Comparing Mechanisms by their Vulnerability to Manipulation*

Parag A. Pathak[†] Tayfun Sönmez[‡]

January 2009

Abstract

This paper introduces a method to compare direct mechanisms based on their vulnerability to manipulation or deviation from truthful reporting. We explore the following idea: if a player can manipulate mechanism ψ whenever some player can manipulate mechanism φ, then ψ is more manipulable than φ. Our notion generates a partial ordering on mechanisms based on their degree of manipulability. We illustrate the concept by comparing several well-known mechanisms in the matching and auction literature. The applications include comparisons between stable matching mechanisms, school choice mechanisms, auctions for internet advertising, and multi-unit auctions.

KEYWORDS: straightforward incentives, strategy-proofness, matching, auctions

*We thank seminar participants at the Arne Ryde Symposium in Lund (Sweden), CEMFI (Madrid), CORE (Louvain), ECARES (Brussels), Harvard-MIT, Princeton, SED 2008 (Michigan), SITE 2008, and the 2008 Social Choice and Welfare conference (Montreal). Lars Ehlers, Bengt Holmström, Fuhito Kojima, Stephen Morris, and Muhamet Yildiz provided helpful suggestions.

[†]Department of Economics, MIT, e-mail: ppathak@mit.edu

[‡]Department of Economics, Boston College, e-mail: sonmezt@bc.edu

1 Introduction

Mechanisms for collective decision making and resource allocation are often evaluated on numerous dimensions. Some criteria, such as whether a mechanism is efficient or generates an equitable allocation, are based on the outcomes of the mechanism. Evaluation of these dimensions depends on the predicted behavior of participants. Other criteria relate to the procedural dimensions of a mechanism or the process by which the outcomes are achieved. A common desideratum is that a mechanism is "simple" or provide "straightforward" incentives. One of the most demanding requirements is to focus on strategy-proof mechanisms, where truth-telling is a dominant strategy for all participants.

Procedural aspects of mechanisms and the desire for straightforward incentives have been important in many policy discussions. In the 1950s, a committee of the American Medical Association supported the use of the Boston Pool Plan as the basis for a centralized clearinghouse, which turned out to be formally equivalent to a deferred acceptance algorithm when selecting the procedure to place graduating medical students to residencies in the US. One rationale for their recommendation was the possibility of strategic behavior in the system used to place medical school graduates to hospital residency programs in New York City in the 1930s (Roth 1984, 2003).[1] Over half a century later, school officials in Boston adopted a strategy-proof mechanism to assign elementary, middle, and high school students to schools over their existing system, known as the *Boston mechanism*, because it simplifies the problem faced by students when they decide how to reveal their preferences over schools to the mechanism. Importantly, the policy change was seen to protect families who were unaware of the strategic aspects of school admissions in Boston (Pathak and Sönmez 2008). One reason that New York City's Department of Education decided to change their student assignment system, which is used to assign over 90,000 students to high school every year, was "to reduce the amount of gaming families had to undertake to navigate a system with a shortage of good schools" (Kerr 2003).

For auctions, the desire for straightforward incentives is one of the leading arguments in favor of mechanisms based on the Vickrey-Clarke-Groves mechanism, which is strategy-proof. In the single unit case, dominant strategy incentive compatibility is one argument for Vickrey's

[1]Roth (2003) reports that the National Student Internship Committee noted that under the originally proposed algorithm, a student could suffer by submitting a rank-order list that listed as first choice a position he or she was unlikely to obtain.

celebrated second-price auction (Vickrey 1961). In their description of the first FCC Spectrum Auction, McAfee, McMillan, and Wilkie (2008) write "A second aspect of simplicity, and one harder to implement, requires that a simple strategy be optimal, or nearly optimal, behavior ... Economists were very much concerned that they could articulate simple bidding strategies for bidders that would perform well. It was expected that novice bidders would probably adopt such strategies."

Providing straightforward incentives to participants has a number of virtues in real-life allocation problems. One virtue is that it is easier to guide participants, and easier for participants to learn or find their optimal strategy. Another virtue is that simplified incentives may encourage entry by unsophisticated players, and this may be desirable. Moreover, truthfully eliciting information from participants can also be desirable if the information is used for purposes in addition to computing the allocation, such as in elections where determining the preferences of the electorate may also influence the policy choices of responsive candidates or in matching problems, where revealed preferences may be used to decide whether to expand or shut down a school or residency program.[2]

Unfortunately, strategy-proof mechanisms are known to exist only in particular domains, and in many cases strategy-proof mechanisms perform poorly on other dimensions. Moreover, there are sometimes informal intuitions about the degree to which a mechanism encourages straightforward play even though a mechanism is either strategy-proof or not. Our goal in this paper is to develop a notion which can be used to compare mechanisms that are not strategy-proof based on the degree to which they encourage manipulation. We show how the notion can formalize intuitions about manipulability in a unified framework.

Our focus is on direct mechanisms, where players report their types. The notion we develop is based on a comparison of the states under which two mechanisms are manipulable. There are two versions: Under the first, mechanism ψ is weakly more manipulable than mechanism φ if whenever a player can profitably manipulate φ, some player can profitably manipulate ψ. Equivalently, whenever truth-telling is a Nash equilibrium under ψ, truth-telling is a Nash equilibrium under φ as well. Under the second, mechanism ψ is strongly more manipulable than mechanism φ if whenever a player profitably manipulates φ, she can manipulate ψ as well.

[2]There is some evidence that this is taking place in New York City's public school choice program. See Gootman (2006).

We investigate these two definitions in several well-known mechanisms in the matching and auction literature. Many of our examples are inspired by the recent literature on "market design" (cf. Roth 2002, Milgrom 2004). We begin by considering various classes of matching problems. In many-to-one matching problems (commonly known as the college admissions model), the student-optimal stable mechanism is strongly more manipulable for colleges than the college-optimal stable mechanism. In school choice problems, the constrained student-optimal stable mechanism when students can rank at most k schools is weakly more manipulable than the constrained student-optimal stable mechanism when students can rank at most ℓ schools for $\ell > k$. Finally, the constrained Boston mechanism when students can rank at most k schools is weakly more manipulable than the constrained version of the student-optimal stable mechanism when students can rank at most k schools.

We then consider examples from various auction models. Our first result is for the auction of a single item: the ℓ^{th} price auction is strongly more manipulable than the k^{th} price auction for $\ell > k \geq 2$. The last two results are for auctions involving multiple units. When there are multiple items for sale but bidders can only win one item, as in Internet keyword advertisement auctions, the Generalized First Price Auction is strongly more manipulable than the Generalized Second Price Auction. In auctions where there are multiple units and bidders can win more than one unit, as in the case of Treasury bill auctions, the discriminatory auction is strongly more manipulable than the uniform-price auction.

The next section introduces the general framework and discusses related literature. Section 3 presents examples from problems in two-sided matching and indivisible good allocation. Section 4 presents examples from auction models. The last section concludes.

2 General Framework

2.1 Primitives

There are a finite number of players indexed by $i = 1, ..., N$. The set of allocations is A which is finite and has generic element $a = (a_1, ..., a_N)$ where a_i specifies player i's assignment. Each player has a preference relation R_i defined over the set of assignments. Let P_i be the strict counterpart of R_i. Let $R = (R_i)$ denote a profile of preferences. The set of possible types for player i is T_i with generic element t_i. We adopt the convention that t_{-i} are the types of players other than player i, and define R_{-i} and P_{-i} accordingly.

We refer to a **problem** as a profile of player types $(t_1, ..., t_N)$. Let $T^N = T_1 \times \ldots \times T_N$. A direct mechanism is a function $\varphi : T^N \rightarrow A$. It is a single-valued mapping of the profile of player types to an element in A. Let $\varphi_i(t)$ be the assignment that player i obtains from the mechanism φ when the reported types are t.

Definition 1. *A mechanism φ is* **manipulable by player i** *at problem t if there exists a type t_i' such that $\varphi_i(t_i', t_{-i})P_i\varphi_i(t)$.*

For a mechanism to be manipulated by player i, she must strictly prefer her allocation when she reports a type other than her true type over what she obtains when she reports her true type.

