Non-Obvious Manipulability of the Rank-Minimizing Mechanism

Peter Troyan
Department of Economics
University of Virginia*

June 22, 2022

Abstract

In assignment problems, the average rank of the objects received is often used as key measure of the quality of the resulting match. The rank-minimizing (RM) mechanism optimizes directly for this objective, by returning an assignment that minimizes the average rank. While a naturally appealing mechanism for practical market design, one shortcoming is that this mechanism fails to be strategyproof. In this paper, I show that, while the RM mechanism is manipulable, it is not *obviously* manipulable.

Keywords: matching, rank distribution, obviousness, manipulability

JEL Codes: C78, D47

*Address: P.O. Box 400182, Charlottesville, VA, 22904. Email: troyan@virginia.edu.

1 Introduction

Many institutions make assignments by collecting agents' ordinal preferences over the possible alternatives and using them as an input to a rule that outputs an assignment. Common examples include public school choice (Abdulkadiroğlu and Sönmez, 2003), medical residency matching (Roth and Peranson, 1999), teacher assignment (Combe et al., 2022), course allocation (Budish and Cantillon, 2012; Budish and Kessler, 2017), and refugee resettlement (Delacrétaz et al., 2016). A natural metric to measure the success of the outcome is the rank distribution: how many students get their first choice, how many get their second choice, and so on. Indeed, many school districts such as those in New York City and San Francisco publicly release statistics on the ranks as a measure of the goodness of the match.

In the context of school choice specifically, most cities rely on either some version of Gale and Shapley's celebrated deferred acceptance mechanism (Gale and Shapley, 1962) or on the so-called Boston mechanism (also sometimes called the immediate acceptance mechanism) to determine the assignment, and then evaluate the rank distribution produced by these mechanisms.[1] However, given the importance placed on the rank distribution, it is also natural to consider mechanisms that optimize directly for this objective. Indeed, Featherstone (2020) notes that Teach for America does exactly this, and uses the rank distribution when *selecting* its assignment. Featherstone (2020) is also one of the few papers that has undertaken a serious analysis of mechanisms based explicitly on the rank distribution (a few other papers in this relatively small but growing literature are discussed in Section 4).

While using mechanisms that select assignments based explicitly on the rank distribution is naturally appealing, an important consideration in any mechanism design problem is the incentives of the agents. Indeed, one of the most appealing properties arguing for the use of DA-based mechanisms is that they generally give agents strong incentives to report their true preferences. On the other hand, Proposition 10 of Featherstone (2020) shows that no

[1] There is another class of mechanisms based on the top trading cycles (TTC) algorithm of Shapley and Scarf (1974) that has been studied extensively in the theoretical literature, but has found few adherents in practice. The only use of it in a real-world school choice setting that I am aware of is New Orleans, where it was used for one year before being abandoned (Abdulkadiroğlu et al., 2020).

ordinal assignment mechanism is both rank-efficient (a refinement of Pareto efficiency that he defines that takes into account the rank distribution) and strategyproof.

However, strategyproofness is a very demanding property, and just because a mechanism *can* be manipulated does not mean that it is *will* be manipulated. As part of a recent strand of literature on "obviousness" in mechanism design, Troyan and Morrill (2020) introduce the concept of a *(non-)obvious manipulation* as a way to relax strategyproofness. They use their definition to taxonomize non-strategyproof mechanisms into two classes: those that are obviously manipulable (such as the Boston mechanism) and thus are likely to be easily manipulated in practice,[2] and those that are non-obviously manipulable (such as school-proposing DA, or, as they show, any mechanism that is stable-dominating in the sense of Alva and Manjunath, 2019) which, while formally manipulable, have manipulations that are more difficult for cognitively-limited agents to recognize and enact successfully.

