$ 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 $\widehat{\mathbf{P}}'$, 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			
	Risk Neutral	Any Risk Attitude	Risk Averse	Risk Neutral	Lexicographic
	(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 Comparisons at the Social Level

We now evaluate social welfare. Clearly, just based on the assumption of responsive preferences, we cannot pareto rank truthful play and strategic play because 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 $\widehat{\mathbf{P}}$ to play $\widehat{\mathbf{P}}'$, if, for any j, the expected number of top-j courses allocation to students under $\widehat{\mathbf{P}}$ exceeds that under $\widehat{\mathbf{P}}'$. Society prefers $\widehat{\mathbf{P}}'$ 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 $\widehat{\mathbf{P}}$ to play $\widehat{\mathbf{P}}'$ if $-\overline{r}.(\widehat{\mathbf{P}},\cdot)$ first-order stochastically dominates $-\overline{r}.(\widehat{\mathbf{P}}',\cdot)$. (the notation $\overline{r}.(\widehat{\mathbf{P}},\cdot)$ indicates that the distribution is taken over priority orders and students)
(ii) If students are risk averse, society prefers play $\widehat{\mathbf{P}}$ to play $\widehat{\mathbf{P}}'$ if $-\overline{r}.(\widehat{\mathbf{P}},\cdot)$ second-order stochastically dominates $-\overline{r}.(\widehat{\mathbf{P}}',\cdot)$.
(iii) If students are risk neutral, society prefers play $\widehat{\mathbf{P}}$ to play $\widehat{\mathbf{P}}'$ if $\sum\limits_{\lambda}\sum\limits_{s}\overline{r}_s(\widehat{\mathbf{P}},\lambda) < \sum\limits_{\lambda}\sum\limits_{s}\overline{r}_s(\widehat{\mathbf{P}},\lambda)$.

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

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.[23] Thus CR5 obtains and by consequence 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

[23]The kink in the HBS Truthful line at rank 6 is a mechanical effect due to the way we construct truthful preferences. 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.

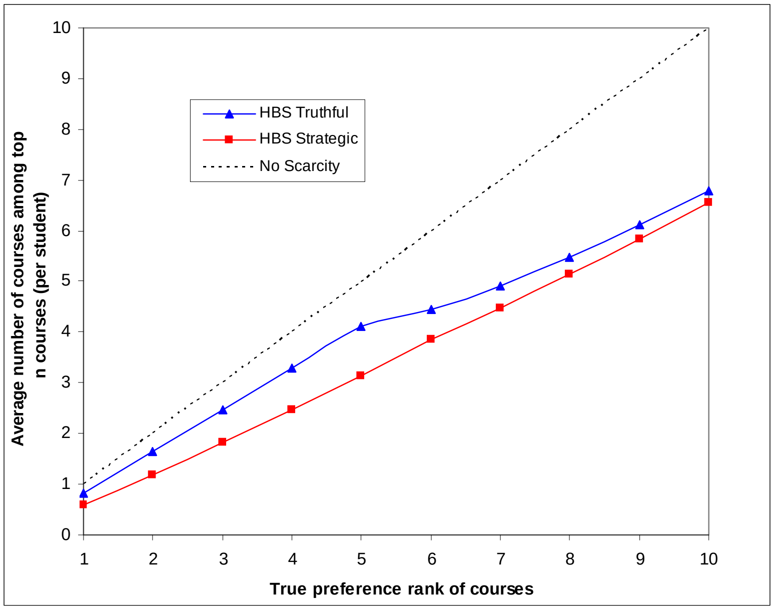

Figure 1: Aggregate outcomes by preference ranks: truthful versus strategic play of the HBS mechanism (CR5, CR7)

play of the HBS mechanism. The difference appears to be 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.[24]

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

[24] 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.