Definition 2. *A mechanism φ is* **manipulable** *if there exists some player i and problem t such that the mechanism is manipulable by player i at problem t.*

A mechanism is **strategy-proof** if truth-telling is a dominant strategy for all players. A strategy-proof mechanism is not manipulable. We consider the following two notions to compare the manipulability of mechanisms.

Definition 3. *A mechanism ψ is* **weakly more manipulable** *than mechanism φ if*

i) for any problem where φ is manipulable, ψ is manipulable, and

ii) there is at least one problem where ψ is manipulable although φ is not.

Definition 4. *A mechanism ψ is* **strongly more manipulable** *than mechanism φ if*

i) for any problem where φ is manipulable, ψ is manipulable by any player who can manipulate φ, and

ii) there is at least one problem where ψ is manipulable although φ is not.

The second requirement in either concept eliminates the possibility that any one of the two mechanisms is more manipulable than the other. It follows from the definitions that if ψ is strongly more manipulable than φ, then ψ is also weakly more manipulable than φ.

If φ is a strategy-proof mechanism and ψ is not, then ψ is both weakly and strongly more manipulable than φ. Our main interest is in the case where ψ and φ are both not strategy-proof mechanisms. While both notions make no explicit reference to equilibrium behavior, the weak notion is equivalent to the following: if at any problem, truth-telling is a Nash equilibrium under φ, it is also a Nash equilibrium under ψ even though the converse does not hold.

2.2 Related literature

There are other complementary approaches to comparing the ease of manipulation in mechanisms which are not strategy-proof. The most related literature is the one which characterizes the domains under which a mechanism is not manipulable (see, for instance, Moulin 1980, Ergin 2002). Our weak notion is related to this literature because it is equivalent to a comparison of domains. An equivalent definition of weak manipulability is to require that the set of types for which truth-telling is a Nash equilibrium in ψ is a strict subset of the set of types for which truth-telling is a Nash equilibrium under φ. One main difference is that this earlier literature has characterized non-manipulable domains for specific mechanisms, while our aim is to make comparisons across mechanisms in a variety of problems.

Another related paper is Day and Milgrom (2008)'s study of core selecting auctions. They define an incentive profile to be vector of each bidder's maximal gain from truthful reporting when all other bidders report truthfully and consider auctions which minimize the incentive profile subject to selecting a core allocation. The idea of making comparisons across mechanisms is also related to Dasgupta and Maskin (2008). They show that if a voting rule satisfies various axioms for a set of preferences, then simple majority voting rule also satisfies those axioms on the same set of preferences. Other than focusing on voting rules, another major difference is that our focus is on comparing mechanisms based on the extent to which they encourage manipulation, while Dasgupta and Maskin focus on other properties.

3 Matching mechanisms

3.1 Comparing stable matching mechanisms

In the college admissions model, there are a number of students each of whom should be assigned a seat at one of a number of colleges. Each student has a strict preference ordering over all colleges as well as remaining unassigned and each college has a strict preference ordering over students. Each college has a maximum quota.

Formally, a college admissions problem is tuple $\Gamma = (S, C, P_S, P_C)$. S and C are sets of students and colleges and $P_S = (P_s)_{s \in S}$ and $P_C = (P_c)_{c \in C}$. For each student $s \in S$, P_s is a strict preference relation over C and being unmatched (being unmatched is denoted by s). Each college c has maximum capacity q_c, and we assume that each college's preferences are responsive (Roth 1985). That is, the ranking of a student is independent of her colleagues, and any set of students exceeding quota is unacceptable.[3] Given this assumption, we sometimes abuse notation and let P_c be the preference list of college c defined over singleton sets and the empty set. If $sP_c\emptyset$, then s is said to be acceptable to c. Similarly, c is acceptable to s if cP_ss. Non-strict counterparts of P_s and P_c are denoted by R_s and R_c, respectively.

A matching μ is a mapping from $S \cup C$ to $S \cup C$ such that (i) for every s, $|\mu(s)| = 1$, and $\mu(s) = s$ if $\mu(s) \notin C$, (ii) $\mu(c) \subseteq S$ for every $c \in C$, and (iii) $\mu(s) = c$ if and only if $s \in \mu(c)$. Given a matching μ, we say that it is **blocked** by pair (s, c) if s prefers c to $\mu(s)$ and either (i) c prefers s to some $s' \in \mu(c)$ or (ii) $|\mu(c)| < q_c$ and s is acceptable to c. A matching μ is **individually rational** if for each student $s \in S$, we have that $\mu(s)R_ss$ and for each $c \in C$ and each $s \in \mu(c)$, we have that $sP_c\emptyset$. A matching μ is **stable** if it is individually rational and is not blocked. A mechanism is a systematic way of assigning students to colleges. A stable mechanism is a mechanism that yields a stable matching for any college admissions problem.

Gale and Shapley (1962) introduce the following student-proposing deferred acceptance algorithm:

Round 1: Each student applies to her first choice college. Each college rejects the lowest-ranking students in excess of its capacity and all unacceptable students among those who applied

[3] The preference relation over sets of students is responsive if, whenever $S' = S'' \cup \{s\} \backslash \{s''\}$ for some $s'' \in S''$ and $s \notin S''$, college c prefers S' to S'' if and only if college c prefers s to s''.

to it, keeping the rest of students temporarily (so students not rejected at this step may be rejected in later steps.)

In general,

Round k: Each student who was rejected in Round k-1 applies to her next highest choice (if any). Each college considers these students *and* students who are temporarily held from the previous step together, and rejects the lowest-ranking students in excess of its capacity and all unacceptable students, keeping the rest of students temporarily (so students not rejected at this step may be rejected in later steps.)

The algorithm terminates either when every student is matched to a college or every unmatched student has been rejected by every acceptable college. Since there are a finite number of students and colleges, the algorithm terminates in a finite number of steps.

Gale and Shapley (1962) show that this procedure results in a stable matching that each student weakly prefers, the student-optimal stable matching, to any other stable matching. We refer to the mechanism employing this algorithm as the student-optimal stable mechanism and denote it as $GS^{\mathcal{S}}$. Dubins and Freedman (1981) and Roth (1982) show that truth-telling is a dominant strategy for each student under $GS^{\mathcal{S}}$.

Responsiveness of college preferences allows us to define a college-proposing variant of the deferred acceptance algorithm, which yields the most preferred stable matching for colleges.[4] We refer to this variant of the mechanism as $GS^{\mathcal{C}}$.

While truth-telling is a dominant strategy for each student under $GS^{\mathcal{S}}$, an analogous result does not hold for colleges under $GS^{\mathcal{C}}$. In fact, there is no stable mechanism where truth-telling is a dominant strategy for colleges in the college admissions model (Roth 1985). The following example illustrates this possibility.

Example 1. There are two students, s_1 and s_2, and two colleges, c_1 and c_2, where c_1 has two seats and c_2 has one seat. The preferences are:

$$R_{s_1} : c_1, c_2, s_1 \qquad R_{c_1} : \{s_1, s_2\}, \{s_2\}, \{s_1\}, \emptyset$$
$$R_{s_2} : c_2, c_1, s_2 \qquad R_{c_2} : \{s_1\}, \{s_2\}, \emptyset.$$

[4]Responsiveness is not necessary to define the college-proposing version of deferred acceptance (Kelso and Crawford 1982).

The only stable matching for this problem is:

$$\begin{pmatrix} s_1 & s_2 \\ c_1 & c_2 \end{pmatrix},$$

which means that student s_1 is matched to college c_1 and student s_2 is matched to college c_2.

Now suppose college c_1 submits the manipulated preference R'_{c_1} where only student s_2 is acceptable. With this report, the only stable matching is:

$$\begin{pmatrix} s_1 & s_2 \\ c_2 & c_1 \end{pmatrix}.$$

Hence college c_1 benefits by manipulating its preferences under any stable mechanism (including the college-optimal stable mechanism).

We are ready to present our first result. Fix all student preferences so that only colleges can potentially manipulate.

