

Comparing Mechanisms by their Vulnerability to Manipulation*

Parag A. Pathak[†] Tayfun Sönmez[‡]

June 19, 2008

Preliminary

Abstract

Measuring the ease of manipulation—deviating from reporting truthfully—is a central issue in social choice, mechanism design and game theory. This paper introduces a new measure of the ease of manipulation for direct revelation mechanisms based on 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 direct mechanisms based on their degree of manipulability. We illustrate the concept by comparing several well-known mechanisms in the matching and auction literature. Our applications include comparisons among stable matching mechanisms, school choice mechanisms, internet advertising mechanisms, and multi-unit auctions.

KEYWORDS: market design, straightforward incentives, matching, auctions

*We thank seminar participants at CEMFI, CORE, ECARES, Princeton, and SED 2008 at Michigan for comments. 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 based on numerous dimensions and potential tradeoffs across these dimensions. Some of the criteria are based on the outcomes of a mechanism. A common objective is whether a mechanism produces an efficient outcome. For instance, in the discussions surrounding the first FCC spectrum auctions in the United States, a mandate was given by then Vice President Al Gore that the auctions should put “licenses into the hands of those who value them the most” (Milgrom 2004).¹ However, efficiency is not the only desirable property of resource allocation systems. Sometimes fairness concerns can trump the desire for efficiency. School policymakers in Boston placed a fairness objective, the *elimination of justified envy*, ahead of efficiency in their adoption of a new student assignment system (Abdulkadiroğlu and Sönmez 2003; Abdulkadiroğlu, Pathak, Roth, Sönmez 2005; 2006). Even in the first US spectrum auction, motivated by fairness arguments, bidding credits and favorable financing terms were given to small businesses and woman- and minority-controlled businesses. Both efficiency, fairness and other objectives based on the outcomes of mechanisms must be evaluated based on what the analyst predicts to be the behavior of participants in the mechanism.

Another set of criteria for evaluating mechanisms is on the process by which its outcomes are achieved. A common objective is to focus on mechanisms that are “simple” or provide “straightforward” incentives. For instance, following the comments of Robert Wilson (1987) that mechanism design theory places too much emphasis on common knowledge assumptions, a critique which has come to be known as the Wilson critique, numerous researchers have examined achieving certain social goals in a way that relaxes these requirements, and does not depend in a critical way on assumptions on what participants know (e.g., Bergemann and Morris 2005, Chung and Ely 2007, Neeman 2004). The most demanding form of requiring straightforward incentives is to focus on strategy-proof mechanisms, which means that truth-telling is a dominant strategy for all participants.

The desire for straightforward incentives, in particular strategy-proofness, has been important in many policy discussions. In problems involving the matching of agents on two sides of a market to one another, the desire to provide straightforward incentives has been central

¹Quoted from Vice President Gore’s speech at the beginning of FCC auction #4. See Section 1.2 of Milgrom (2004), “Designing for Multiple Goals” for a description of many of the objectives of this auction.

in policy discussions regarding the selection of matching mechanisms. In the 1950s, a medical 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 deciding on which procedure to use to place graduating medical students to residencies in the US. Their recommendation was based on possibility of strategic play in the system used to place medical school graduates to hospital residency programs in New York City (Roth 1984, 2003).² Over half a century later, school officials in Boston adopted a strategy-proof mechanism to place elementary, middle, and high school students to schools over their existing system, known as the *Boston mechanism*, because it simplifies the complex problem faced by students who had to determine how to reveal their rank order list over schools to the mechanism. Importantly, the policy change was seen to protect families who were unaware of the strategic aspects of student placement in Boston (Abdulkadiroğlu, Pathak, Sönmez, Roth 2006; Pathak and Sönmez 2007). In New York City’s decision to change their student assignment system which is used to place over 90,000 students to high school every year, a major aim of the policy change 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 simple 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 often supports the selection of the Vickrey’s celebrated second price auction (Vickrey 1961).

Providing straightforward incentives to participants has a number of virtues in real-life allocation problems.³ 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 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,

²Roth (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.

³Indeed, Milgrom (2007) has commented that “simplification is the essence of practical market design.”

where determining whether to expand or shut down a school or residency program.⁴

Unfortunately, strategy-proof mechanisms are known to exist only in particular domains, and in many cases strategy-proof mechanisms perform poorly on other dimensions. 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 focus on direct mechanisms, where participants report their preferences. Our notion of manipulability is based on a comparison of the states under which two mechanisms are manipulable. We introduce two versions of this notion: Under the first notion, mechanism ψ is weakly more manipulable than mechanism φ if whenever an individual can profitably manipulate φ , some individual 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 notion, mechanism ψ is strongly more manipulable than mechanism φ if whenever an individual profitably manipulates φ , she can manipulate ψ as well.

We use our notion to compare several well-known mechanisms in the matching and auction literature. Many of our examples are directly inspired by the recent literature on “market design” (cf. Roth 2002, Milgrom 2007). We begin by considering various classes of matching problems. In many-to-one matching problems (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 student proposing deferred acceptance mechanism when students can only rank k schools is weakly more manipulable than the student proposing deferred acceptance mechanism when participants can only rank ℓ schools for $k < \ell$. Finally, the Boston mechanism when participants can only rank k schools is weakly more manipulable than the student proposing deferred acceptance mechanism when participants can only rank k schools.

We next consider examples from various auction environments. 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 $2 \leq k < \ell$. The last two results are for auctions involving multiple items. In Internet keyword advertisement auctions, where there are multiple items for sale but bidders can only win one item, the Generalized First Price Auction is strongly more manipulable than the Generalized Second Price Auction. In auctions where there are multiple items and bidders

⁴There is some evidence that this is indeed taking place in New York City. See Gootman (2006).

can win more than one item, the discriminatory price 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 auctions. The last section concludes.

2 General Framework

2.1 Primitives

We begin by defining the problem in the abstract. Let the preferences of player i be given by R_i , and let $R = (R_i)$ denote a profile of preferences. Let P_i be the strict counterpart of R_i . The set of allocations is denoted by A . A mechanism is a function φ which maps a preference profile R to an allocation $\varphi(R) \in A$.⁵ Let $\varphi_i(R)$ be the allocation obtained by player i .

We first define what it means for a mechanism to be manipulable.

Definition 1. *A mechanism φ is **manipulable** by agent i at profile R if there exists a preference \hat{R}_i such that*

$$\varphi_i(\hat{R}_i, R_{-i}) P_i \varphi_i(R).$$

This definition allows us to define two versions of our notion of manipulability.

Definition 2. *A mechanism ψ is **weakly more manipulable** than mechanism φ if whenever φ is manipulable, the mechanism ψ is also manipulable, even though the converse does not hold.*

Definition 3. *A mechanism ψ is **strongly more manipulable** than mechanism φ if at any profile R , any agent who can manipulate φ can also manipulate mechanism ψ , even though the converse does not hold.*

If φ is a strategy-proof mechanism and ψ is not, then ψ is both weakly and strongly more manipulable than φ . The interesting and motivating case is when neither ψ nor φ is not

⁵For the purposes of this paper we focus on deterministic mechanisms, but it is easy to extend this notion to stochastic mechanisms, which yield lotteries over allocations.

strategy-proof. If ψ is strongly more manipulable than φ , then ψ is weakly more manipulable than φ . Under the strong notion, the player who finds it beneficial to deviate from truth-telling is the same across the two mechanisms. An equivalent formulation of the weak notion is that if truth-telling is not a Nash equilibrium of mechanism φ , then it is not a Nash equilibrium of mechanism ψ .