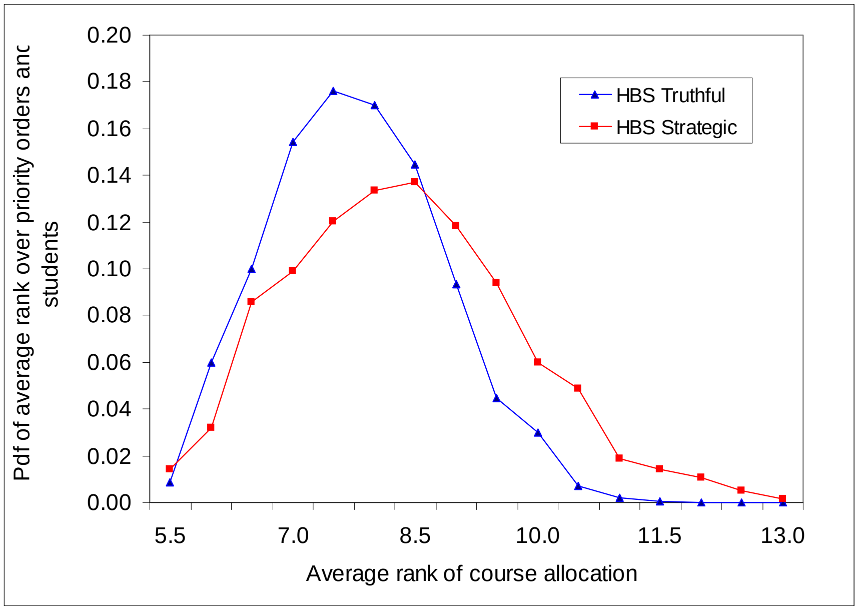

Figure 2: Distribution of the average rank received: truthful versus strategic play of the HBS mechanism (CR6)

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 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 will be entirely indeterminate.

As soon as we put additional structure on preferences we find that the HBS mechanism appears to be more attractive ex-ante than RSD. RSD's ex-ante unattractiveness may be surprising since ex-post it is efficient. We provide a theoretical explanation of RSD's poor performance at the end of this section.

7.1 Comparisons at the Individual Level

We repeat the methodology of Section 6.2. 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			
	Risk Neutral	Any Risk Attitude	Risk Averse	Risk Neutral	Lexicographic
	(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, unlike in Table 6.

7.2 Comparisons at the Social Level

Figure 3 implements CR5 and CR7 by comparing the aggregate rank distributions of HBS and RSD. The distribution under strategic play of the HBS mechanism first-order stochastically dominates that under truthful play of RSD. So a utilitarian social planner prefers HBS to RSD when students are risk neutral and have any additive preferences. This is surprising because it suggests that RSD's ex-post Pareto efficiency is not a good proxy for social welfare.

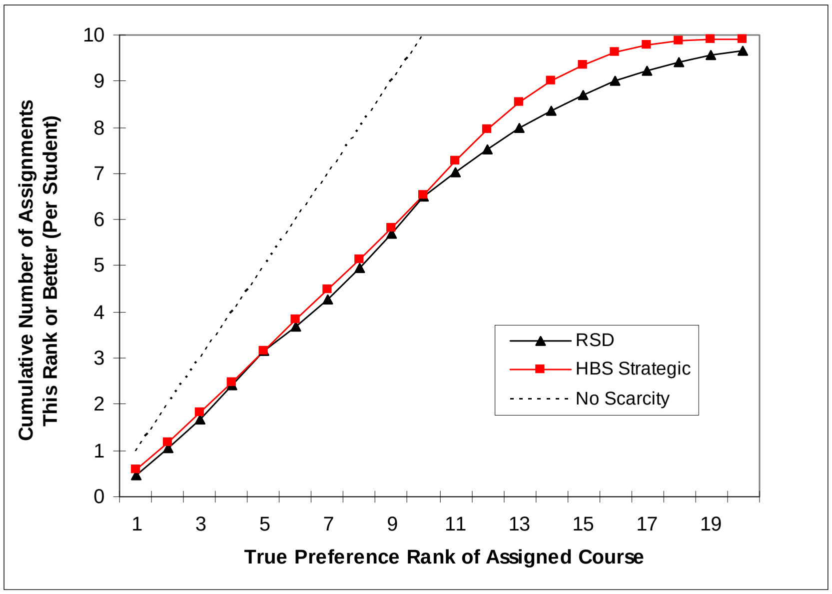

Figure 3: Aggregate outcomes by preference ranks: RSD versus strategic play of the HBS mechanism (CR5, CR7)

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.[25] 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.

[25] 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).

We can get a better understanding of the risk to which RSD exposes students by examining the distribution of average ranks. Figure 4 implements CR6. RSD puts much more weight on the tails of the distribution, and indeed is second-order stochastically dominated by HBS, i.e., CR6 (ii) obtains. 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.