Proposition 1. *The student-optimal stable mechanism ($GS^{\mathcal{S}}$) is strongly more manipulable than the college-optimal stable mechanism ($GS^{\mathcal{C}}$).*

Proof. Fix student preferences, let P denote college preferences, and let P_{-c} denote the preferences of colleges other than college c. Suppose there is some college c and preference $\hat{P}_c$ such that

$$GS^{\mathcal{C}}_c(\hat{P}_c, P_{-c}) \;\; P_c \;\; GS^{\mathcal{C}}_c(P). \tag{1}$$

We want to show that there exists some $\tilde{P}_c$ such that

$$GS^{\mathcal{S}}_c(\tilde{P}_c, P_{-c}) \;\; P_c \;\; GS^{\mathcal{S}}_c(P).$$

First, by Gale and Shapley (1962), the college-optimal stable matching is weakly more preferred by colleges than the student-optimal stable matching:

$$GS^{\mathcal{C}}_c(P) \;\; R_c \;\; GS^{\mathcal{S}}_c(P). \tag{2}$$

Construct $\tilde{P}_c$ as follows: for any $s \in S$,

$$s \tilde{P}_c \emptyset \;\Leftrightarrow\; s \in GS^{\mathcal{C}}_c(\hat{P}_c, P_{-c}).$$

That is, only students in $GS^{\mathcal{C}}_c(\hat{P}_c, P_{-c})$ are acceptable to college c under $\tilde{P}_c$.

Since matching $GS^{\mathcal{C}}(\hat{P}_c, P_{-c})$ is stable under $(\hat{P}_c, P_{-c})$, it is also stable under $(\tilde{P}_c, P_{-c})$. Moreover by Roth (1984), college c is assigned the same number of students at any stable matching under profile $(\tilde{P}_c, P_{-c})$. Since only students in $GS^{\mathcal{C}}_c(\hat{P}_c, P_{-c})$ are acceptable to college c under $\tilde{P}_c$, we have

$$GS^{\mathcal{S}}_c(\tilde{P}_c, P_{-c}) = GS^{\mathcal{C}}_c(\hat{P}_c, P_{-c}). \tag{3}$$

Hence, by (1), (2), and (3), we have

$$\underbrace{GS^{\mathcal{C}}_c(\hat{P}_c, P_{-c})}_{=GS^{\mathcal{S}}_c(\tilde{P}_c, P_{-c})} \; P_c \; GS^{\mathcal{C}}_c(P) \; R_c \; GS^{\mathcal{S}}_c(P),$$

which shows that college c can manipulate $GS^{\mathcal{S}}$ with report $\tilde{P}_c$.

Finally, we describe a problem where $GS^{\mathcal{C}}$ is not manipulable by any college, while some college can manipulate $GS^{\mathcal{S}}$. Suppose there are two students, s_1 and s_2, and two colleges, c_1 and c_2, each with one seat. The student and college preferences are

$$\begin{array}{ll} R_{s_1} : c_1, c_2, s_1 & \quad R_{c_1} : \{s_2\}, \{s_1\}, \emptyset \\ R_{s_2} : c_2, c_1, s_2 & \quad R_{c_2} : \{s_1\}, \{s_2\}, \emptyset. \end{array}$$

Since each college obtains her top choice under $GS^{\mathcal{C}}$, no college can manipulate. However, if college c_1 declares that only s_2 is acceptable, it can manipulate $GS^{\mathcal{S}}$. This completes the proof. □

The same argument can be used to generalize this result in a few directions. We state these here and omit the complete proofs.

Let φ be an arbitrary stable mechanism. Then

a) φ is strongly more manipulable than $GS^{\mathcal{C}}$ for colleges,

b) $GS^{\mathcal{S}}$ is strongly more manipulable than φ for colleges, and

c) $GS^{\mathcal{C}}$ is strongly more manipulable than φ for students.

This result is also related to the recent policy discussion about the reforms of the National Resident Matching Program (NRMP), the job market clearinghouse that annually fills more than 25,000 jobs for new physicians in the United States. Prior to 1998, the mechanism was inspired by the college-proposing deferred acceptance algorithm. As we have discussed, in the college-optimal stable mechanism truth-telling is not a dominant strategy for students or colleges. In the mid-1990s, the NRMP came under increased scrutiny by students and their advisors who believed that the NRMP did not function in the best interest of students and was open to the possibility of different kinds of strategic behavior (Roth and Rothblum 1999). The mechanism was changed to one based on the student-proposing deferred acceptance algorithm (Roth and Peranson 1999).[5] One reason for this change was that truth-telling is a dominant strategy for students. In contrast, the result states that the student-optimal stable mechanism is the most manipulable stable matching mechanism for colleges.

3.2 School choice mechanisms

The college admissions model is closely related to another model introduced by Abdulkadiroğlu and Sönmez (2003), known as the school choice problem. In this model, students are the only players and school seats are objects to be consumed even though for each school there is a priority ranking of students. The priority ranking of students has the same mathematical structure as the preferences of colleges in the college admissions model. Let $\pi = (\pi)_{c \in C}$ denote the school priorities. For any school c, the function $\pi_c : \{1, ..., n\} \rightarrow \{s_1, ..., s_n\}$ is the priority ordering at school c, where $\pi_c(1)$ indicates the student with the highest priority, $\pi_c(2)$ indicates the student with the second highest priority and so on.

Vulnerability of school choice mechanisms to manipulation has played a major role in the adoption of new student assignment procedures in Boston and New York City (see Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005) and Abdulkadiroğlu, Pathak, and Roth (2005)). For instance, in Fall 2003, the New York City Department of Education changed their assignment mechanism to one based on the student-proposing deferred acceptance algorithm. One

[5]This reform was mimicked in a number of other clearinghouses. A comprehensive list of 43 clearinghouses is presented in Table 1 in Roth (2008).

of the features of this system is that it only allows students to submit a rank order list of their top 12 choices. Based on the strategy-proofness of the student-optimal stable mechanism, the following advice was given to students:

> You must now rank your 12 choices according to your true preferences.

However, for a student who has more than 12 acceptable schools, truth-telling is no longer a dominant strategy. In practice, between 20 to 30 percent of students rank 12 schools.[6] This issue was first theoretically investigated by Haeringer and Klijn (2007).

Our next result formalizes the idea that the greater the number of choices a student can make, the less vulnerable the constrained version of student-optimal stable mechanism is to manipulation. Let GS be the student-optimal stable mechanism, and GS^k be the constrained version of the student-optimal stable mechanism where only the top k choices are considered.

Proposition 2. *Let $\ell > k > 0$ and suppose there are at least ℓ schools. Then GS^k is weakly more manipulable than GS^ℓ.*[7]

Proof. Suppose there is a student i and preference $\hat{P}_i$ such that

$$GS_i^\ell(\hat{P}_i, P_{-i}) \ P_i \ GS_i^\ell(P). \tag{4}$$

For any student j, let P_j^ℓ be the truncation of P_j after the ℓ^{th} choice. This means that in P_j^ℓ any choice after the top ℓ in P_j are unacceptable, and choices among the top ℓ are ordered according to P_j. Observe that relation (4) implies that

$$GS_i(\hat{P}_i^\ell, P_{-i}^\ell) \ P_i \ GS_i(P^\ell). \tag{5}$$

Since GS is strategy-proof, relation (5) implies that student i does not receive one of her top ℓ choices from the GS mechanism under profile P^ℓ. Hence, $GS_i(P^\ell) = GS_i^\ell(P) = i$.

[6] These details together with the entire description of the new assignment procedure is contained in Abdulkadiroğlu, Pathak and Roth (2008).

[7] We thank an NSF reviewer who pointed out that Proposition 2 is implied by Theorem 6.5 of Haeringer and Klijn (2007) which states that the set of Nash equilibria of the preference revelation game induced by GS^k is a subset of the set of Nash equilibria of the preference revelation game induced by GS^ℓ.

For $k < \ell$, there are two cases to consider.

Case 1: $GS_i^k(P) = i$.