Both definitions require that the deviator strictly benefit from the deviation, but neither definition considers the extent to which a player benefits from the manipulation. Finally, both definitions also require that if ψ is more manipulable than φ , there is at least one profile for which ψ is manipulable, but φ is not.

2.2 Related literature

There are many complimentary approaches to comparing the ease of manipulation in mechanisms which are not strategy-proof. We review only a few approaches here. One approach relies on equilibrium analysis, and mechanisms are compared based on the properties of strategies at equilibrium. Another approach studies under what informational or asymptotic assumptions mechanisms which are not strategy-proof have equilibrium where players report the truth, and mechanisms are compared based on the nature of these assumptions (e.g., Roth and Rothblum (1999), Ehlers (2004), Ehlers (2006), Kojima and Pathak (2007)). Researchers also construct measures of the difficulty of finding the optimal strategy. For instance, the computational social choice literature computes the computational complexity of computing a manipulating ballot in an election (e.g., Bartholdi and Orlin (1991)). Finally, another approach is to find domain restrictions under which mechanisms are strategy-proof, such as single-peakedness in voting problems and various conditions including Ergin-acyclicity in priority based resource allocation problems (e.g., Moulin (1980), Ergin (2002), Kesten (2006)). Comparisons between two mechanisms can be made based on the size of these domains.

Our approach is closest in spirit to the literature examining domain restrictions. The main difference is that rather than exogenously imposing domain restrictions, the domains we consider are based on comparisons between two mechanisms. For instance, if the set of preferences where mechanism φ has a Nash equilibrium in truthful strategies is a strict subset of the set of preferences where mechanism ψ has a Nash equilibrium in truthful strategies, then we con-

clude that φ is weakly more manipulable than ψ . Another related contribution is Dasgupta and Maskin (2007) who show that if a voting rule satisfies various axioms for a set of preference profiles, then simple majority voting rule also satisfies those axioms on the same domain. However, their interest is not in making comparisons based on the incentives generated by mechanisms, while this is the central focus here.

3 Matching mechanisms

3.1 Comparing stable matching mechanisms

In the college admissions model (Gale and Shapley 1962), there are a number of students S 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 of schools. Each college has a maximum quota.

Formally, a market is tuple $\Gamma = (S, C, P_S, P_C)$. S and C are sets of students and colleges and $P_S = (P_s)_{s \in S}, P_C = (P_c)_{c \in C}$. 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 quota q_c , and we assume that each college's preferences are responsive. That is, the ranking of a student is independent of her colleagues, and any set of students exceeding quota is unacceptable (see Roth 1985 for a formal definition of responsiveness). Given this assumption, P_c is 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 $cP_s s$. 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 $C \cup S$ to $C \cup S$ 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 (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 $s \in S \cup C$, $\mu(s)R_s\emptyset$ and for each $c \in C$ and each $s \in \mu(c)$, $sP_c\emptyset$. 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 market.

Gale and Shapley (1962) first analyzed 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 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. The algorithm always 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, known as the student-optimal stable matching, to any other stable matching. We refer to the mechanism employing this algorithm as GS^S . Dubins and Freedman (1981) and Roth (1982) show that truth-telling is a dominant strategy for each student under GS^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. We refer to this variant of the mechanism as GS^C .

Unlike the case of students, in the college admissions problem, truth-telling is not a dominant strategy for each college under GS^C . In fact, there is no stable mechanism that is strategy-proof in the college admissions problem (Roth 1985). The following example illustrates this possibility.

Example 1. *There are three colleges c_1, c_2 , and c_3 and four students s_1, \dots, s_4 . Suppose $q_{c_1} = 2$*

and $q_{c_2} = 1$ and $q_{c_3} = 1$, and that the preferences are as follows:

$$P_{c_1} : s_1, s_2, s_3, s_4$$

$$P_{c_2} : s_1, s_2, s_3, s_4$$

$$P_{c_3} : s_3, s_1, s_2, s_4$$

$$P_{s_1} : c_3, c_1, c_2$$

$$P_{s_2} : c_2, c_1, c_3$$

$$P_{s_3} : c_1, c_3, c_2$$

$$P_{s_4} : c_1, c_2, c_3.$$

The college-optimal stable matching (and unique stable matching) is

$$\mu = \begin{pmatrix} c_1 & c_1 & c_2 & c_3 \\ s_3 & s_4 & s_2 & s_1 \end{pmatrix},$$

which means that c_1 is matched to s_3, s_4 and c_2 is matched to s_2 and c_3 is matched to s_1 .

Consider what happens when college c_1 reports $P'_{c_1} = s_1, s_4$. The college-optimal stable matching (and unique stable matching) is

$$\nu = \begin{pmatrix} c_1 & c_1 & c_2 & c_3 \\ s_1 & s_4 & s_2 & s_3 \end{pmatrix},$$

so college c_1 receives assignment $\{s_1, s_4\}$ which it prefers to $\{s_3, s_4\}$ it can profitably manipulate.

The National Resident Matching Program (NRMP) annually fills more than 25,000 jobs for new physicians in the United States. In 1998, under the direction of Alvin E. Roth and Elliot Peranson (Roth and Peranson 1999), one of the reforms of the NRMP changed the algorithm from one based on the college-proposing deferred acceptance algorithm to one based on the student-proposing deferred acceptance algorithm. This reform was mimicked in a number of other residency programs (Roth 2002).⁶

One of the reasons for this change is that those who advised students about the medical job

⁶A more comprehensive list of markets is presented in Table 1 in Roth (2007). The table lists 43 labor markets where the Roth-Peranson clearinghouse design is employed after 1998.

market started reporting in the mid 1990s that many students believed that the NRMP did not function in the best interest of students, and that students were discussing the possibility of different kinds of strategic behavior. In GS^s , students have a weakly dominant strategy to reveal their true preferences. However, as we have seen in the previous example, colleges have an incentive to manipulate either mechanism. Our first result investigates how much a college's incentives to manipulate depend on whether the student-proposing versus college-proposing mechanism is employed.

Proposition 1. *The mechanism GS^S is strongly more manipulable than GS^C for colleges.*

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

$$GS_c^C(\hat{P}_c, P_{-c})P_cGS_c^C(P). \quad (1)$$

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

$$GS_c^S(\tilde{P}_c, P_{-c})P_cGS_c^S(P).$$

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

$$GS_c^C(P)R_cGS_c^S(P). \quad (2)$$

Construct \tilde{P}_c as follows: for any $s \in S$, $s\tilde{P}_c\emptyset \Leftrightarrow s \in GS_c^C(\hat{P}_c, P_{-c})$. That is, only students in $GS_c^C(\hat{P}_c, P_{-c})$ are acceptable under \tilde{P}_c .