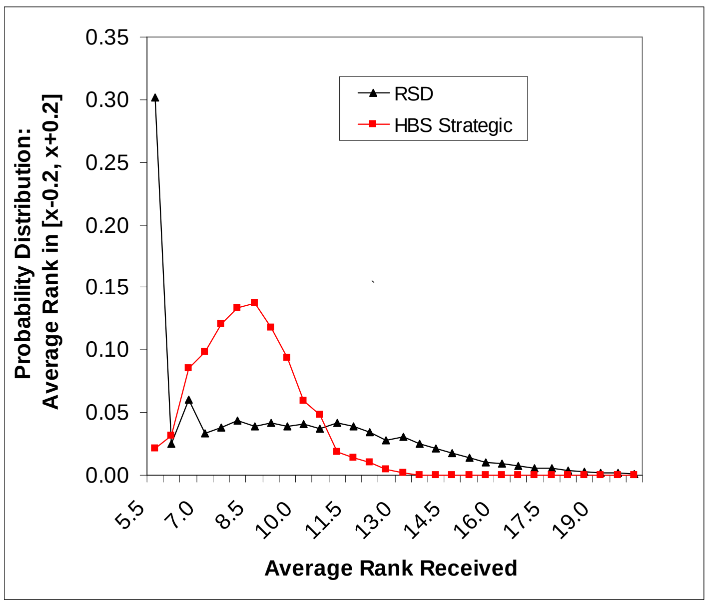

Figure 4: Distribution of the average rank received: RSD vs. strategic play of the HBS mechanism (CR6)

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. For both average-rank and lexicographic preferences, the ex-post utility benefit to these fortunate students is exceeded by the ex-post harm they cause to the unfortunate students. The fortunate students "callously" disregard the preferences of others. In the absence of monetary transfers, the allocation reached by RSD is ex-post Pareto efficient. Ex-ante, though, students do not know if they will be fortunate or unfortunate, and the large majority regard this distribution over callous outcomes as unattractive.

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

Example 3 (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. Consider the RSD choosing order $AABB$ and the HBS choosing order $ABBA$.

RSD. A 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.

HBS. A gets his 1st favorite object, and either his 2nd, 3rd, or 4th, each with equal probability. 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.

Table 8. Results of Example 3

	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 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 3 generalizes to larger economies. For instance, in an HBS-sized version of Example 3 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.[26]

The following simple theorem shows that the first-choice-course finding in Example 3 generalizes.

Theorem 5 (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 $S \to \infty$ the proportion converges to $1 - \frac{1}{2m}$ 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 Theorem 5 illustrates that the Callousness of RSD in

[26] We also are able to show theoretically that Example 3 generalizes to any number of students S, with $m = 2$ and Sm courses each in unit supply. The proof is available upon request, as are simulation results and code.

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

Our analysis of a specific imperfect solution to a specific open problem yields two generalizable lessons that contribute to the broader market design literature. The first lesson concerns the relationship between ex-post and ex-ante efficiency in random allocation mechanisms. Researchers have long acknowledged that ex-ante efficiency is the more compelling 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 lotteries over allocations tend to be less tractable to work with than sure allocations, especially in the case of multi-unit demand. As a result, the literature on random allocation mechanisms has largely focused on ex-post efficiency.[27]

Our paper sounds a cautionary note against advocating for a mechanism on the basis of its ex-post efficiency properties if ex-ante efficiency is what we care about.[28] In our environment, RSD is ex-post efficient, whereas the HBS mechanism is severely inefficient ex-post. Yet, ex-ante, the HBS is actually more attractive. This wedge between ex-ante efficiency and ex-post efficiency in multi-unit assignment is driven by what we call callousness.

Our second lesson concerns the role of strategyproofness in market design. 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): equilibrium behavior is not informationally demanding on the part of participants so we can reasonably

[27] Two important exceptions are Hylland and Zeckhauser (1979), who propose a single-unit assignment mechanism that is ex-ante efficient, and Bogomolnaia and Moulin (2001), who propose a single-unit assignment mechanism that satisfies an intermediate notion of efficiency, ordinal efficiency, that is stronger than ex-post yet weaker than ex-ante.