Let $GS_i^\ell(\hat{P}_i, P_{-i}) = c$ and let $\tilde{P}_i$ be such that c is the only acceptable school.

Claim: $GS_i^k(\tilde{P}_i, P_{-i}) = c$.

Proof: First note that $GS_i^\ell(\tilde{P}_i, P_{-i}) = c$. Moreover, by definition

$$GS^\ell(\tilde{P}_i, P_{-i}) = GS(\tilde{P}_i, P_{-i}^\ell) \qquad \text{and} \qquad GS^k(\tilde{P}_i, P_{-i}) = GS(\tilde{P}_i, P_{-i}^k).$$

Gale and Sotomayor (1985) (see also Theorem 5.34 of Roth and Sotomayor 1990) implies that

$$GS_i(\tilde{P}_i, P_{-i}^k) \ R_i \ GS_i(\tilde{P}_i, P_{-i}^\ell).$$

Substituting the definitions,

$$GS_i^k(\tilde{P}_i, P_{-i}) \ R_i \ \underbrace{GS_i^\ell(\tilde{P}_i, P_{-i})}_{=c}.$$

Since c is the only acceptable school in $\tilde{P}_i$, the claim follows. ⋄

Thus, in the first case, student i can manipulate GS^k:

$$\underbrace{GS_i^k(\tilde{P}_i, P_{-i})}_{=c} \ P_i \ \underbrace{GS_i^k(P)}_{=i}.$$

Case 2: $GS_i^k(P) \neq i$.

Claim 1: $\exists j \in S$ such that $GS_j^k(P) = j$ although $GS_j^\ell(P) \neq j$.

Proof: Suppose not. Then since $GS_i^\ell(P) = i$ and $GS_i^k(P) \neq i$, there is a college who is assigned strictly more students under $GS^k(P)$ than $GS^\ell(P)$. This is a contradiction to Gale and Sotomayor (1985), which requires that each colleges is weakly worse off under GS^k (since profile P^k is a truncation of profile P^ℓ). ⋄

Pick any $j \in S$ such that $GS_j^k(P) = j$ although $GS_j^\ell(P) \neq j$. Let $GS_j^\ell(P) = c$ and let $\tilde{P}_j$ be such that c is the only acceptable school.

Claim 2: $GS_j^k(\tilde{P}_j, P_{-j}) = c$.

Proof: Since $GS_j^\ell(P) = c$, we have $GS_j^\ell(\tilde{P}_j, P_{-j}) = c$ as well. Moreover, by definition

$$GS^{\ell}(\tilde{P}_j, P_{-j}) = GS(\tilde{P}_j, P^{\ell}_{-j}) \quad \text{and} \quad GS^{k}(\tilde{P}_j, P_{-j}) = GS(\tilde{P}_j, P^{k}_{-j}).$$

Gale and Sotomayor (1985) implies that

$$GS_j(\tilde{P}_j, P^{k}_{-j}) \ R_j \ GS_j(\tilde{P}_j, P^{\ell}_{-j}).$$

Substituting the definitions,

$$GS^k_j(\tilde{P}_j, P_{-j}) \ R_j \ \underbrace{GS^{\ell}_j(\tilde{P}_j, P_{-j})}_{=c}.$$

Since c is the only acceptable school in $\tilde{P}_j$,

$$GS^k_j(\tilde{P}_j, P_{-j}) = c,$$

which establishes the claim. ◇

Thus, for the second case, student j can manipulate GS^k:

$$\underbrace{GS^k_j(\tilde{P}_j, P_{-j})}_{=c} \ P_j \ \underbrace{GS^k_j(P)}_{=j}.$$

Finally, we describe a problem where GS^{ℓ} is not manipulable by any students, but GS^k is manipulable by some student. Suppose there are two students, s_1 and s_2, and two schools, c_1 and c_2, each with one seat. The students have identical preferences which rank c_1 ahead of c_2 and both schools have identical priority orderings: s_1 is ordered ahead of s_2. Under GS^2, no student can manipulate because each obtains her top or second choice and GS is strategy-proof. Under GS^1, s_2 is unassigned, and can benefit from ranking c_2 as her top choice. This example can be generalized to the case of GS^k and GS^{ℓ}. This completes the proof. □

The result of Proposition 2 does not extent to strong manipulability as the following example illustrates:

Example 2. There are three students, s_1, s_2, and s_3, and three schools, c_1, c_2, and c_3, each with one seat. Suppose that the student preferences and school priorities are:

$$
\begin{aligned}
R_{s_1} &: c_1, c_2, c_3, s_1 & \qquad \pi_{c_1} &: s_2, s_3, s_1 \\
R_{s_2} &: c_2, c_3, c_1, s_2 & \qquad \pi_{c_2} &: s_2, s_3, s_1 \\
R_{s_3} &: c_2, c_1, c_3, s_3 & \qquad \pi_{c_3} &: s_1, s_2, s_3.
\end{aligned}
$$

We show that while student s_1 can manipulate GS^2, she cannot manipulate GS^1. The outcome of GS^2 is:

$$\begin{pmatrix} s_1 & s_2 & s_3 \\ s_1 & c_2 & c_1 \end{pmatrix}.$$

Student s_1 is unassigned, and she can manipulate by declaring that c_3 is her only acceptable school. The outcome of GS^1 is:

$$\begin{pmatrix} s_1 & s_2 & s_3 \\ c_1 & c_2 & s_3 \end{pmatrix}.$$

Since student s_1 obtains her top choice, she cannot manipulate GS^1.

3.3 The Boston mechanism

Another widely studied and popular mechanism for the school choice problem is the Boston mechanism. From July 1999 to July 2005, the mechanism has been used by school authorities in Boston to assign over 75,000 students to public school. Variants of the mechanism have been used in many different US school districts including: Cambridge MA, Charlotte-Mecklensburg NC, Denver CO, Miami-Dade FL, Minneapolis MN, Providence RI, and Tampa-St. Petersburg FL.

For given student preferences and school priorities, the outcome of the Boston mechanism is determined with the following procedure:

Round 1: Only the first choices of students are considered. For each school, consider the students who have listed it as their first choice and assign seats of the school to these students one at a time following their priority order until there are no seats left or there is no student left who has listed it as her first choice.

In general,

Round k: Consider the remaining students. In Round k, only the k^{th} choices of these students are considered. For each school with still available seats, consider the students who have listed it as their k^{th} choice and assign the remaining seats to these students one at a time following their priority order until there are no seats left or there is no student left who has listed it as her k^{th} choice.

The procedure terminates when each student is assigned a seat at a school.

The Boston mechanism is vulnerable to preference manipulation. Loosely speaking, the Boston mechanism attempts to assign as many students as possible to their first choice school, and only after all such assignments have been made does it consider assignments of students to their second choices, and so on. If a student is not admitted to her first choice school, her second choice may be filled with students who have listed it as their first choice. That is, a student may fail to get a place in her second choice school that would have been available had she listed that school as her first choice.

Some families understand these features of the Boston mechanism and have developed rules of thumb for submitting preferences strategically. See, for instance, the description of the strategies employed by the West Zone Parents Group in Boston in Pathak and Sönmez (2008). Similar heuristics have developed in other school districts as well (see Ergin and Sönmez 2006 for more examples). Finally, in controlled experiments, Chen and Sönmez (2006) show that more than 70% of participants in their experiment do not reveal their preferences truthfully under the Boston mechanism. Of course, the Boston mechanism is strongly more manipulable than the student-optimal stable mechanism, which is strategy-proof.

In practice, many school districts using mechanisms based on the Boston mechanism limit the number of schools that participants may rank. In Providence Rhode Island, students may only list two schools, while in Cambridge Massachusetts, students may only list three schools.[8]

[8]See Parent Handbook, Providence Public Schools available at http://www.providenceschools.org/ and Controlled Choice Plan, Cambridge Public Schools, available at http://www.cpsd.us/Web/PubInfo/ControlledChoice.pdf.

Let β be the Boston mechanism and β^k be the Boston mechanism when only the top k choices of students are considered.

Our next result shows that not only is the Boston mechanism more manipulable than the student-optimal stable mechanism, its constrained version is more manipulable than the constrained version of the student-optimal stable mechanism.