Since matching $GS_c^C(\hat{P}_c, P_{-c})$ 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 $GS_c^C(\hat{P}_c, P_{-c})$ are acceptable to college c under \tilde{P}_c , we have

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

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

$$GS_c^S(\tilde{P}_c, P_{-c}) = GS_c^C(\hat{P}_c, P_{-c})P_cGS_c^C(P)R_cGS_c^S(P),$$

which shows that college c can manipulate GS^S .

Example 2.10 in Roth and Sotomayor (1990) exhibits a profile where GS^C is not manipulable by any college, while some college can manipulate GS^S . This completes the proof. \square

Remarks.

1. This theorem can be generalized using the same type of proof in a few ways. Let φ be an arbitrary stable mechanism. Then
 - a) φ is strongly more manipulable than GS^C for colleges,
 - b) GS^S is strongly more manipulable than φ for colleges, and
 - c) GS^C is strongly more manipulable than φ for students.
2. Sönmez (1997) introduced to the literature on college admissions problems the possibility of capacity manipulation, whereby a college can benefit from not revealing its true quota. He shows that there is no stable mechanism which is immune to capacity manipulation. Since any manipulation by preferences and capacities performs at least as well as a manipulation by preferences alone (Kojima and Pathak 2007), it is also possible to generalize the theorem to an environment when colleges report their preferences and quotas and can manipulate both.

3.2 Comparing 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**. The main difference is that in college admissions each school is a (possibly strategic) agent whose welfare matters, whereas in school choice each school is a collection of indivisible goods to be allocated and only the welfare of students is considered.

In the school choice problem, there are a number of students each of whom should be assigned a seat at one of a number of schools. Each student has a strict preference ordering over all schools as well as remaining unassigned and each school has a strict priority ranking of all students. Each school has a maximum capacity.

Formally, a school choice problem consists of a set of students I and schools S with capacities $q = (q_s)_{s \in S}$, student preferences $P_I = (P_i)_{i \in I}$, and strict school priorities $\pi = (\pi_s)_{s \in S}$. The key difference between the model in the last section and this one are the priorities π . For any school s , the function $\pi_s : \{1, \dots, n\} \rightarrow \{i_1, \dots, i_n\}$ is the priority ordering at school s where $\pi_s(1)$ indicates the student with highest priority, $\pi_s(2)$ indicates the student with second highest priority, and so on. Priority rankings are determined by the school district and schools have no control over them. We assume that $n > m$, so that the number of students is greater than the number of schools.

The outcome of a school choice problem, as in college admissions, is a **matching**. In the school choice problem, a matching is defined as $\mu : S \cup I \rightarrow S \cup I$ such that

1. $\mu(i) \notin S \Rightarrow \mu(i) = i$ for any student i , and
2. $|\mu(s)| \leq q_s$ for any school s .

We refer $\mu(i)$ as the assignment of student i under matching μ .

In Fall 2003, the New York City Department of Education changed their assignment mechanism to one based on the student-proposing deferred acceptance algorithm (Abdulkadiroğlu, Pathak, and Roth 2005). One of the features of this system is that it only allows students to submit a rank order list of their top 12 choices. For student who prefer more than 12 schools, these students no longer have a dominant strategy to reveal their true preferences. In practice, between 20 to 30 percent of students rank 12 schools.⁷ Haeringer and Klijn (2006) study this issue and characterize the Nash equilibrium of the preference revelation game induced by various school choice mechanisms when students can only declare a fixed number of schools to be acceptable.

Our next result investigates the impact of a constraint on the number choices a student on the scope for manipulation.

⁷These details together with the entire description of the new allocation procedure is contained in Abdulkadiroğlu, Pathak and Roth (2008).

Proposition 2. *Let $\ell > k > 0$. The student-proposing deferred acceptance mechanism when students may rank only k programs is weakly more manipulable than the student-proposing deferred acceptance mechanism when students can only rank ℓ programs.*

Proof. Let GS denote the student-proposing deferred acceptance mechanism. Let GS^m denote the student-proposing deferred acceptance mechanism when students can only rank m choices. Let k and ℓ be less than the number of students.

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

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

For any student j , let P_j^ℓ be the truncation of P_j after the ℓ th choice. Equation (4) implies that student i does not receive one of her top ℓ choices from the GS mechanism under profile (P_i^ℓ, P_{-i}^ℓ) , otherwise GS would not be strategy-proof. Hence, $GS_i^\ell(P) = i$.

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

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

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

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

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

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

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

Theorem 5.34 of Roth and Sotomayor (1990) implies that

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

Thus,

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

Since s is the only ranked school of student i in preference list \tilde{P}_i , the claim follows. \diamond

Thus, in the first case,

$$s = GS_i^k(\tilde{P}_i, P_{-i})P_iGS_i^k(P) = i.$$

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

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

Proof: Suppose not. Since $GS_i^\ell(P) = i$, there is a college who is assigned strictly more students under $GS^k(P)$ than $GS^\ell(P)$. This is a contradiction to Theorem 5.34, which requires that all colleges are weakly worse off under GS^k (since profile P^k is a truncation of profile P^ℓ).

◇

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

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

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

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

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

Since profile (\tilde{P}_j, P_{-j}^k) is a truncation of $(\tilde{P}_j, P_{-j}^\ell)$,

$$GS_j^k(\tilde{P}_j, P_{-j}) = GS_j(\tilde{P}_j, P_{-j}^k)R_jGS_j(\tilde{P}_j, P_{-j}^\ell) = GS_j(\tilde{P}_j, P_{-j}^\ell) = s,$$

from Theorem 5.34 of Roth and Sotomayor (1990). Since s is the only acceptable school under \tilde{P}_j ,

$$GS_j^k(\tilde{P}_j, P_{-j}) = s,$$

proving the claim. ◇

Thus, for the second case, we have

$$s = GS_j^k(\tilde{P}_j, P_{-j})P_jGS_j^k(P) = j.$$

Finally, the following exhibits a profile where the converse does not hold. Consider a problem

with two students and two schools each with unit capacity. The student preferences are:

$$P_{s_1} : i_1, i_2,$$

$$P_{s_2} : i_1, i_2,$$

while the priority at both schools orders s_1 ahead of s_2 . Under GS^1 , s_2 is unassigned, and can benefit from ranking i_2 as her top choice, while under GS^2 no student can manipulate. This completes the proof. \square

The proof reveals the situations in which we cannot extend the result for strong manipulability. It might be possible that a student cannot manipulate the student-proposing deferred acceptance mechanism when she can only rank k choices, but she can manipulate the student-proposing deferred acceptance mechanism when she can rank $\ell > k$ choices. To see this case, consider the following example:

Example 2. *There are three students i_1, i_2 and i_3 and three schools s_1, s_2 , and s_3 each with unit capacity. Suppose that the preferences of the students are as follows:*

$$P_{i_1} : c_1, c_2, c_3$$

$$P_{i_2} : c_2, c_1, c_3$$

$$P_{i_3} : c_2, c_3, c_1,$$

and the priorities of the schools are:

$$\pi_{s_1} : i_2, i_1, i_3$$

$$\pi_{s_2} : i_3, i_2, i_1$$

$$\pi_{s_3} : i_2, i_1, i_3.$$

When students can only rank one school, the outcome of the student proposing deferred acceptance mechanism is:

$$\mu = \begin{pmatrix} i_1 & i_2 & i_3 \\ s_1 & i_2 & s_s \end{pmatrix},$$

which means that both i_1 and i_3 obtain their top choice, while i_2 is unassigned. Since i_1 obtains her top choice, she cannot manipulate the mechanism.