More recently, an exciting series of papers has taken an ex-ante perspective to the single-unit assignment problem of school choice, using either simulation evidence (Abdulkadiroglu, Che and Yasuda, 2009; Miralles, 2009) or laboratory evidence (Featherstone and Niederle, 2008).

[28] 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."

trust that they will play equilibrium strategies. A second and related reason is that strategyproof mechanisms make it easy to advise market participants and they help 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.[29]

Our paper sounds a cautionary note against imposing strategyproofness as a strict design requirement. In our environment, strategic behavior by market participants indeed harms welfare. The magnitudes are large. However, we also find that the highly manipulable HBS mechanism is preferable to the strategyproof RSD on natural measures of welfare, including all of the measures emphasized by the actual market administrators. That is, the costs of manipulability are large, but the costs of requiring strategyproofness are larger.

One final point to note, as we think about ex-post versus ex-ante efficiency, and the costs and benefits of requiring strategyproofness, is the revealed preference of school administrators. Despite its being the only known anonymous mechanism that is ex-post efficient and strategyproof, we are not aware of any university that has adopted RSD as its course-allocation mechanism.

[29] Budish (2008) suggests that a mechanism that is not strategyproof but that is difficult to manipulate in large markets may yield many of the same benefits, and proposes a course-allocation mechanism that satisfies such a criterion of approximate incentive compatibility.

9 Appendices

9.1 Proof of Theorem 1

Proof: (i) Identical preferences: Let $P_s = P_{s'} : c_1, c_2, c_3, ...$ 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 $\widehat{P}_s^{cc'\uparrow}$ 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 $\widehat{P}_s^{cc'c''\uparrow}$, $\widehat{P}_s^{cc'c''c'''\uparrow}, ...$

We show that the sequence $\widehat{P}_s^{c_1\uparrow}, \widehat{P}_s^{c_1 c_2\uparrow}, ..., \widehat{P}_s^{c_1 ... c_{C-1}\uparrow} = P_s$ constitutes a chain of profitable deviations. 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' requests are "in synch" again and so are other students' requests. This implies there are no additional discrepancies between the two outcomes.■

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

Claim 2: Student s gets either exactly the same courses under $(\widehat{P}_s^{c_1 \cdots c_k\uparrow}, \mathbf{P}_{-s})$ and $(\widehat{P}_s^{c_1 \cdots c_{k-1}\uparrow}, \mathbf{P}_{-s})$ or his two allocations differ by exactly one course: he gets c_k under $(\widehat{P}_s^{c_1 \cdots c_k\uparrow}, \mathbf{P}_{-s})$ 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 λ 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.$\blacksquare$

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 1(i) then follows from transivity.

(ii) Independent preferences: Let $\overline{r}$ be such that $D_c(\overline{r}-1) < 1$ and $D_c(\overline{r}) \geq 1$. Under truthful play, all **P**-popular courses run out exactly in round $\overline{r}$. 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 $\overline{r}-1$ courses. Moreover, it maximizes the chance that he gets $\overline{r}$ **P**-popular courses, and conditional on getting $\overline{r}$ **P**-popular courses, the probability distribution it generates on those $\overline{r}$-course bundles first order stochastically dominates the outcome from any alternative (here we use the assumption of responsiveness and the fact that all $\overline{r}$-course bundles differ by a single course, the one requested in round $\overline{r}$, to generate an order over them). Finally, the fact that **P**-unpopular courses are listed in order of preferences ensures that he gets his $m-\overline{r}$ (or $m+1-\overline{r}$) most preferred courses among them. The claim then follows from responsiveness and the assumption of vNM preferences over uncertain outcomes. QED

9.2 Proof of the Downgrade Lemma

Fix λ and $\widehat{\mathbf{P}}_{-s}$. For an arbitrary student s, we compare the outcome of the following alternative strategies:

$\widehat{P}_s : c_1, c_2, ..., c_k, c_{k+1}, ..., c_l, c_{l+1}, ...$

$\widehat{P}_s^{c_k \downarrow l} : c_1, c_2, ..., c_{k+1}, ...c_l, c_k, c_{l+1}, ...$

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. Let r_k be the round at which this happens. Without loss of generality we can focus on the case where student s gets c_k in round r_k if he plays $\widehat{P}_s$ (otherwise, the two strategies are equivalent for this particular λ and part (i) of the lemma follows trivially). Because student s has zero mass, her change of strategy does not affect course seats availabilities and thus, a fortiori, the allocation and requests in any given round of other students.