Proposition 3. *Suppose there are at least k schools where $k > 1$. Then β^k is weakly more manipulable than GS^k.*

Proof. For any student j, let P_j^k be the truncation of P_j after the k^{th} choice. By definition,

$$\beta^k(P) = \beta(P^k) \qquad \text{and} \qquad GS^k(P) = GS(P^k).$$

Suppose that no student can manipulate β^k. We will show that no student can manipulate GS^k either. Consider two cases:

Case 1: $\beta^k(P) = \beta(P^k)$ is stable under profile P.

Since $\beta(P^k)$ is stable under P, it is stable under P^k as well. Moreover, $GS(P^k)$ is stable for P^k by definition. Since the set of unmatched students across stable matchings is the same (McVitie and Wilson 1970), for all $s \in S$,

$$GS_s(P^k) = s \iff \beta_s(P^k) = s. \tag{6}$$

Pick some $s \in S$. If $GS_s^k(P^k) \neq s$, then student s receives one of her top k choices. This implies that s receives one of her top k choices under GS. Since GS is strategy-proof, student s cannot manipulate GS^k.

Suppose $GS_s^k(P^k) = s$ and s can manipulate. We derive a contradiction. Since s can manipulate, there exists some school c and preference $\hat{P}_s$ such that

$$\underbrace{GS_s^k(\hat{P}_s, P_{-s}^k)}_{=c} \;\; P_s \;\; s.$$

Observe that c is not one of the top k choices of student s under P_s for otherwise student s could manipulate GS. Construct $\tilde{P}_s$ which lists c as the only acceptable school.

Matching $GS^k(\hat{P}_s, P^k_{-s})$ remains stable under $(\tilde{P}_s, P^k_{-s})$ and therefore

$$GS^k_s(\tilde{P}_s, P^k_{-s}) = c.$$

Since $GS(P^k)$ is stable under P^k and $GS^k_s(P^k) = s$ by assumption, relation (6) implies

$$\beta_s(P^k) = s.$$

By Roth (1984), matching $\beta(P^k)$ is not stable under $(\tilde{P}_s, P^k_{-s})$ since student s remains single under $\beta(P^k)$ although not under stable matching $GS^k(\hat{P}_s, P^k_{-s})$. Since matching $\beta(P^k)$ is not stable under $(\tilde{P}_s, P^k_{-s})$, but it is stable for P^k, the only possible blocking pair of $\beta(P^k)$ in $(\tilde{P}_s, P^k_{-s})$ is (s, c). But since $\beta_s(P^k) = s$, this implies that (s, c) also blocks $\beta(P^k)$ under P^k, which is the desired contradiction. Thus, in case 1, no student can manipulate GS^k.

Case 2: $\beta(P^k)$ is not stable for profile P.

In this case, some pair (s, c) blocks $\beta(P^k)$, so that there exists $s' \in \beta_c(P^k)$ such that s obtains higher priority than s' at school c and $cP_s\beta_c(P^k)$.

Construct $\tilde{P}_s$ so that school c is the only acceptable school for student s. Since $s' \in \beta_c(P^k)$ and student s has higher priority than student s' at school c, we must have $s \in \beta_c(\tilde{P}_s, P^k_{-s})$. But this means that

$$\underbrace{\beta_s(\tilde{P}_s, P^k_{-s})}_{=c} \; P_s \; \beta_s(P^k),$$

contradicting the assumption that no student can manipulate β at P^k.

Finally, the following example describes a problem where the constrained version of the Boston mechanism is manipulable although the constrained version of the student-optimal stable mechanism is not. There are three students and three schools each with one seat. The student preferences and school priorities are:

$$\begin{array}{ll} R_{s_1} : c_1, c_2, c_3, i_1 & \qquad \pi_{c_1} : s_1, s_3, s_2 \\ R_{s_2} : c_2, c_3, c_1, i_2 & \qquad \pi_{c_2} : s_3, s_2, s_1 \\ R_{s_3} : c_1, c_2, c_3, i_3 & \qquad \pi_{c_3} : s_3, s_1, s_2. \end{array}$$

The matchings produced by β^2 and GS^2 are:

$$\begin{pmatrix} s_1 & s_2 & s_3 \\ c_1 & c_2 & s_3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} s_1 & s_2 & s_3 \\ c_1 & c_3 & c_2 \end{pmatrix},$$

respectively. Since no student receives an outcome worse than her second choice from GS^2, no student can manipulate GS^2 by the strategy-proofness of GS. On the other hand, student s_3 can manipulate β^2 by declaring that c_2 is her only acceptable school. This example can be generalized to the case of GS^k and β^k. This completes the proof. □

The comparison does not extend to strong manipulability, as shown in the next example.

Example 3. There are three students, s_1, s_2, and s_3, and three schools, c_1, c_2, and c_3, each with one seat. Suppose that the student preferences and school priorities are:

$$\begin{aligned} R_{s_1} &: c_1, c_2, c_3, s_1 & \qquad \pi_{c_1} &: s_2, s_3, s_1 \\ R_{s_2} &: c_2, c_3, c_1, s_2 & \qquad \pi_{c_2} &: s_2, s_3, s_1 \\ R_{s_3} &: c_2, c_1, c_3, s_3 & \qquad \pi_{c_3} &: s_1, s_2, s_3. \end{aligned}$$

We will show that while student s_1 can manipulate GS^2, she cannot manipulate β^2. The matching produced by GS^2 is:

$$\begin{pmatrix} s_1 & s_2 & s_3 \\ s_1 & c_2 & c_1 \end{pmatrix}.$$

Student s_1 can manipulate GS^2 by declaring that c_3 is her top choice. The matching produced by β^2 is:

$$\begin{pmatrix} s_1 & s_2 & s_3 \\ c_1 & c_2 & s_3 \end{pmatrix}.$$

Since student s_1 receives her top choice, she cannot manipulate β^2.

4 Auction mechanisms

4.1 Single unit auctions: k^{th} price auction

Our next application involves a seller who wishes to sell a single unit of an object. There are N bidders and bidder i has value v_i for the object. We assume that the utilities are quasi-linear, so the utility of bidder i who receives the object at a price of p is

$$u_i = v_i - p.$$

We consider k^{th} price sealed-bid auctions pioneered by Vickrey (1961), when $k = 2$ and later examined by Kagel and Levin (1993) for $k = 3$. Under the k^{th} price auction, each player simultaneously bids for the object. The highest bidder receives the object and pays the k^{th} highest price.[9]

Vickrey's second-price auction is perhaps the most well-known strategy-proof mechanism. None of the other k^{th} price auctions are strategy-proof. Hence, they are all strongly more manipulable than the second-price auction. Our next result extends this well-known result:

Proposition 4. *For any $\ell > k \geq 2$, the ℓ^{th} price auction is strongly more manipulable than k^{th} price auction.*

Proof. Fix the bids of all bidders except bidder i. Suppose that bidder i can manipulate the k^{th} price auction. This means that bidder i's valuation is not higher than all of the other bids, and her valuation is higher than the k^{th} highest bid. In this case, she can bid higher than the highest bid and obtain the object at a price of the k^{th} highest bid. The same bidder could also manipulate the ℓ^{th} price auction exactly the same way since her valuation is necessarily higher than the ℓ^{th} highest bid. This shows that any bidder who can manipulate the k^{th} price auction can manipulate the ℓ^{th} price auction as well.

Finally, consider an example where the highest k valuations are the same although the $(k+1)^{\text{th}}$ valuation is strictly lower. In this case, the k^{th} price auction cannot be manipulated, although the ℓ^{th} price auction can be manipulated by any of the bidders who do not receive the object and are one of the highest valuation bidders. □

[9]For the purposes of the exposition, we adopt the convention that all ties are broken in favor of the bidder with the lower index i.