Consider next the outcome of the student-proposing deferred acceptance mechanism when students can rank two schools:

$$\nu = \begin{pmatrix} i_1 & i_2 & i_3 \\ i_1 & s_1 & s_2 \end{pmatrix},$$

so student i_1 is unassigned, student i_2 receives her second choice, and student i_3 receives her top choice. In this example, student i_1 could manipulate by listing s_3 as her only acceptable school. Under this manipulation, she is be matched to s_3 , which is preferred to what she obtains when she reports the truth. Hence, in this example, the same student does not profitably manipulate both mechanisms.

3.3 The Boston mechanism

Another mechanism that has been widely studied in the school choice problem is the Boston mechanism. This mechanism is by far the most popular mechanism that is used in school districts throughout the U.S. 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 school districts, including: Cambridge MA, Charlotte-Mecklenburg NC, Denver CO, Miami-Dade FL, Minneapolis MN, Providence RI, and Tampa-St. Petersburg FL.

For given set of student preferences and strict school priorities, the outcome of the Boston mechanism is determined in 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.

In the Boston mechanism, students may benefit from manipulating their preferences over schools. 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. If a student is willing to take a risk with her first choice, then she should be careful to rank a second choice that she has a chance of obtaining.

Some families understand these features of the Boston mechanism and have developed rules of thumb for submitting preferences strategically. For instance, the West Zone Parents Group (WZPG), a well-informed group of approximately 180 members who meet regularly prior to admissions time to discuss Boston school choice for elementary school (grade K2), recommends two types of strategies to its members. Their introductory meeting minutes on 10/27/2003 state:

One school choice strategy is to find a school you like that is undersubscribed and put it as a top choice, OR, find a school that you like that is popular and put it as a first choice and find a school that is less popular for a “safe” second choice.

Participants in other school districts employing the Boston mechanism that have developed similar rules of thumb. The following is from the St. Petersburg Times:

Make a realistic, informed selection on the school you list as your first choice. Its the cleanest shot you will get at a school, but if you aim too high you might miss. Heres why: If the random computer selection rejects your first choice, your chances of getting

your second choice school are greatly diminished. That's because you then fall in line behind everyone who wanted your second choice school as their first choice. You can fall even farther back in line as you get bumped down to your third, fourth and fifth choices.

Finally, in experiments, Chen and Sönmez (2006) document that more than 70% of participants in their experiment do not reveal their preferences truthfully under the Boston mechanism.

Using data on stated choices from Boston Public Schools from 2000-2004, Abdulkadiroğlu, Pathak, Roth and Sönmez (2006) describe several empirical patterns which suggest that there are different levels of sophistication among the families who participate in the mechanism. Some fraction of parents behave as the WZPG suggest and avoid ranking two overdemanded schools as their top two choices. The fact that these rules of thumb developed suggest that Boston mechanism is “easy to manipulate.” Of course, the Boston mechanism is strongly more manipulable than the student-proposing deferred acceptance mechanism, which is strategy-proof.

In practice, many school districts employing 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. We can extend the comparison between the Boston mechanism and the student-proposing deferred acceptance mechanism by considering both mechanisms when there is a constraint on the number of schools a student is allowed to rank.

Proposition 3. *Suppose there are more than k schools where $k > 1$. The Boston mechanism when students may rank at most k schools is weakly more manipulable than the student-proposing deferred acceptance mechanism when students may rank at most k schools.*

Proof. Let β be the Boston mechanism, β^k be the Boston mechanism when students may rank at most k schools, GS be the student-proposing deferred acceptance mechanism, and GS^k be the student-proposing deferred acceptance mechanism when students may rank at most k schools. For any student j , let P_j^k be the truncation of P_j after the k th choice.

By definition,

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

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

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

Since $\beta(P^k)$ is stable for P , $\beta(P^k)$ is stable for P^k . Moreover, $GS(P^k)$ is stable for P^k by definition. Since the set of unmatched students across stable matchings is the same,

$$GS_i(P^k) = i \Leftrightarrow \beta_i(P^k) = i. \quad (5)$$

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

Suppose $GS_i^k(P^k) = i$ and i can manipulate. We derive a contradiction. Since i can manipulate, there exists some \hat{P}_i such that

$$GS_i^k(\hat{P}_i, P_{-i}^k) = sP_i i.$$

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

$GS^k(\hat{P}_i, P_{-i}^k)$ remains stable under (\tilde{P}_i, P_{-i}^k) and therefore

$$GS_i^k(\tilde{P}_i, P_{-i}^k) = s. \quad (6)$$

Since $GS(P^k)$ is stable under P^k and $GS_i^k(P^k) = i$ by assumption, relation (7) and stability imply

$$\beta_i(P^k) = i.$$

$\beta(P^k)$ is not stable under (\tilde{P}_i, P_{-i}^k) since student i remains single under $\beta(P^k)$ although not under a stable matching $GS^k(\hat{P}_i, P_{-i}^k)$. Since matching $\beta(P^k)$ is not stable under (\tilde{P}_i, P_{-i}^k) , but it is stable for P^k , the only possible blocking pair of $\beta(P^k)$ in (\tilde{P}_i, P_{-i}^k) is (i, s) . But since $\beta_i(P^k) = i$, this implies that (i, s) also blocks $\beta(P^k)$ under P^k , which is a contradiction.

Thus, under case 1, no student can manipulate GS^k .

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

In this case, some (i, s) blocks $\beta(P^k)$, so that there exists $j \in \beta_s(P^k)$ such that $iP_s j$ and $sP_i \beta_i(P^k)$.

Construct \tilde{P}_i so that s is the first choice. Since $j \in \beta_s(P^k)$ and student i has higher priority than j at school s , $i \in \beta_s(\tilde{P}_i, P_{-i}^k)$. But this means that

$$s = \beta_i(\tilde{P}_i, P_{-i}^k)P_i\beta_i(P^k),$$

a contradiction as we have assumed that no student can manipulate β at P^k .