In this paper, I consider a canonical assignment model in which there is a set of agents (such as students) to be assigned to a set of objects (such as schools), each of which has some fixed capacity. Agents participate in a mechanism in which they report their preferences over the objects. I consider the *rank-minimizing (RM) mechanism,* which takes the reported agent preferences, and implements an assignment that minimizes the average rank of the objects received. My main result is to show that the RM mechanism, while manipulable, is not *obviously* manipulable. Thus, to the extent that minimizing the average rank is a desirable policy objective, and that the main shortcoming of the RM mechanism is its lack of strategyproofness, my results suggest that this shortcoming may not be so severe, and the benefits of such a mechanism in terms of potential efficiency gains may outweigh the costs in terms of incentives. At the very least, given their other appealing properties, RM mechanisms are worthy of further consideration for practical market design.[3]

[2]There is a plethora of evidence that the Boston mechanism is commonly manipulated; for a detailed analysis, see, for instance, Dur et al. (2018).

[3]Featherstone (2020) shows that rank efficiency is incompatible with strategyproofness, but also shows that in a low-information environment where beliefs over other agents' preferences satisfy a symmetry condition as in Roth and Rothblum (1999), agents are weakly better off by telling the truth. A main advantage of non-obvious manipulability is that it does not rely on any assumption about agents' beliefs. This is discussed further in

2 Model

There is a set of N agents $I = \{i_1, \dots, i_N\}$ and a set of M objects $O = \{o_1, \dots, o_M\}$. Each object o_m has a **capacity** q_m which denotes the maximum number of agents who can be assigned to it. I assume that $\sum_{m=1}^{M} q_m \geq N$, which is common in school choice settings where all students must be offered a seat at a school, and is without loss of generality if one of the objects is an "outside option" that has enough capacity for all agents. Each agent has a **strict preference relation** $\succ_i$ defined on the set of objects O, where $o \succ_i o'$ denotes that agent i strictly prefers object o to object o'. I use $o \succsim_i o'$ when either $o \succ_i o'$ or $o = o'$. I will sometimes refer to $\succ_i$ as an agent's **type**. I will also write $\succ_i$: $o, o', \dots$, to denote an agent who has preferences such that her favorite object is o, second favorite object is o', and the rest of her preferences can be arbitrary.

For an agent i with type $\succ_i$, I define $r_i(o) = |o' \in O : o' \succsim_i o|$ to be the **rank** of object o according to i's preferences; in other words, i's favorite object has rank 1, i's second-favorite object has rank 2, etc.[4] Notice that I use the convention that lower numbers correspond to more preferred objects.

An **allocation** $\alpha : I \to O$ is a function that assigns each agent to an object. I use α_i to denote the object assigned to agent i in allocation α. Any allocation must satisfy $|i \in I : \alpha_i = o| \leq q_m$ for all $o_m \in O$, i.e., each object cannot be assigned to more agents than its capacity. Let $\mathcal{A}$ be the set of all possible allocations. Let $\mathcal{P}_i$ be agent i's preference domain, which consists of all strict rankings over O, and let $\mathcal{P}^I = \mathcal{P}_1 \times \cdots \times \mathcal{P}_N$. I write $\succ_I = (\succ_1, \dots, \succ_N) \in \mathcal{P}^I$ to denote a profile of preferences, one for each agent $i_1, \dots, i_N$, and sometimes write $\succ_I = (\succ_i, \succ_{-i})$ to separate i's preferences $\succ_i$ from those of the remaining agents, $\succ_{-i}$. The **average rank** of any allocation α is

$$\bar{r}(\alpha) = \frac{1}{N} \sum_{i \in I} r_i(\alpha_i).$$

As above, I suppress the dependence of $\bar{r}(\alpha)$ on the preference profile $\succ_I$ for readability. A **mechanism** is a function $\psi : \mathcal{P}^I \to 2^{\mathcal{A}}$. Notice that I define the range of a mechanism ψ to be the power set of $\mathcal{A}$. The reason for this

Section 3.

[4] The rank function depends implicitly on agent i's preferences $\succ_i$, but for readability, I write this as just $r_i(\cdot)$.

choice is discussed below, after introducing the main mechanism of interest.

The focus of this paper is on one mechanism in particular, the rank-minimizing mechanism. The **rank-minimizing (RM) mechanism**, ψ^{RM}, is given by:

$$\psi^{RM}(\succ_I) = \{\alpha \in \mathcal{A} : \bar{r}(\alpha) \leq \bar{r}(\alpha') \text{ for all } \alpha' \in \mathcal{A}\}.$$

In words, ψ^{RM} takes any preference profile as an input and returns *all* of the allocations that minimize the average rank. The reason that the range of the function is set-valued is that there can in general be many allocations that achieve the minimum average rank, and because the incentive notion considered in the next section uses a worst-case analysis, it becomes more natural to formally define a mechanism in this way.