4.2 Auctions for Internet Advertising

The next application we consider involves the model for internet advertising pioneered by Edelman, Ostrovsky and Schwarz (2007) and Varian (2006). When an Internet user enters a search term into an online search engine, she obtains a webpage with search results and sponsored links. The advertisements are ordered on the webpage in different positions, with an advertisement shown at the top of the page more likely to be clicked than one at the bottom of the page. The process by which these advertisement slots are allocated to webpages is currently one of the largest auction markets: in 2005, Google generated more than 6 billion dollars in revenue via their auction mechanism (Edelman et. al 2007).

Our notation and model follow Edelman, Ostrovsky and Schwarz (2007). There are N bidders, and $S < N$ ordered slots on a webpage. Each slot has a *click-through* rate of α_s, where $\alpha_1 > \alpha_2 > \ldots > \alpha_S$. We assume that the click-through rates are common knowledge among bidders. Bidder i has a value of v_i per click. The highest value bidder wins the first slot, the second highest value bidder wins the second slot, and so on. When there are ties, we assume that they are broken with some fixed tie-breaking rule. If bidder i wins slot s and pays price p, then her utility is:

$$u_i = \alpha_s v_i - p.$$

Edelman, Ostrovsky and Schwarz (2007) present a detailed historical overview of the origins of this market. In 1997, Overture introduced an auction for selling Internet advertising. In the original design, each advertiser simultaneously bids for a slot for a particular keyword. The highest bidder receives the first slot at a price of her bid times the click-through rate of slot 1, the second highest bidder receives the second slot at a price of her bid times the click-through rate of slot 2, and so on. Overture's search platform was adopted by major search engines including Yahoo! and MSN. This auction format is known as the Generalized First Price (GFP) auction.

In February 2002, Google introduced its own pay-per-click system, AdWords Select, based on a different payment rule. The highest bidder receives the first slot at a price of the second highest bid times the click-through rate of slot 1, the second highest bidder receives the second slot at a price of the third highest bid times the click-through rate of slot 2, and so on. This auction format has come to be known as the Generalized Second Price (GSP) auction. Once Google introduced this new format, many search engines including Yahoo!/Overture also switched to the GSP.

While neither mechanism is strategy-proof, Edelman, Ostrovsky, and Schwarz (2007) argue that

> The second-price structure makes the market more user friendly and less susceptible to gaming.

Our next result formalizes their insight.

Proposition 5. *The Generalized First Price Auction is strongly more manipulable than the Generalized Second Price Auction.*

Proof. Fix the bids of all bidders except bidder i and order the bids of the others from highest to lowest: $b_1, b_2, \ldots$. Suppose bidder i can manipulate the GSP. We argue that bidder i's valuation has to be at least as large as b_{S-1}, the $(S-1)^{\text{th}}$ highest bid.

There are two cases. First, suppose bidder i's valuation is less than or equal to the S^{th} highest bid. If bidder i declares her true valuation, she receives a payoff of zero. To win a slot, she must declare a value greater than the S^{th} highest bid. Consider such a manipulation, so the bidder obtains the ℓ^{th} slot, where $\ell \in \{1, ..., S\}$. Under the GSP, the payoff from this manipulation is:

$$\alpha_\ell v_i - \alpha_\ell b_\ell \leq 0,$$

since $v_i \leq b_\ell$ for any $\ell \in \{1, ..., S\}$. Therefore, bidder i cannot manipulate the GSP if her value is less than or equal to the S^{th} highest bid.

Next, suppose that bidder i's valuation is less than the $(S-1)^{\text{th}}$ bid, but greater than the S^{th} bid. If bidder i declares her true valuation, she wins slot S and receives payoff of

$$\alpha_S v_i - \alpha_S b_S > 0,$$

since $v_i > b_S$. It is not profitable to manipulate by declaring that her value is less than b_S because she does not win a slot. Suppose she declares her value is greater than b_{S-1}. Consider such a manipulation, so the bidder obtains the ℓ^{th} slot, where $\ell \in \{1, ..., S-1\}$. Under the

GSP, the payoff from this manipulation is:

$$\alpha_\ell v_i - \alpha_\ell b_\ell \leq 0,$$

since $v_i \leq b_\ell$ for any $\ell = \{1, ..., S-1\}$. Therefore, if bidder i can manipulate, her value must be at least as large as b_{S-1}.

If bidder i has a value at least as large as b_{S-1}, then she wins a slot when she truthfully reveals her value. Let the slot she receives when she declares her true valuation be the s^{th} slot. This means that her value v_i is not larger than b_{s-1} but not smaller than b_s. Under the GSP, her payoff is

$$\alpha_s v_i - \alpha_s b_s \geq 0. \tag{7}$$

We claim that bidder i cannot manipulate by reporting a value greater than her true value. If bidder i's value is greater than b_1, she will not affect her payoff by declaring a higher value. Otherwise, suppose she declares her value to obtain slot $\ell \in \{1, ..., s-1\}$ with a greater click-through rate than slot s. Her payoff is

$$\alpha_\ell v_i - \alpha_\ell b_\ell \leq 0,$$

since $v_i \leq b_{s-1} \leq b_\ell$ for any $\ell \in \{1, ..., s-1\}$ and so this manipulation is unprofitable.

Therefore, the only remaining possibility is that bidder i manipulates to win a slot with a lower click-through rate. It is not profitable to report a valuation less than b_S. Suppose instead, that her report is greater than b_ℓ but less than $b_{\ell-1}$ for some $\ell \in \{s+1, ..., S\}$. In this case, she wins slot ℓ and her payoff is

$$\alpha_\ell v_i - \alpha_\ell b_\ell \geq 0.$$

Without further assumptions on values and click-through rates, this manipulation might be profitable. Suppose that values and click-through rates are such that for some $\ell \in \{s+1, ..., S\}$,

$$\alpha_\ell(v_i - b_\ell) > \alpha_s(v_i - b_s) \geq 0 \tag{8}$$

where $\alpha_s(v_i - b_s) \geq 0$ from inequality (7). In this case bidder i can profitably manipulate by submitting a bid to win the ℓ^{th} slot.

Under the GFP, when this bidder declares her true value, she obtains a payoff of zero. Consider the deviation where she reports her true value to be less than v_i such that she wins the ℓ^{th} slot for some $\ell \in \{s+1, ..., S\}$. For simplicity, let her report be $b_\ell + \epsilon$ for some $\epsilon > 0$ small.

Under the GFP, her payoff from this report is

$$\alpha_\ell v_i - \alpha_\ell (b_\ell + \epsilon) = \alpha_\ell(v_i - b_\ell) - \alpha_\ell \epsilon.$$

Since $\alpha_\ell(v_i - b_\ell) > 0$ from inequality (8), there exists an ϵ small enough that her GFP payoff is strictly greater than zero.

To complete the proof, we describe a problem where some bidder can manipulate the GFP, but no bidder can manipulate the GSP. Suppose $v_1 > v_2 = \ldots = v_S = v_{S-1} > v_{S-2} > \ldots > v_N$. First we show that no bidder can manipulate the GSP. Under the GSP, the highest value bidder's payoff from reporting her true valuation is $\alpha_1(v_1 - v_2)$. If she reports her value to be v_2, she obtains a zero payoff. If she reports her value to be less than v_2, she also obtains zero payoff, so she cannot manipulate the GSP. Any bidder with value equal to v_2 cannot manipulate the GSP because if she still obtains a slot after manipulation, the next highest report is unchanged, so her payoff is zero. Moreover, if such a bidder manipulates with a report where she does not win a slot, she also receives zero. Therefore, no bidder can manipulate the GSP. In contrast, in the GFP, suppose the highest value bidder declares her value to be less than v_1 but greater than v_2. With this report, she wins the first slot, but pays a lower price for the first slot than she would if she reported her true value. This shows that the highest value bidder can manipulate the GFP. □

4.3 Multi-unit auctions

Unlike the two previous examples, the next model we consider involves the auctioning of multiple units of identical objects. The US Treasury's bond issue auctions, auctions for electricity and other commodities, and financial market auctions such as the opening batch auctions at the NYSE, Paris, and Amsterdam exchanges are examples of auctions involving multiple identical objects.[10] We are interested in comparing two sealed-bid auction formats. In each format, a bidder is asked to submit bids for each of the k units indicating how much she is willing to pay for each unit.