Finally, the following exhibits a profile where the converse does not hold. Consider a problem with three students and three schools each with unit capacity. The student preferences are:

$$P_{s_1} : i_1, i_2,$$

$$P_{s_2} : i_1, i_2,$$

$$P_{s_1} : i_2, i_3,$$

and the priorities of the schools are:

$$\pi_{s_1} : i_2, i_1, i_3$$

$$\pi_{s_2} : i_1, i_3, i_2$$

$$\pi_{s_3} : i_3, i_2, i_1.$$

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

$$\beta^2(P) = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_1 & i_1 & i_2 \end{pmatrix} \quad \text{and} \quad GS^2(P) = \begin{pmatrix} s_1 & s_2 & s_3 \\ i_2 & i_1 & i_3 \end{pmatrix}.$$

Since each student obtains her second choice or higher, no student can manipulate GS^2 . On the other hand, student i_1 can manipulate β^2 by declaring that s_2 is her only acceptable school. This completes the proof. \square

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

Example 3. There are three students i_1, i_2 and i_3 and three schools s_1, s_2 , and s_3 each with unit capacity. Suppose that the preferences of the students are as follows:

$$P_{i_1} : s_1, s_2, s_3$$

$$P_{i_2} : s_2, s_1, s_3$$

$$P_{i_3} : s_1, s_2, s_3,$$

and the priorities of the schools are:

$$\pi_{s_1} : i_3, i_1, i_2,$$

$$\pi_{s_2} : i_1, i_2, i_3,$$

$$\pi_{s_3} : i_1, i_2, i_3$$

Consider the matching produced by β^2 :

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ i_1 & s_2 & s_1 \end{pmatrix}.$$

In this matching, student i_1 is unassigned, and the other two students receive their top choice. The two students who receive their top choice cannot manipulate.

If student i_1 reported \hat{P}_{i_1} declaring that only s_3 is acceptable, then the matching produced by β^2 is:

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ s_3 & s_2 & s_1 \end{pmatrix},$$

and student i_1 receives her third choice.

Consider, next, the matching produced by GS^2 :

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ s_2 & i_2 & s_1 \end{pmatrix}.$$

In this matching, student i_2 is now unassigned. Student i_1 receives her second choice, and she cannot manipulate because student i_3 , who is assigned to s_1 , has higher priority at s_1 than

i_1 . Thus, we have an example the same individual does not manipulate both mechanisms, illustrating that β^k is not strongly more manipulable than GS^k . Note, of course, that student i_2 can manipulate GS^2 : by ranking s_3 as her top choice, she is be assigned there.

4 Auctions

4.1 Single unit auctions: k-price auction

Consider a seller who wants to sell a single unit. There are N bidders each with value v_i for the unit. In the auctions we consider, the bidder with the highest report wins the object, and the payment rule determines the amount each bidder must pay. Vickrey's (1961) pioneering analysis established that truth-telling is a dominant strategy in the second price auction. Kagel and Levin (1993) were the first to study a 3rd price auction in their study of the winner's curse. In this section, we investigate how the incentives for manipulation vary with the payment rule.

The payment rules we consider are k -price payment rules. When $k = 1$, we have the first price auction. When $k = 2$, we have the 2nd price auction. In general, we consider k -price auctions, for $k \geq 2$. Our next result considers the manipulability of this larger class of mechanisms.⁸

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 valuations of the N bidders and order them from highest to lowest: v_1, v_2, \dots, v_N . Consider the k th price auction. Bidders with values less than or equal to the k th highest valuation can only win the object by reporting a value greater than the highest value. If a bidder does this, she pays an amount greater than their value for the object, and does not obtain a payoff greater than zero. As a result, these bidders cannot manipulate.

When the $(k - 1)$ th highest valuation declares her true value, she does not receive the object. If she manipulates by reporting that her value is higher than the highest valuation bidder, then she does not affect the price she pays because this is equal to v_k , the k th highest value. Her manipulation will be profitable only if $v_{k-1} > v_k$. In this case, in an ℓ th price

⁸Our results do not depend on the particular tie-breaking rule as long as the tie-breaking rule is the same across the mechanisms, so we work with any exogenously specified tie-breaking rule.

auction, if she reports that her value is higher than the highest valuation bidder, she similarly does not affect the price she pays because this is equal to the ℓ th highest bid and $l < k$. This manipulation will be profitable in the ℓ th price auction only when $v_{k-1} > v_\ell$, which is true because $v_{k-1} > v_k \geq v_{k+1} \dots \geq v_\ell$.

When the $(k - 2)^{\text{th}}$ highest valuation bidder declares her true value, she does not receive the object. If she manipulates by reporting that her value is higher than the highest valuation bidder, then she does not affect the k th highest bid. Her deviation is profitable only if $v_{k-2} > v_k$. In this case, in an ℓ th price auction, if she bid more than the highest valuation bidder, she also does not affect the payment, and this manipulation is profitable only when $v_{k-2} > v_\ell$, which is true because $v_{k-2} > v_k \geq v_{k+1} \dots \geq v_\ell$.

Iterating the same argument for the bidders with values in $\{v_2, \dots, v_{k-2}\}$ demonstrates that if a bidder can manipulate the k th price auction, then she can manipulate the ℓ th price auction.

The highest valuation bidder receives the object in the k th price auction. If she reports that her value is strictly greater than v_2 , then she still receives the object and pays v_k , and thus cannot manipulate in this manner. If she submitted a bid strictly less than v_2 , then she does not receive the object, and receives zero payoff. Therefore, the highest valuation bidder cannot manipulate the k th price auction. \square

4.2 Internet Advertisement Auctions

When an Internet user types enters a search term into a 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. Edelman, Ostrovsky and Schwarz (2007) and Varian (2006) pioneered the analysis of the internet advertisement auction problem.⁹ The process by which these advertisement slots are allocated to webpages is 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 follows these earlier contributions. There are N bidders, and $M < N$ ordered slots on a webpage. Each slot has a *click-through* rate of α_i , where $\alpha_i > \alpha_j, \forall i > j$. We assume that the click-through rates are common knowledge among bidders. Bidder k has a value of

⁹See also Athey and Ellison (2007) and Milgrom (2007).

v_k 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 exogenous tie-breaking rule. If bidder k wins slot j , then her utility is given by

$$\pi_k = \alpha_j v_k - p,$$

where p is the price that the bidder pays for the slot.

Edelman, Ostrovsky and Schwarz (2007) present a detailed historical overview of the origins of this market. In 1997, Overture (now part of Yahoo!) introduced an auction for selling Internet advertising. In the original design, each advertiser submitted a bid reporting the advertiser’s willingness to pay on a per-click basis for a particular keyword. Advertising was sold per click, and every time a consumer clicked on a sponsored link, and advertiser’s account was automatically billed the amount of the advertiser’s recent bid. 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, where the winner of slot j pays her bid times the click-through rate of slot j , α_j .

In February 2002, Google introduced its own pay-per-click system, AdWords Select, based on a different payment rule. The payment rule used in this system specifies that an advertiser in position j pays a price-per click equal to the bid of an advertiser in position $i + 1$. This auction format has been dubbed the **Generalized Second Price** (GSP) auction, where the winner of slot j pays the next highest bid times the click-through rate of slot j , α_j . Once Google introduced this new format, 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.

Example. Let us start with an example of 2 slots and four bidders to illustrate the main intuition. Start with the case there are 2 slots, and four bidders. Order the valuations from highest to lowest: (v_1, \dots, v_4) .

We first examine when the GSP can be manipulated. Start with the 3rd highest valuation bidder. This bidder can only manipulate by reporting a value higher than v_2 or by announcing a value higher than v_1 . In the first case, she receives the 2nd slot, which yields utility $\alpha_2 v_3$,

and pays $\alpha_2 v_2$. Since $v_3 \leq v_2$, this does not benefit the bidder (and strictly so unless the two valuations are equal). In the second case, she receives the 1st slot, which yields utility $\alpha_1 v_3$ and pays $\alpha_1 v_1$, also yielding a loss. The same argument shows that any bidder with valuation lower than the 3rd highest bidder also cannot manipulate the auction.