3 Main Result

Before giving the main result, I first introduce the notion of non-obvious manipulability from Troyan and Morrill (2020). They define the concept for any direct revelation mechanism in which agents report their type. Here, I adapt their definition to my setting and write it using the rank notation introduced in the previous section.

Given any set of allocations $A \in \mathcal{A}$, let

$$\bar{\rho}_i(A) = \max_{\alpha \in A} r_i(\alpha_i)$$

$$\underline{\rho}_i(A) = \min_{\alpha \in A} r_i(\alpha_i).$$

In words, $\bar{\rho}_i(A)$ is the rank of agent i's worst assignment over all of the allocations in the set A, where once again the dependence on $\succ_i$ is implicit. That is, $\bar{\rho}_i(A)$ is the rank of i's least-preferred school among all allocations in A, where recall that a higher rank corresponds to a worse school, and thus the worst-case is given by taking the maximum. Similarly, $\underline{\rho}_i(A)$ is the rank of agent i's best assignment over all of the allocations in the set A.

Definition 1. A mechanism ψ is **not obviously manipulable** if, for any agent i of type $\succ_i$ and any $\succ_i' \neq \succ_i$, the following are true:

(i) $\max_{\succ_{-i}} \bar{\rho}_i(\psi(\succ_i, \succ_{-i})) \leq \max_{\succ_{-i}} \bar{\rho}_i(\psi(\succ_i', \succ_{-i}))$

(ii) $\min_{\succ_{-i}} \underline{\rho}_i(\psi(\succ_i, \succ_{-i})) \leq \min_{\succ_{-i}} \underline{\rho}_i(\psi(\succ_i', \succ_{-i}))$

If either of (i) or (ii) does not hold for some agent and type, then $\succ_i'$ is an **obvious manipulation** for agent i of type $\succ_i$, and the mechanism ψ is said to be **obviously manipulable**.

To understand Definition 1, first consider part (i). On the left-hand side of the inequality, $\psi(\succ_i, \succ_{-i})$ is the set of all possible allocations when i reports her true preferences, $\succ_i$, and the other agents report $\succ_{-i}$. Next, $\bar{\rho}_i(\psi(\succ_i, \succ_{-i}))$ is the rank of i's worst-case outcome in this set. Finally, we take the maximum over all of the other agent reports, $\succ_{-i}$, to give the rank of the worst-case outcome. The RHS is the same, just replacing $\succ_i$ with a misreport $\succ_i'$. The inequality in (i) says that the rank of the worst-case outcome under the misreport should be worse (i.e., higher) than that under the truth. Part (ii) of Definition 1 is analogous, except it compares the best-case outcomes instead, and, since lower numbers correspond to more preferred outcomes, we replace max with min.

There are several justifications for why a designer might be concerned with obvious manipulations. As discussed in the introduction, just because a mechanism *can* be manipulated does not mean that it *will* be manipulated, and Definition 1 is a way to separate those manipulations that are "obvious", and are thus likely to be identified by participants, from those that are not. Formally, Theorem 1 of Troyan and Morrill (2020) shows that obvious manipulations are precisely those manipulations that can be recognized by an agent who is cognitively-limited in the sense defined by Li (2017), and is unable to contingently reason about outcomes state-by-state. Mathematically, such agents know the range of the function ψ conditional on their own reports, but not the full function itself, state-by-state.[5] Allowing some manipulations so long as they are non-obvious widens the space of mechanisms available to the designer, which may allow for improvements on other dimensions, such as the average rank. Further, unlike other relaxations of

[5]This is particularly relevant in the context of school choice. For instance, Troyan and Morrill (2020) write: "...this could be a neighborhood parent group that does not fully understand (or has not been told) the assignment algorithm but has kept track of what preferences have been submitted and what the resulting assignments were." Such parent groups are indeed quite common; see Pathak and Sönmez (2008).

strategyproofness such as Bayesian incentive compatibility or approximate notions such as strategyproofness-in-the-large (SPL, Azevedo and Budish, 2019), NOM requires no assumptions on how preferences are drawn or agent beliefs. Rather, NOM is defined using best and worst case scenarios, which are likely to be particularly salient.[6]

Given Definition 1, I can now state the main result of the paper, Theorem 1 below.