Under $\widehat{P}_s^{c_k \downarrow l}$, student s requests course c_{k+1} in round r_k and all other courses one round earlier than under strategy $\widehat{P}_s$ until we either reach a course, say $c_{k'}$, in $\{c_{k+1}, ..., c_l\}$ 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 $\{c_{k+1}, ..., c_l\}$ 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. From round $r_{k'}$ onwards, student s' requests are in synch under both strategies and thus he gets the same outcome until the algorithm reaches position l in his ROL. In particular, there is no other course in $\{c_{k+1}, ..., c_l\}$ that one strategy gets and not the other.

   When the algorithm reaches the requests in position l, student s requests course c_k under $\widehat{P}_s^{c_k \downarrow l}$. From then on, we need to distinguish two subcases. If c_k is no longer available, the student's requests remain in synch under both strategies and so there is no additional differences. The final outcomes differ in one course: $\widehat{P}_s^{c_k \downarrow l}$ gets $c_{k'}$ at the cost of c_k. If c_k is available, student s now 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 that is no longer available when $\widehat{P}_s^{c_k \downarrow l}$ requests it (after which his requests are in synch and thus 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 $\{c_{l+1}, ..., c_C\}$ 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

   Two outcomes are possible at that round. If c_k is available, student s's requests become in synch again and there is thus no difference in outcomes. If, on the other hand, $\underline{c_k \text{ is no longer}}$ $\underline{\text{available}}$ in that round, student s continues to request courses one round earlier under strategy $\widehat{P}_s^{c_k \downarrow l}$, and in particular the courses in $\{c_{l+1}, ...c_C\}$. This has two possible consequences: either there exists a course that he gets under $\widehat{P}_s^{c_k \downarrow l}$ but that is no longer available when the original strategy requests it (after which his requests are in synch and thus 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 $\widehat{P}_s^{c_k \downarrow l}$ is never requested by the original strategy). In both cases, the two final allocations differ in a single course: there exists a course in $\{c_{l+1}, ..., c_C\}$ that student s gets under $\widehat{P}_s^{c_k \downarrow l}$ instead of c_k that he gets under $\widehat{P}_s$.

9.3 Proof of Theorem 2

Let ρ denote the position in the ROL of the lowest ranked course that student s ever receives in the HBS mechanism (over all possible λ's). If $\rho = m$, then student s always receives his top m courses and $\widehat{P}_s^{\text{simple}}$ gets exactly the same courses as P_s so that the claim follows trivially. Thus assume for the remainder that $\rho > m$. The strategy of the proof is to show that a sequence of deviations from P_s, that consist in downgrading the $\widehat{\mathbf{P}}$-unpopular courses to the bottom half of the top m courses in student s's ROL while preserving the relative ordering of the $\widehat{\mathbf{P}}$-popular and $\widehat{\mathbf{P}}$-unpopular courses, leaves student s weakly better off. For ease of reference, relabel courses such that $P_s : c_1, c_2, c_3, ...c_C$.

Claim 1: Let c_k be the lowest-ranked $\widehat{\mathbf{P}}$-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

Proof of claim 1: By the downgrade lemma (iii), $\widehat{P}_s^1$ gets exactly the same courses among $\{c_1, ..., c_m\}$ (because c_k is unpopular, student s gets it for sure), or exactly one additional course in $\{c_{k+1}, ..., c_m\}$ than P_s, at the cost of a course $\{c_{m+1}, ..., c_C\}$. Because all courses in $\{c_{k+1}, ..., c_m\}$ are strictly preferred to courses in $\{c_{m+1}, ..., c_C\}$, 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 2: Let c_j be the n^{th} lowest $\widehat{\mathbf{P}}$-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 $\widehat{\mathbf{P}}$-unpopular courses that he has already downgraded). Student s is weakly better off using $\widehat{P}_s^n$ than $\widehat{P}_s^{n-1}$.