Consider the 2nd highest valuation bidder. When this bidder submits her true value, she pays v_3 per-click for the 2nd slot. Her utility is then

$$\alpha_2 v_2 - \alpha_2 v_3 = \alpha_2 (v_2 - v_3) \geq 0.$$

If this bidder submitted a bid less than the 3rd highest bidder, she does not receive the object and thus would obtain utility 0. If this bidder reported that her valuation is larger than the highest bidder, she would receive the first slot, and pay

$$\alpha_1 v_2 - \alpha_1 v_1 = \alpha_1 (v_2 - v_1) \leq 0,$$

and therefore this is weakly dominated.

As a result, the only bidder who can potentially manipulate in this example is the highest valuation bidder. When this bidder reveals her true value, she pays v_2 per-click for the 1st slot. Her utility is then

$$\alpha_1 v_1 - \alpha_1 v_2 = \alpha_1 (v_1 - v_2) \geq 0.$$

If instead, she shades her bid and reports her valuation to be less than the second highest valuation, but higher than the third highest valuation, she receives the second slot. In this case, her utility is

$$\alpha_2 v_1 - \alpha_2 v_3 = \alpha_2 (v_1 - v_3) \geq 0.$$

Without further assumptions on the click-through rates and values, $\alpha_2 (v_1 - v_3)$ could either be larger or smaller than $\alpha_1 (v_1 - v_2)$

The highest value bidder can only manipulate in the case when

$$\alpha_2 (v_1 - v_3) > \alpha_1 (v_1 - v_2).$$

Under the generalized first price auction, when the highest valuation bidder reports her true valuation, she receives the slot, obtain utility $\alpha_1 v_1$ from the slot, and pays $\alpha_1 v_1$ for the slot.

Thus her total payoff is 0. In contrast, if she reported her valuation to be $v_3 + \frac{\alpha_1}{\alpha_2}(v_1 - v_2)$, then her utility is

$$\alpha_2 v_1 - \alpha_2 \left(v_3 + \frac{\alpha_1}{\alpha_2}(v_1 - v_2) \right) = \alpha_2(v_1 - v_3) - \alpha_1(v_1 - v_2) > 0,$$

since $\alpha_2(v_1 - v_3) > \alpha_1(v_1 - v_2)$ by assumption. This demonstrates that the highest valuation bidder can also manipulate the generalized first price auction.

The proof follows directly from this example.

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

Proof. Fix a vector of valuations ordered from highest to lowest: v_1, v_2, \dots, v_N . Bidders with value v_k for $k = M + 1, \dots, N$ do not have a value high enough to win a slot in the auction. Any bidder with a valuation in this range could potentially manipulate by reporting that their value is greater than v_M . This manipulation allows such a bidder to win a slot. Consider a manipulation such that the bidder obtains the ℓ th slot where $\ell = 1, \dots, M$. Under the GSP, the payoff from this manipulation is

$$\alpha_\ell v_k - \alpha_\ell v_{\ell+1},$$

which is less than or equal to 0 because $v_k \leq v_{\ell+1}$ for $k = M + 1, \dots, N$ and $\ell = 1, \dots, M$. Therefore, bidders value values less than the M th highest value cannot manipulate.

Next, consider a bidder with value v_M . If this bidder reports her true value, then her payoff under the GSP is

$$\alpha_M v_M - \alpha_M v_{M+1} = \alpha_M(v_M - v_{M+1}) \geq 0.$$

If the bidder manipulates by reporting that her value is lower than v_{M+1} , then she does not receive the slot and the manipulation is not unprofitable. Next, consider a manipulation where the bidder reports her value is higher than v_{M-1} . Suppose this manipulation is such that the bidder wins the ℓ th slot for $\ell = 1, \dots, M - 1$. Under the GSP, her payoff is

$$\alpha_\ell v_M - \alpha_\ell v_{\ell+1} = \alpha_\ell(v_M - v_{\ell+1}),$$

which is not greater than zero because $\ell = 1, \dots, M - 1$. Therefore, the lowest value bidder cannot manipulate.

Finally, consider bidders who have a value such that they win a slot under truth-telling. These are bidders with value v_k for $k = 1, \dots, M - 1$. The bidder with value v_k wins the k th slot, and under the GSP, her payoff is

$$\alpha_k v_k - \alpha_k v_{k+1} = \alpha_k (v_k - v_{k+1}) \geq 0.$$

Suppose she shades her bid and reports some value less than v_k . Her manipulation could be low enough such that she does not win a slot and obtains a payoff of zero, in which case her manipulation is not profitable. Suppose instead, that her report wins her spot ℓ for $\ell = k + 1, \dots, M - 1$. Then her payoff under the GSP is

$$\alpha_\ell v_k - \alpha_\ell v_\ell = \alpha_\ell (v_k - v_\ell) \geq 0.$$

Without further assumptions on values, this manipulation might be profitable. In particular, if

$$\alpha_\ell (v_k - v_{\ell+1}) \leq \alpha_k (v_k - v_{k+1}), \quad \forall k = 1 \dots M - 1, \ell = k + 1, \dots, M - 1$$

then it is not profitable to manipulate. Suppose instead that values and click-through rates are such that for some k and ℓ ,

$$\alpha_\ell (v_k - v_{\ell+1}) > \alpha_k (v_k - v_{k+1}).$$

Under the GFP, the bidder with value v_k wins the k th slot, and obtains payoff of $\alpha_k v_k - \alpha_k v_k = 0$.

Consider what happens if the bidder deviates by reporting her value to be less than v_k such that she wins the ℓ th slot, for $\ell = k + 1, \dots, M - 1$. Let her manipulated report declare her value to be $v_\ell + \frac{\alpha_k}{\alpha_\ell} (v_k - v_{k+1})$. Under the GFP,

Under the GFP, her payoff is

$$\alpha_\ell v_k - \alpha_\ell \left(v_\ell + \frac{\alpha_k}{\alpha_\ell} (v_k - v_{k+1}) \right) = \alpha_\ell (v_k - v_\ell) - \alpha_k (v_k - v_{k+1}) > 0,$$

since $\alpha_\ell(v_k - v_\ell) > \alpha_k(v_k - v_{k+1})$ by assumption. Thus the same bidder can manipulate the GFP.

Finally, we consider the potential for bidders with values v_k where $k = 2, \dots, M - 1$ to manipulate so that they receive a slot with a higher click-through rate. Under the GSP, such a bidder receives payoff

$$\alpha_k v_k - \alpha_k v_{k+1}.$$

If these bidders report a value such that they obtain the ℓ th spot, where $\ell = 1, \dots, k - 1$, then such a bidder receives payoff

$$\alpha_\ell v_k - \alpha_\ell v_\ell = \alpha_\ell(v_k - v_\ell) \leq 0.$$

Therefore, any manipulation which yields a slot with a higher click-through rate is unprofitable.

It is simple to construct an example where the GSP is not manipulable, while the GSP is manipulable, so the details are omitted. \square

4.3 Multi-unit auctions

Unlike the two previous examples, the next environment 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. See Krishna (2002) for more background.