Theorem 1. *The rank-minimizing mechanism is not obviously manipulable.*

The proof of this theorem can be found at the end of the paper, following the Discussion and Conclusion section.

4 Discussion and Conclusion

The rank-minimizing mechanism is an appealing mechanism for assignment problems, because it directly optimizes a natural objective that is desirable to policy-makers. Thus, it is somewhat striking that there has been thus far relatively little written about this mechanism in the literature, while there have been hundreds of papers written about mechanisms such as DA and TTC. One possible explanation for this gap is that the literature is overly-focused on strategyproofness as an incentive property, which is satisfied by both DA and TTC. While strategyproofness is a very appealing desideratum, it limits the flexibility for a designer to optimize on other important dimensions. This paper shows that, while not strategyproof, the rank-minimizing mechanism is at least not obviously manipulable.

Besides Featherstone (2020), who provides a detailed analysis of rank efficiency criteria and related mechanisms and was discussed in the Introduction, there are only two other papers I am aware of in the economics literature that analyze the RM mechanism. Nikzad (2022) studies the RM mechanism in large markets, and provides an upper bound on the expected average rank

[6]Following Troyan and Morrill (2020), several other papers have applied non-obvious manipulability to various settings, including Aziz and Lam (2021), Ortega and Segal-Halevi (2019), Archbold et al. (2022), Troyan et al. (2020), and Cerrone et al. (2022). Also similar in spirit, though technically different, is Li and Dworczak (2021) who show that a designer can sometimes be better off using a non-SP mechanism even when agents are unsophisticated, no matter how they resolve their "strategic confusion".

(see also Sethuraman (2022) for another proof of the same result). He also shows that the RM mechanism is Bayesian incentive compatible when agent preference rankings are drawn uniformly at random. Ortega and Klein (2022) study the average rank of RM, DA, and TTC in large markets, both theoretically and using simulations, and find that RM outperforms DA and TTC on important dimensions such as efficiency and fairness. They also use data from secondary school admissions in Hungary to analyze the three mechanisms in an empirical setting, and find similar support for RM. The data is from DA, a strategyproof mechanism, and they write:

> One potential concern is that students would submit different preferences when allocations are determined by RM, which is not strategy-proof. This concern is mitigated because the potential gains from misrepresentation are small...Moreover, there is evidence than [*sic*] truthful behavior can depend more on the complexity of the mechanism rather than actual strategy-proofness (Cerrone et al., 2022).

My result showing the non-obvious manipulability of RM can be seen as a formalization of this final point of Ortega and Klein (2022). Whether the rank-minimizing mechanism will actually be manipulated in practice is of course ultimately an empirical question, but given my result and the other desirable properties of the mechanism, RM is a mechanism that seems worthy of further investigation.

Proof of Theorem 1

I start with the following lemma.

Lemma 1. *Consider a preference profile $\succ'_I$ such that for all $j \in I$, agent j's preferences are $\succ'_j$: $o_{m_1}, o_{m_2}, \ldots, o_{m_M}$, and let $k^* = \min\{k : \sum_{k'=1}^{k'} q_{m_{k'}} \geq N\}$. Further, let $O' = \{o_{m_1}, \ldots, o_{m_{k^*}}\}$ and $A^* \subseteq \mathcal{A}$ be the subset of allocations that satisfy:*

(I) All $o_{m_k} \in O' \setminus \{o_{m_{k^*}}\}$ are assigned to exactly q_{m_k} agents.

(II) Object $o_{m_{k^*}}$ is assigned to exactly $N - \sum_{k'=1}^{k^*-1} q_{m_{k'}}$ agents.

Then, $\psi^{RM}(\succ'_I) = A^*$. Further, for any $\tilde{o} \in O' = \{o_{m_1}, \ldots, o_{m_{k^*}}\}$, there is at least one allocation $\alpha \in \psi^{RM}(\succ'_I)$ such that $\alpha_i = \tilde{o}$.