In the discriminatory format, also known as the pay-your-bid auction, each bidder pays an amount equal to the sum of her bids that are winning bids. The discriminatory auction is a natural multi-unit extension of the first-price sealed bid auction. Milton Friedman (1960) initially proposed a uniform-price auction, where all k units are sold at a "market-clearing" price such that the total amount demanded is equal to the total amount supplied.

Formally, a seller wishes to sell k of identical items to N bidders, where we assume $N \geq k$. The bidders are asked to report their valuations for the k objects, where v_i^ℓ is bidder i's valuation for the ℓ^{th} unit of the item for sale. In both auctions we consider, a total of N k-dimensional reports are collected, and the k units are awarded to the bidders with the k highest reported valuations.[11]

The utility of bidder i who wins ℓ objects at a total price of p is:

$$u_i = v_i^1 + ... + v_i^\ell - p.$$

We will assume that marginal values are declining for each bidder: $v_i^1 \geq v_i^2 \geq ... \geq v_i^k$.

The two payment rules we consider are:

- *discriminatory auction*: for the units awarded, the bidder pays the value declared for each unit.
- *uniform-price auction*: for the units awarded, the bidder pays the $(k+1)^{\text{th}}$ highest value for each unit.[12]

The US Treasury has employed a discriminatory format since 1929 for the sale of short-term treasury securities. Since the 1970s, the US Treasury also employed a discriminatory format to auction Treasury bonds. In 1992, the US Treasury switched to a uniform-price auction for 2 and

[10] See Krishna (2002) for more examples and discussion.

[11] For both the discriminatory and uniform-price auction, we adopt the convention that when there is a tie, it is broken in favor of the bidder with fewer units, and if the bidders have the same number of units, it is broken in favor of the bidder with the lower index i.

[12] It is possible to consider other "market clearing" rules such as paying the k^{th} value or paying a value between the k^{th} and $(k+1)^{\text{th}}$ value. The comparison between formats is not sensitive to this choice.

5 year notes. Throughout these policy changes, the Treasury has been influenced by a number of arguments. One of the most influential arguments is from Milton Friedman's testimony to the Joint Economic Committee of the US Congress in 1959. In this testimony, Friedman argued that a uniform-price format levels the playing field by reducing the importance of specialized knowledge among dealers. According to Friedman, more bidders would be induced to bid directly in uniform-price auctions because the fear of being awarded securities at too high a price is eliminated. Merton Miller supported this argument stating, "All of that [gaming] is eliminated if you use the [uniform-price] auction. You just bid what you think it's worth." A US government report issued around that time jointly signed by the Treasury Department, SEC, and Federal Reserve Board states: "Moving to a uniform-price award method permits bidding at the auction to reflect the true nature of investor preferences."[13]

Both the discriminatory and uniform-price auctions are not strategy-proof. In particular, in both formats, bidders have an incentive to shade their bids. In a discriminatory auction, bidders have an incentive to report that their bids are just above the lowest bid that wins a unit. In a uniform-price auction, a bidder has an incentive to shade her bid for the units other than the first one because these bids have the potential to influence the market-clearing price if she wins. This "demand-reduction" feature of the uniform-price auction prevents it from being strategy-proof.

The next proposition supports Milton Friedman's original argument about the incentive properties of the uniform-price auction relative to the discriminatory auction.

Proposition 6. *The discriminatory auction is strongly more manipulable than the uniform-price auction.*

Proof. Fix the reports of every bidder except for bidder i. Let $b_1, b_2, ...$ be these per unit bids ordered from highest to lowest.

The first case we consider is if bidder i's highest value for a unit is less than the k^{th} highest bid of the other bidders (e.g., $v_i^1 \leq b_k$).

If bidder i reports her true values, then in a uniform-price auction, the market-clearing price is $\max\{v_i^1, b_{k+1}\}$. Since bidder i's highest value is less than k of the other bids, if $v_i^1 < b_k$, she

[13] For more discussion on the influence of Friedman's argument, see Malvey, Archibald and Flynn (1995) and Ausubel and Cramton (2002).

does not obtain any units in a uniform-price auction and if $v_i^1 = b_k$ she only wins a unit if she wins the tie-breaking. In either case, her payoff is 0.

Bidder i could manipulate to win n units by reporting that n of her valuations are greater than b_{k-n+1}. If $n < k$, denote her declared valuation for the $(n+1)^{\text{th}}$ unit by $\hat{v}_i^{n+1}$. This manipulation yields a market-clearing price of $\tilde{p} = \max\{\hat{v}_i^{n+1}, b_{k-n+1}\}$. The payoff from this manipulation is:

$$(v_i^1 + \ldots + v_i^n) - n \cdot \tilde{p}.$$

Since $v_i^1 \leq b_k$ and $b_k \leq \tilde{p}$, this is not profitable.

The next case we consider is when bidder i's m highest values where $m \in \{1, ..., k\}$ are among the top k highest bids (e.g., $v_i^m \geq b_{k-m+1}$ and if $m \neq k$, $v_i^{m+1} \leq b_{k-m+1}$).

If bidder i reports her true values, then in a uniform-price auction, the market-clearing price is $p = \max\{v_i^{m+1}, b_{k-m+1}\}$ and bidder i wins m units. Her profit is

$$(v_i^1 + \ldots + v_i^m) - m \cdot p \geq 0. \tag{9}$$

We argue that a manipulation by bidder i in the uniform-price auction to win more than m units is unprofitable. If $m = k$, bidder i wins all units and the market-clearing price is set by another bidder. In this case, it is impossible to win more units. When $m \neq k$, suppose bidder i reports bids such that she wins $n > m$ units where $n \leq k$ and if $n < k$, declaring that her $(n+1)^{\text{th}}$ highest value is $\hat{v}_i^{n+1}$. In this case, in a uniform-price auction, the market-clearing price is $\hat{p} = \max\{\hat{v}_i^{n+1}, b_{k-n+1}\}$ and bidder i wins n units. Her profit is

$$(v_i^1 + \ldots + v_i^n) - n \cdot \hat{p}.$$

We will show that

$$(v_i^1 + \ldots + v_i^m) - m \cdot p \geq (v_i^1 + \ldots + v_i^n) - n \cdot \hat{p},$$

which demonstrates this manipulation is not profitable.

First, since $n > m$, we have that $b_{k-n+1} \geq b_{k-m+1}$ and $b_{k-m+1} \geq v_i^{m+1}$. As a result,

$$\hat{p} \geq p,$$

or the manipulation leads to a weakly higher market-clearing price. Hence, for $j = 1, ..., m$:

$$v_i^j - p \geq v_i^j - \hat{p}.$$

For each of the first m units, when the bidder manipulates, she receives no additional utility because the price is higher. Next, for $j = m+1, ..., n$, we have that

$$v_i^j - \hat{p} \leq 0,$$

since $v_i^{m+1} \leq b_{k-m+1} \leq b_{k-n+1}$. For each extra unit that the bidders wins when she manipulates, she loses utility. Therefore, a manipulation by bidder i to win more than m units is not profitable in the uniform-price auction.

The last remaining case is if bidder i manipulates to win fewer units. Suppose bidder i reports bids such that she wins $\ell < m$ units and that her valuation for the $(\ell+1)^{\text{th}}$ unit is $\hat{v}_i^{\ell+1}$. In this case, in a uniform-price auction, the market-clearing price is $p' = \max\{\hat{v}_i^{\ell+1}, b_{k-\ell+1}\}$ and bidder i obtains ℓ units. Her profit is

$$(v_i^1 + \ldots + v_i^\ell) - \ell \cdot p'.$$

Suppose the values are such that

$$(v_i^1 + \ldots + v_i^\ell) - \ell \cdot p' > (v_i^1 + \ldots + v_i^m) - m \cdot p. \tag{10}$$

In this case, it is profitable for bidder i to manipulate and win fewer units than she would if she reported her true values. In a discriminatory auction, if bidder i reported her true valuations, she receives m units. Since she pays her reported value for each unit, her payoff is zero. Suppose that bidder i manipulates by reporting that her ℓ highest values are equal to $b_{k-\ell+1}$ and all other values are less than $b_{k-\ell+1}$. In a discriminatory auction, with this report, she wins ℓ units, and pays $b_{k-\ell+1}$ for each unit. Her profit is

$$(v_i^1 + \ldots + v_i^\ell) - \ell \cdot b_{k-\ell+1}.$$

Since $b_{k-\ell+1} \leq \max\{\hat{v}_i^{\ell+1}, b_{k-\ell+1}\} = p'$, equations (9) and (10) imply that

$$(v_i^1 + \ldots + v_i^\ell) - \ell \cdot b_{k-\ell+1} > 0.$$

Therefore, bidder i also finds it profitable to manipulate the discriminatory auction.