Three sealed-bid auction formats are typically discussed when multiple identical units are to be sold. In each of these formats, a bidder is asked to submit bids for each of the K items indicating how much she is willing to pay for each additional unit. We assume, like most other analysis, that each bidder's value for the first unit is no less than her value for the second unit, and so on.

The Vickrey auction is the multi-unit generalization of the Vickrey-Clarke-Groves mechanism and, although not widely used, is strategy-proof for bidders.¹⁰ In the discriminatory

¹⁰Milgrom (2005) describes some of the reasons why Vickrey auctions are not employed in multi-unit contexts including various monotonicity problems. See also Rothkopf (2007).

format, also known as the pay-your-bid auction, each bidder pays an amount equal to the sum of his bids that are winning. The discriminatory auction is the natural multi-unit extension of the first-price sealed bid auction (it is identical when $k = 1$). In the uniform-price auction, all k units are sold at a “market-clearing” price such that the total amount demanded is equal to the total amount supplied.

Starting in the 1970s, the US Treasury employed a discriminatory format to auction Treasury bonds. In 1992, the US Treasury switched to a uniform-price auction for 2 and 5 year notes. During the time, there was a lot of policy discussion on the switch. One of the earliest proponents of the virtues of a uniform-price auction over a discriminatory price auction was Milton Friedman (1960). He argued that a uniform-price auction was strategically simpler, and that a uniform-price format is perceived as leveling the playing field by reducing the importance of specialized knowledge among dealers. Moreover, according to Friedman, more bidders would be induced to bid directly in auctions because the fear of being awarded securities at too high a price is eliminated. In the policy change of the US Treasury in the 1990s, Merton Miller wrote in favor of the change writing, “All of that [gaming] is eliminated if you use the [uniform-price] auction.”

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 price auction, bidders have an incentive to report that their bids are just above the lowest bid to receive a unit. In a uniform price auction, it is a weakly dominant strategy for a bidder to bid truthfully for the first unit. However, for additional units, a bidder has an incentive to shade their bid because that bid has the potential to influence the price she pays. This “demand-reduction” feature of the uniform price auction prevents it from being strategy-proof.

We now formally investigate how our notion of manipulability allows us to compare uniform and discriminatory formats. A single seller wishes to sell a fixed number k of identical items to N bidders. The bidders are asked to report their valuations for the k objects, where v_i^j is bidder i 's valuation for the j th unit of the item for sale. In the auction, a total of N reports of valuations for the k items are collected, and the k units are awarded to the k highest reported valuations.

The utility of agent i who receives ℓ objects at a total price of p is:

$$u_i = v_i^1 + \dots + v_i^\ell - p,$$

where $v_i^1 \geq v_i^2 \geq \dots \geq v_i^\ell$.

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 each of the units awarded, bidders pay the $(k + 1)^{\text{th}}$ highest value.¹¹

The next proposition supports Milton Friedman’s intuition that the uniform price auction is “strategically simpler” than the discriminatory price auction.

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

Proof. Fix the reports of every bidder except for bidder i . Denote these bids as b_1, b_2, \dots , as without loss of generality, let b_1 be the highest, b_2 be second highest, and so on. We consider the following cases:

- 1) 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 valuation, then in a uniform price auction, the market-clearing price is $p = \max\{v_i^1, b_{k+1}\}$ and bidder i does not obtain any units. Bidder i could manipulate by reporting that $j = 1, \dots, k$ of her valuations are greater than b_{k-j+1} to win j units, and when $j < k$, her valuation for the $(j + 1)^{\text{th}}$ unit is \hat{v}_i .

Such a manipulation would yield a market clearing price of $p = \max\{\hat{v}_i, b_{k-j+1}\}$. The payoff from this manipulation is:

¹¹We can consider other “market clearing” rules such as paying the k th value or paying a value between the k th and $k+1$ th value. This does not affect the analysis.

$$(v_i^1 + \dots + v_i^j) - j \cdot \max\{\hat{v}_i, b_{k-j+1}\}.$$

Since $v_i^1 \leq b_k$ and $b_k \leq \max\{\hat{v}_i, b_{k-j+1}\}$, this is not profitable.

- 2) Bidder i 's j highest values where $j = 1, \dots, k$ are among the top k highest bids (e.g., $v_i^j \geq b_{k-j+1}$ and if $j \neq k$, $v_i^{j+1} < b_{k-j+1}$)

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

$$(v_i^1 + \dots + v_i^j) - j \cdot \max\{v_i^{j+1}, b_{k-j+1}\} \geq 0. \quad (7)$$

Bidder i could manipulate by reporting bids such that she wins $\ell < j$ units and that her valuation for the $(\ell + 1)^{\text{th}}$ unit is \hat{v}_i . In this case, in a uniform-price auction, the market-clearing price is $p = \max\{\hat{v}_i, b_{k-\ell+1}\}$ and bidder i obtains ℓ units. Her profit is

$$(v_i^1 + \dots + v_i^\ell) - \ell \cdot \max\{\hat{v}_i, b_{k-\ell+1}\}.$$

Suppose the values are such that

$$(v_i^1 + \dots + v_i^\ell) - \ell \cdot \max\{\hat{v}_i, b_{k-\ell+1}\} > (v_i^1 + \dots + v_i^j) - j \cdot \max\{v_i^{j+1}, b_{k-j+1}\}, \quad (8)$$

so that the manipulation is profitable. In a discriminatory auction, if bidder i reported her true valuations, then she receives j 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

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

Since $b_{k-\ell+1} \leq \max\{\hat{v}_i, b_{k-\ell+1}\}$, equations (7) and (8) imply that

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

Therefore, bidder i can also manipulate the discriminatory auction.

When $j \neq k$, bidder i could also manipulate by reporting bids such that she wins $m > j$ units where $m \leq k$ and if $m < k$, declaring that her $(m + 1)^{th}$ highest value is \hat{v}_i . In this case, in a uniform-price auction, the market-clearing price is $p = \max\{\hat{v}_i, b_{k-m+1}\}$ and bidder i obtains m units. Her profit is

$$(v_i^1 + \dots + v_i^m) - m \cdot \max\{\hat{v}_i, b_{k-m+1}\}.$$

However, since $m > j$, $b_{k-m+1} > b_{k-j+1}$ and $b_{k-m+1} > v_i^{j+1}$. This implies that

$$\begin{aligned} m \cdot \max\{\hat{v}_i, b_{k-m+1}\} - j \cdot \max\{v_i^{j+1}, b_{k-j+1}\} &\geq m \cdot \max\{\hat{v}_i, b_{k-m+1}\} - j \cdot b_{k-m+1} \\ &\geq (m - j) \cdot b_{k-m+1} \geq (m - j) \cdot b_{k-j+1}. \end{aligned}$$

Since $v_i^{j+1} < b_{k-j+1}$,

$$(m - j) \cdot b_{k-j+1} > (m - j) \cdot v_i^{j+1} \geq (v_i^{j+1} + \dots + v_i^m).$$

Thus,

$$m \cdot \max\{\hat{v}_i, b_{k-m+1}\} - j \cdot \max\{v_i^{j+1}, b_{k-j+1}\} > (v_i^{j+1} + \dots + v_i^m),$$

which demonstrates that a manipulation by bidder i to win more than j units is not profitable in the uniform price auction.