In words, object $o_{m_{k^*}}$ is the critical object, in the sense that the set $O' = \{o_{m_1}, \ldots, o_{m_{k^*}}\}$ has enough total capacity to accommodate all agents, while the set $O' \setminus \{o_{m_{k^*}}\}$ does not. The set A^* is then the set of all allocations that consist of all possible ways of assigning the N agents to the objects in O' such that $o_{m_1}, \ldots, o_{m_{k^*-1}}$ are filled to capacity, and all remaining agents are assigned to $o_{m_{k^*}}$. The lemma says that when all agents have the same preferences, any rank-minimizing allocation must satisfy (I) and (II).

Proof of Lemma 1. For any allocation α, let $\alpha_{m_k} = \{i \in I : \alpha_i = o_{m_k}\}$ be the set of agents assigned to o_{m_k}. Under the preference profile given, the average rank of any allocation is:

$$\bar{r}(\alpha) = \frac{1}{N}\left(\sum_{k=1}^{M} k|\alpha_{m_k}|\right).$$

This is clearly minimized by an allocation that assigns q_{m_1} students to o_{m_1}, q_{m_2} students to o_{m_2}, etc., until all students are assigned. This is precisely the set of allocations A^*, and so by construction, any allocation $\alpha \in A^*$ achieves the minimum average rank, i.e., $\alpha \in \psi^{RM}(\succ_I)$. For any allocation $\alpha' \notin A^*$, either (I) or (II) must fail. If (I) fails, let $o_{\hat{m}} \in \{o_1, \ldots, o_{m_{k^*-1}}\}$ be an object that is not filled to capacity. Then, there must be some agent j assigned to some object $o_{m_{k^*}}, \ldots, o_{m_M}$. Reassigning agent j to object $o_{\hat{m}}$ and leaving all other assignments the same lowers the average rank, and so $\alpha' \notin \psi^{RM}(\succ_I)$. Therefore, we have $\psi^{RM}(\succ'_I) = A^*$. Lastly, because all agents have the same preferences, it does not matter precisely which agents are assigned to which schools, and so by symmetry, there exists at least one $\alpha \in \psi^{RM}(\succ'_I)$ such that $\alpha_i = \tilde{o}$ for all $\tilde{o} \in O'$.

■

Now, consider an agent i with type $\succ_i$. I show that for any $\succ'_i \neq \succ_i$, each part of Definition 1 holds for the rank-minimizing mechanism.

Part (i): $\max_{\succ_{-i}} \bar{\rho}_i(\psi(\succ_i, \succ_{-i})) \leq \max_{\succ_{-i}} \bar{\rho}_i(\psi(\succ'_i, \succ_{-i}))$.

Index agent i's true type as $\succ_i$: $o_{m_1}, o_{m_2}, \ldots, o_{m_M}$. I first show that when i reports her true preferences, $\max_{\succ'_{-i}} \bar{\rho}_i(\psi^{RM}(\succ_i, \succ'_{-i})) = k^*$, where k^* is as defined in Lemma 1. When $\succ_j = \succ_i$ for all j, there is at least one allocation $\alpha \in \psi^{RM}(\succ_I)$ such that $\alpha_i = o_{m_{k^*}}$, by Lemma 1. This implies $\max_{\succ'_{-i}} \bar{\rho}_i(\psi^{RM}(\succ_i, \succ'_{-i})) \geq k^*$. To show equality, assume that $\max_{\succ_{-i}} \bar{\rho}_i(\psi^{RM}(\succ_i, \succ_{-i})) > k^*$. Then, there must be some $\succ'_{-i}$ and some $\alpha' \in \psi^{RM}(\succ_i, \succ'_{-i})$ such that $\alpha'_i = o_{m_{k'}}$ for some $k' > k^*$. But, this implies that there is some $k'' \leq k^*$ such that object $o_{m_{k''}}$ is not filled to capacity. Thus, consider an alternative allocation α'' where agent i is reassigned to $o_{m_{k''}}$ and all other agents have the same assignment as in α'. Then, we have $\bar{r}(\alpha'') < \bar{r}(\alpha')$, i.e., this lowers the average rank, which contradicts that $\alpha' \in \psi^{RM}(\succ_i, \succ'_{-i})$. Therefore, $\max_{\succ'_{-i}} \bar{\rho}_i(\psi^{RM}(\succ_i, \succ'_{-i})) = k^*$.