Proof of claim 2: By the downgrade lemma (iii), $\widehat{P}_s^n$ gets either exactly the same courses among $\{c_1, ..., c_m\}$ or exactly one additional course among the $\widehat{\mathbf{P}}$-popular 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 $\{c_{m+1}, ..., c_C\}$. 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 $\widehat{\mathbf{P}}$-unpopular course to downgrade. At each deviation, student s is weakly better off. The claim then follows by transitivity. QED

9.4 Proof of Lemma 2 (Downgrade Lemma)

Suppose for a contradiction 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 $r_s(c) \leq m$ and $\Pr(c \in a_s|\widehat{\mathbf{P}}) \in (0,1)$. Let u denote the last such $\widehat{\mathbf{P}}$-unpopular course to appear before c on $\widehat{P}_s$.

We construct an alternative rank-order-list, $\widetilde{P}_s$, by making two changes versus $\widehat{P}_s$. First, switch the positions of c and u. Second, if there is any $\widehat{\mathbf{P}}$-popular course ranked lower than c on $\widehat{P}_s$, find the lowest-ranked such course, say p, and downgrade u until the position right after p. The two strategies can be written as:

$$\widehat{P}_s : ...u...c...p...$$
$$\widetilde{P}_s : ...c... ...pu...$$

where the ellipses denote courses that do not change in their relative position between $\widehat{P}_s$ and $\widetilde{P}_s$. Partition the set of priority orders into four sets:

$$\mathcal{L}_1 : c \in a_s|(\widehat{P}_s, \widehat{\mathbf{P}}_{-s}) \text{ and } c \in a_s|(\widetilde{P}_s, \widehat{\mathbf{P}}_{-s})$$
$$\mathcal{L}_2 : c \notin a_s|(\widehat{P}_s, \widehat{\mathbf{P}}_{-s}) \text{ and } c \notin a_s|(\widetilde{P}_s, \widehat{\mathbf{P}}_{-s})$$
$$\mathcal{L}_3 : c \notin a_s|(\widehat{P}_s, \widehat{\mathbf{P}}_{-s}) \text{ and } c \in a_s|(\widetilde{P}_s, \widehat{\mathbf{P}}_{-s})$$
$$\mathcal{L}_4 : c \in a_s|(\widehat{P}_s, \widehat{\mathbf{P}}_{-s}) \text{ and } c \notin a_s|(\widetilde{P}_s, \widehat{\mathbf{P}}_{-s})$$

If $\lambda \in \mathcal{L}_1$ then in 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 $\widehat{P}_s$. If there is such a difference, than there is some unpopular course that $\widehat{P}_s$ receives that $\widetilde{P}_s$ does not. This follows from a slight modification of the proof of the downgrade lemma: from the time that $\widehat{P}_s$ obtains u and $\widetilde{P}_s$ obtains c, the two strategies are in synch, requesting the same courses at the same time, until the round at which $\widehat{P}_s$ receives c. Now, $\widetilde{P}_s$ requests the popular courses between c and p on $\widehat{P}_s$ one round earlier than does $\widehat{P}_s$, so there is the possibility that $\widetilde{P}_s$ receives a course that $\widehat{P}_s$ does not. Because unpopular courses are available with probability one in the aftermarket, the final schedule that s is able to form having used $\widetilde{P}_s$ is at least weakly preferred to that from using $\widehat{P}_s$.

If $\lambda \in \mathcal{L}_2$, then, again, $\widetilde{P}_s$ gets every popular course that $\widehat{P}_s$ receives, plus possibly one additional. Now, though, the additional course might be any popular course ranked lower than u on $\widehat{P}_s$. Because of the aftermarket, $\widetilde{P}_s$ yields a weakly better outcome than does $\widehat{P}_s$.

If $\lambda \in \mathcal{L}_3$, then $\widetilde{P}_s$ gets c while $\widehat{P}_s$ does not, and otherwise they get exactly the same popular courses. Since unpopular courses are available with probability one in the aftermarket, and since c is one of s's favorite m courses, the final schedule that s is able to form having used $\widetilde{P}_s$ is strictly preferred to that from using $\widehat{P}_s$.

Last, note that $\Pr(\lambda \in \mathcal{L}_4) = 0$. For all priority orders, $\widetilde{P}_s$ requests c strictly earlier than does $\widehat{P}_s$, and so it is impossible that $\widehat{P}_s$'s request for c is successful while $\widetilde{P}_s$'s is rejected.