To complete the proof, we describe a problem where no bidder can manipulate the uniform-price auction, while some bidder can manipulate the discriminatory format. Suppose that v_1^1 is strictly larger than any other valuation of the other bidders, all valuations of all bidders other than bidder 1 are equal to $\bar{v}$, and $\bar{v} > v_1^2 > \ldots > v_1^k$. Under the uniform-price auction, when bidders report truthfully, each bidder wins a unit. If bidder 1 wishes to win an additional unit, she must pay $\bar{v}$ which is less than her marginal value for the unit. No other bidder would find it profitable to win more than one unit because if a bidder reported any of her values to be greater than $\bar{v}$, she would have to pay at least $\bar{v}$ for each unit she wins. Therefore, no bidder can manipulate the uniform-price auction. In contrast, bidder 1 would find it profitable to report that her value for the first unit is lower than v_1^1, but not lower than $\bar{v}$. In this manipulation, bidder 1 wins one unit, but pays a lower price than what she would pay if she reported her true valuations.

□

5 Conclusion

While strategy-proofness is a very plausible property of a mechanism, it is at the same time very demanding. One goal of this paper is to develop an approach to comparing mechanisms that are not strategy-proof based on the incentives they generate. We believe that our approach complements other approaches, confirms existing views on the manipulability of mechanisms, and is useful in many prominent applications. The applications we have considered in this paper are problems from matching and auctions. We are hopeful about the possibility to investigate the concept in other settings where manipulation has been studied such as in social choice and political economy or other resource allocation contexts, such as those involving public goods or cost-sharing.

Another goal of this paper is to make specific comparisons between mechanisms that have found practical use in real-world allocation problems. Many of our applications were motivated by problems from the recent "market-design" literature. Our first result is inspired by the reform of the National Residency Matching Program, the second result is motivated by the new student assignment system in New York City, and the third result provides a way to formalize the idea that the Boston mechanism is a highly manipulable mechanism. The examples of the internet advertisement auctions and the multi-unit auctions are also cases where ideas from economics have inspired the design of actual mechanisms. In situations like these, where providing straightforward incentives may be desirable, our results may serve as another factor in deciding between mechanisms.

References

[1] Abdulkadiroğlu, Atila and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93: 729-747.

[2] Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez. 2005. "The Boston Public Schools Match." *American Economic Review*, Papers and Proceedings, 96: 368-371.

[3] Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2005. "The New York City High School Match." *American Economic Review*, Papers and Proceedings, 96: 364-367.

[4] Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2008. "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match." *American Economic Review*, forthcoming.

[5] Ausubel, Larry and Peter Cramton. 2002. "Demand Reduction and Inefficiency in Multi-Unit Auctions." *Unpublished mimeo*, University of Maryland.

[6] Chen, Yan and Tayfun Sönmez. 2006. "School Choice: An Experimental Study." *Journal of Economic Theory*, 127: 202-231.

[7] Dasgupta, Partha and Eric Maskin. 2008. "On the Robustness of Majority Rule." *Journal of the European Economic Association*, 6: 949-973.

[8] Day, Robert and Paul Milgrom. 2008. "Core-Selecting Auctions." *International Journal of Game Theory*, 36: 393-407.

[9] Dubins, Lester E. and David A. Freedman. 1991. "Machiavelli and the Gale-Shapley Algorithm." *American Mathematical Monthly*, 88: 485-494.

[10] Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review*, 97: 242-259.

[11] Ergin, Haluk. 2002. "Efficient Resource Allocation on the Basis of Priorities." *Econometrica*, 70: 2489-2497.

[12] Ergin, Haluk and Tayfun Sönmez. 2006. "Games of School Choice Under the Boston Mechanism." *Journal of Public Economics*, 90: 215-237.

[13] Friedman, Milton. 1960. A Program for Monetary Stability, Fordham University Press.

[14] Gale, David and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69: 9-15.

[15] Gale, David and Marilda Sotomayor. 1985. "Some Remarks on the Stable Matching Problem." *Discrete Applied Mathematics*, 11: 223-232.

[16] Gootman, Elissa. 2006. "Lafayette Among 5 High Schools to Close." *New York Times*, Metropolitan Desk, page 3, Published on December 14.

[17] Haeringer, Guillaume and Flip Klijn. 2007. "Constrained School Choice." Barcelona GSE Research Network Working Paper 294.

[18] Kagel, John and Dan Levin. 1993. "Independent-Private-Values Auctions: Bidder Behavior in First-, Second- and Third-Price Auctions with Varying Numbers of Bidders." *Economic Journal*, 103: 868-879.

[19] Kelso, Alexander and Vince Crawford. 1982. "Job Matching, Coalition Formation, and Gross Substitutes." *Econometrica*, 50: 1483-1504.

[20] Kerr, Peter. 2003. "Re: Parents Perplexed by High School Admissions Rule." *New York Times*, Editorial desk, Page 18, Published on November 3.

[21] Krishna, Vijay. 2002. Auction Theory. Academic Press.

[22] Malvey, Paul, Chrstine Archibald, and Sean Flynn. 1995. "Uniform-Price Auctions: Evaluation of the Treasury Experience." Office of Market Finance, U.S Treasury.

[23] McAfee, R. Preston, John McMillan, and Simon Wilkie. 2008. "The Greatest Auction in History" in *Better Living Through Economics*, John Sigfreid, editor, AEA and Harvard University Press, 2008.

[24] McVitie, D.G. and L. B. Wilson. 1970. "Stable Marriage Assignments for Unequal Sets." *BIT*, 10: 295-309.

[25] Milgrom, Paul. 2004. Putting Auction Theory to Work (Churchill Lectures in Economics). Cambridge University Press.

[26] Moulin, Hervé. 1980. "On Strategy-Proofness and Single Peakedness." *Public Choice*, 35: 437-455.

[27] Pathak, Parag A. and Tayfun Sönmez. 2008. "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review*, 98: 1636-1652.

[28] Roth, Alvin E. 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operations Research*, 7: 617-628.

[29] Roth, Alvin E. 1984. "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory." *Journal of Political Economy*, 92: 991-1016.

[30] Roth, Alvin E. 1985. "The College Admissions Problem is Not Equivalent to the Marriage Problem." *Journal of Economic Theory*, 36: 277-288.

[31] Roth, Alvin E. 2002. "Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica*, 70: 1341-1378.

[32] Roth, Alvin E. 2003. "The Origins, History, and Design of the Resident Match." *Journal of the American Medical Association*, 289: 909-212.

[33] Roth, Alvin E. 2008. "Deferred Accpetance Algorithms: History, Theory, Practice, and Open Questions." *International Journal of Game Theory*, Special Issue in Honor of David Gale on his 85th birthday, 36: 537-569.

[34] 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*, 89-4: 748-780.

[35] Roth, Alvin E. and Uriel G. Rothblum. 1999. "Truncation Strategies in Matching Markets - In Search of Advice for Participants." *Econometrica*, 67: 21-43.

[36] Roth, Alvin E. and Marilda Sotomayor. 1990. Two-Sided Matching: A Study in Game Theoretic Modeling and Analysis. *Econometric Society Monographs*, vol. 18, Cambridge University Press.

[37] Varian, Hal R. 2006. "Position Auctions." *International Journal of Industrial Organization.*

[38] Vickrey, W. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, 16: 8-37.