It is simple to construct an example where the uniform price auction is not manipulable and the discriminatory price auction is, so the details are omitted. \square

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 was 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 two different literatures: matching and auction problems. We are hopeful that subsequent work investigates other settings where manipulation has been studied such as in 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, 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 auction and the multi-unit auction 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.

Lastly, there are many potential avenues to build on the ideas in this paper. (1) If manipulation by coalitions is a consideration, one could define a version of our notion for manipulating coalitions. (2) Each of the mechanisms we analyzed in the paper are deterministic; one could also consider mechanisms which are stochastic. (3) One motivation for our weak notion is to make comparisons across situations where truth-telling is a Nash equilibrium. One could also consider different environments where other equilibrium concepts could form a foundation for measuring manipulability. (4) Another avenue for further work is to strengthen the strong version of manipulability by requiring that the deviating strategy is the same across the two mechanisms. (5) Finally, both notions of manipulability apply to situations where there exists a problem where one mechanism is manipulable and the other is not. In situations where one mechanism is manipulable only if another mechanism is manipulable, an alternative may be to consider the maximal gain from deviation from truth-telling to compare mechanisms.¹² These, and other, extensions are left for future work.

¹²Day and Milgrom (2007) consider “core-selecting” auctions and show that the bidder-optimal core-selecting auction minimizes the maximal gain from reporting a deviation from truth-telling.

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, 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, 364-367.
- [4] Ausubel, Larry and Peter Cramton. 2002. "Demand Reduction and Inefficiency in Multi-Unit Auctions." *Unpublished mimeo*, University of Maryland.
- [5] Athey, Susan and Glenn Ellison. 2007. "Position Auctions with Consumer Search." Working paper, MIT.
- [6] Bartholdi, John J. and James B. Orlin. 1991. "Single Transferable Vote Resists Strategic Voting." *Social Choice and Welfare*, 8: 341-354.
- [7] Bergemann, Dirk and Stephen Morris. 2005. "Robust Mechanism Design." *Econometrica*, 73: 1771-1813.
- [8] Chen, Yan and Tayfun Sönmez. 2006. "School Choice: An Experimental Study." *Journal of Economic Theory*, 127: 202-231.
- [9] Chung, Kim-Sau and Jeff Ely. 2007. "Foundations of Dominant Strategy Mechanisms." *Review of Economic Studies*, 74: 447-476.
- [10] Dasgupta, Partha and Eric Maskin. 2007. "On the Robustness of Majority Rule." *Journal of the European Economic Association*, forthcoming.
- [11] Day, Robert and Paul Milgrom. 2007. "Core-Selecting Auctions." *International Journal of Game Theory*, July.
- [12] Dubins, Lester E. and David A. Freedman. 1991. "Machiavelli and the Gale-Shapley Algorithm." *American Mathematical Monthly*, 88: 485-494.
- [13] Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review*, 97-1: 242-259.

- [14] Ehlers, Lars. 2004. "In Search of Advice for Participants in Matching Markets which use the Deferred Acceptance Algorithm." *Games and Economic Behavior*, 48: 249-270.
- [15] Ehlers, Lars. 2006. "Truncation Strategies in Matching Markets." *Mathematics of Operations Research*, forthcoming.
- [16] Ergin, Haluk. 2002. "Efficient Resource Allocation on the Basis of Priorities." *Econometrica*, 70: 2489-2497.
- [17] Friedman, Milton. 1960. A Program for Monetary Stability, Fordham University Press.
- [18] Gale, David and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69: 9-15.
- [19] Gale, David and Marilda Sotomayor. 1985. "Some Remarks on the Stable Matching Problem." *Discrete Applied Mathematics*, 11: 223-232.
- [20] Gootman, Elissa. 2006. "Lafayette Among 5 High Schools to Close." *New York Times*, Education section, 12/14/06.
- [21] Haeringer, Guillaume and Flip Klijn. 2006. "Constrained School Choice." UFAE and IAE Working papers, Universitat Autònoma de Barcelona, report 671.06.
- [22] 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.
- [23] Kesten, Onur. 2006. "On Two Competing Mechanisms for Priority-Based Allocation Problems." *Journal of Economic Theory*, 127: 155-171.
- [24] Kojima, Fuhito and Parag A. Pathak. 2007. "Incentives and Stability in Large Two-Sided Matching Markets." *Unpublished mimeo*, Harvard University.
- [25] Krishna, Vijay. 2002. "Auction Theory." Academic Press.
- [26] Milgrom, Paul. 2004. "Putting Auction Theory to Work (Churchill Lectures in Economics)." Cambridge University Press.
- [27] Milgrom, Paul. 2005. "Package Auctions and Package Exchanges." *Econometrica*, 75; 935-966.
- [28] Milgrom, Paul. 2007. "Simplified Mechanisms with Applications to Sponsored Search and Package Auctions." *Unpublished mimeo*, Stanford University.

- [29] Moulin, Hervé. 1980. "On Strategy-Proofness and Single Peakedness." *Public Choice*, 35: 437-455.
- [30] Neeman, Zvika. 2004. "The Relevance of Private Information in Mechanism Design." *Journal of Economic Theory*, 117: 55-77.
- [31] Roth, Alvin E. 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operations Research*, 7: 617-628.
- [32] 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.
- [33] Roth, Alvin E. 1985. "The College Admissions Problem is Not Equivalent to the Marriage Problem." *Journal of Economic Theory*, 36: 277-288.
- [34] Roth Alvin E. 1991. "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K." *American Economic Review*, 81: 415-440.
- [35] Roth, Alvin E. 2002. "Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica*, 70: 1341-1378.
- [36] Roth, Alvin E. 2003. "The origins, history, and design of the resident match." *Journal of the American Medical Association*, 289 (7): 909-212.
- [37] Roth, Alvin E. 2008. "Deferred Acceptance 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.
- [38] 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.
- [39] Roth, Alvin E. and Uriel G. Rothblum. 1999. "Truncation Strategies in Matching Markets—In Search of Advice for Participants." *Econometrica*, 67-1: 21-44.
- [40] Roth, Alvin E. and Marilda Sotomayor. 1990. Two-Sided Matching: A Study in Game Theoretic Modeling and Analysis. *Econometric Society Monographs*, vo. 18, Cambridge University Press.
- [41] Rothkopf, Michael. 2007. "Thirteen Reasons Why the Vickrey-Clarke-Groves Mechanism is Not Practical." *Operations Research*, 55:2, 191-197.

- [42] Sönmez, Tayfun. 1997. "Manipulation via Capacities in Two-Sided Matching Markets." *Journal of Economic Theory*, 77: 197-204.
- [43] Varian, Hal R. 2006. "Position Auctions." *International Journal of Industrial Organization*.
- [44] Vickrey, W. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, 16, 8-37.
- [45] Wilson, R. 1987. "Game-Theoretic Analysis of Trading Processes," in *Advances in Economic Theory: Fifth World Congress*, ed. Truman Bewley, Cambridge, UK: Cambridge University Press.