To complete the argument, we note that $\Pr(\lambda \in \mathcal{L}_3) > 0$. This follows from the assumption that $\Pr(c \in a_s|\widehat{\mathbf{P}}) \in (0,1)$ and the fact that $\widetilde{P}_s$ requests c strictly earlier than does $\widehat{P}_s$.

9.5 Proof of Theorem 3

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

(ii) Let k denote the number of $\widehat{\mathbf{P}}$-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 $\widehat{\mathbf{P}}$-popular course c among their top-m most preferred courses, but who place it in position $k+1$ or below in their submitted ROL and (2) get course c with probability strictly less than 1. Consider one such student, say s. $\widehat{P}_s$ must contain at least one $\widehat{\mathbf{P}}$-unpopular course, c^*, in the top k positions. This contradicts lemma 2 if $\Pr(c \in a_s|\widehat{\mathbf{P}}) \in (0,1)$. If $\Pr(c \in a_s|\widehat{\mathbf{P}}) = 0$, then there exists a profitable deviation similar to that described in the proof of Lemma 2. Form $\widetilde{P}_s$ by switching the positions of c and c^*, and then downgrading c^* until immediately after the lowest-ranked $\widehat{\mathbf{P}}$-popular course on $\widehat{P}_s$. By the same argument as in Lemma 2, the new strategy $\widetilde{P}_s$ gets every $\widehat{\mathbf{P}}$-popular course that $\widehat{P}_s$ gets, and possibly one additional. With strictly positive probability the additional course is c, because it is ranked earlier on $\widetilde{P}_s$ than its maximum sell-out time $\bar{t}_c$, and so $\Pr(c \in a_s|(\widetilde{P}_s, \widehat{\mathbf{P}}_{-s}) > 0$. Hence, the deviation constitutes a strict improvement.

9.6 Proof of Lemma 3 (Safe Strategies)

Suppose that for some s, c there exists a course c' such that cP_sc' and yet $c'\widehat{P}_sc$, and that $\Pr(c \in \mathbf{a}_s|\widehat{\mathbf{P}}) \in (0,1)$. If there exist multiple such courses, consider the first one before c. Thus, $\widehat{P}_s = \ldots, c', \widetilde{c}_1, \widetilde{c}_2, \ldots, \widetilde{c}_n, c, \ldots$, where $\widetilde{c}_i P_s c$ for all i.

We need to show that there exists an alternative strategy for student s, $\widehat{P}'_s$, that yields strictly greater expected utility. Consider $\widehat{P}'_s = \widehat{P}_s^{c'\downarrow C}$, i.e. $\widehat{P}'_s$ corresponds to $\widehat{P}_s$ except that c' is moved to the end of the ROL.

By Lemma 2, for every λ, this alternative strategy yields exactly the same outcome for student s

or it yields exactly one extra course among those from $\widetilde{c}_1$ until the end of $\widehat{P}_s$ (at the cost of c'). Consider in particular the subset of courses $\{\widetilde{c}_1, \widetilde{c}_2, ..., \widetilde{c}_n, c\}$ which, by construction, are all strictly preferred to c'. We claim that there exists at least one course among them that student s gets with increased probability. Indeed, suppose that for all λ, the alternative strategy gets exactly the same set of courses in $\{\widetilde{c}_1, \widetilde{c}_2, ..., \widetilde{c}_n\}$ as $\widehat{P}_s$ (otherwise we are done). Then it requests course c exactly one round earlier than $\widehat{P}_s$, which strictly increases the probability that he receives it since, by assumption, $\Pr(c \in \mathbf{a}_s | \widehat{\mathbf{P}}) \in (0,1)$. Because student s has lexicographic preferences, this is enough to guarantee that he has strictly higher expected utility with $\widehat{P}'_s$. QED.

9.7 Proof of Theorem 4

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

(i) follows immediately from (ii) given $D_c(m) > 1$.

(iii) If $\widehat{r}_s(c) > r_s(c)$, there exists course c' with $c'\widehat{P}_s c$ yet cP_sc', a contradiction with Lemma 3. QED.

9.8 Proof of Theorem 5

The probability that the jth 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] Abdukadiroğlu, Atila, Yeon-Koo Che and Yosuke Yasuda (2008). "Expanding 'Choice' in School Choice." Mimeo.

[2] Abdukadiroğlu, 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] Abdukadiroğlu, Atila, Parag Pathak, Alvin E. Roth and Tayfun Sonmez (2006). "Changing the Boston School Choice Mechanism." Mimeo.

[4] Abdukadiroğlu, Atila and Tayfun Sönmez (1999). "House Allocation with Existing Tenants." *Journal of Economic Theory*, 88: 233-60.

[5] Bogomolnaia, Anna and Hervé Moulin (2001), A New Solution to the Random Assignment Problem, *Journal of Economic Theory,* 100, 295-328.

[6] Brams, Steven J., Paul H. Edelman and Peter C. Fishburn (2003). "Fair Division of Indivisible Items." *Theory and Decision* 55: 147-80.

[7] Brams, Steven J. and Philip D. Straffin, Jr. (1979). "Prisoners' Dilemma and Professional Sports Drafts." *The American Mathematical Monthly*, 86(2): 80-88.

[8] Brams, Steven J., and Alan D. Taylor (1999). The Win-Win Solution: Guaranteeing Fair Shares to Everybody. W.W. Norton & Company.

[9] Budish, Eric (2008). "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes." Mimeo.

[10] Cantillon, Estelle and Martin Pesendorfer (2006). "Combination Bidding in Multi-Unit Auctions." Mimeo.

[11] Che, Yeon-Koo and Fuhito Kojima (2009). "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms." Mimeo.

[12] Ehlers, Lars and Bettina Klaus (2003). "Coalitional Strategy-proof and Resource-Monotonic Solutions for Multiple Assignment Problems." *Social Choice and Welfare*, 21: 265-80.

[13] Featherstone, Clayton and Muriel Niederle (2008). "Ex-Ante Efficiency in School Choice Mechanisms: An Experimental Investigation." Mimeo.

[14] Gehan, Edmund A. (1965), "A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples." *Biometrika,* 52 (1 and 2), 203-223.

[15] Gollier, Christian (2001). The Economics of Risk and Time. MIT Press.

[16] 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.

[17] Hatfield, John William (2005). "Strategy-proof and Nonbossy Quota Allocations." Mimeo.

[18] Hylland, Aanund and Richard Zeckhauser (1979). "The Efficient Allocation of Individuals to Positions." *Journal of Political Economy*, 87(2), 293-314.

[19] Klaus, Bettina and Eiichi Miyagawa (2001). "Strategy-proofness, Solidarity, and Consistency for Multiple Assignment Problems." *International Journal of Game Theory*, 30: 421-435.

[20] 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.

[21] Kojima, Fuhito (2007). "Random Assignment of Multiple Indivisible Objects." Mimeo.

[22] 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.

[23] Manea, Mihai (2007). "Serial Dictatorship and Pareto Optimality." *Games and Economic Behavior*, 61(2): 316-330.

[24] Mas-Colell, Andrew, Michael D. Whinston and Jerry R. Green (1995). Microeconomic Theory. Oxford University Press.

[25] Miralles, Antonio (2008). "School Choice: The Case for the Boston Mechanism." Mimeo.

[26] Pápai, Szilvia (2001). "Strategyproof and Nonbossy Multiple Assignments." *Journal of Public Economic Theory*, 3(3): 257-71.

[27] Pathak, Parag (2006). "Lotteries in Student Assignment." Mimeo.

[28] Pathak, Parag and Tayfun Sonmez (2008). "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review*, September.

[29] Roth, Alvin E. (1985). "The College Admissions Problem is not Equivalent to the Marriage Problem." *Journal of Economic Theory*, 36: 277-88.

[30] Roth, Alvin E. (2008). "What Have We Learned from Market Design?" Hahn Lecture, *Economic Journal*, 118 (March), 285-310.

[31] 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.

[32] Sönmez, Tayfun (1999). "Strategy-Proofness and Essentially Single-Valued Cores." *Econometrica*, 67(3): 677-89.

[33] Sönmez, Tayfun and Utku Unver (2010). "Course Bidding at Business Schools." Forthcoming, *International Economic Review*.

[34] 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.

[35] Zhou, Lin (1990). "On a Conjecture by Gale About One-Sided Matching Problems." *Journal of Economic Theory*, 52(1), 123-135.