Thus, I have shown that when i reports her true preferences, her worst-case outcome is $\max_{\succ'_{-i}} \bar{\rho}_{-i}(\psi^{RM}(\succ_i, \succ'_{-i})) = k^*$. What remains to show is that for any misreport $\succ'_i \neq \succ_i$, we have $\max_{\succ'_{-i}} \bar{\rho}_i(\psi^{RM}(\succ'_i, \succ'_{-i})) \geq k^*$, where of course $\bar{\rho}_i$ is evaluated with respect to i's true preferences. Consider a misreport $\succ'_i \neq \succ_i$. For notational purposes, index this preference profile as

$$\succ'_i : o_{r_1}, o_{r_2}, \ldots, o_{r_M}.$$

Let $k^{**} = \min\{k : \sum_{k'=1}^{k} q_{r_{k'}} \geq N\}$. Similar to Lemma 1, this is the index of the critical object in the sense that the set $O'' = \{o_{r_1}, \ldots, o_{r_{k^{**}}}\}$ has enough total capacity for all agents, but the set $\{o_{r_1}, \ldots, o_{r_{k^{**}-1}}\}$ does not. Consider a preference profile $\succ'_I$ such that $\succ'_j = \succ'_i$ for all $j \in I$. By Lemma 1, we have $\psi^{RM}(\succ'_I) = A^{**}$, where A^{**} is defined analogously to A^* above, replacing O' with the set O''.

Case 1: $O'' = \{o_1, \ldots, o_{m_{k^*}}\}$.

By Lemma 1, there is at least one allocation $\alpha \in \psi^{RM}(\succ'_I)$ such that $\alpha_i = o_{m_{k^*}}$. Therefore, $\max_{\succ'_{-i}} \bar{\rho}_i(\psi^{RM}(\succ'_i, \succ'_{-i}) \geq k^*$, as desired.

Case 2: $O'' \neq \{o_1, \ldots, o_{m_{k^*}}\}$.

In this case, there is some $\tilde{o} \in O''$ such that $o_{m_{k^*}} \succ_i \tilde{o}$ according to i's true preferences $\succ_i$. By Lemma 1, there is at least one allocation $\alpha \in \psi^{RM}(\succ'_I)$ such that $\alpha_i = \tilde{o}$. Because $o_{m_{k^*}} \succ_i \tilde{o}$, this implies that $\max_{\succ'_{-i}} \bar{\rho}_i(\psi^{RM}(\succ'_i, \succ'_{-i}) > k^*$, as desired.

This completes the argument for part (i).

Part (ii): $\min_{\succ_{-i}} \varrho_i(\psi(\succ_i, \succ_{-i})) \leq \min_{\succ_{-i}} \varrho_i(\psi(\succ'_i, \succ_{-i}))$.

Without loss of generality, consider agent i_1 whose preferences are $\succ_1$: $o_1, \ldots, o_M$. As in part (i), let $m^* = \min\{m' : \sum_{m=1}^{m'} q_{m'} \geq N\}$. Consider preference profile $\succ_{-i}$ for the other agents constructed as follows:

- Exactly $q_1 - 1$ agents have preferences such that $\succ_j$: $o_1, \ldots$.
- For all $m' = 2, \ldots, m^* - 1$, exactly $q_{m'}$ agents have preferences such that $\succ_j$: $o_{m'}, \ldots$.
- Exactly $N - \sum_{m'=1}^{m^*-1} q_{m'}$ agents have preferences such that $\succ_j$: $o_{m^*}, \ldots$.

In words, the constructed profile $\succ_I$ is such that each object o_m has exactly q_m agents who have ranked it first. This is possible by the definition of m^* and the assumption that $\sum_m q_m \geq N$. Now, notice that at this profile, there is a unique rank-minimizing allocation, $\psi^{RM}(\succ_I) = \{\alpha^*\}$, where α^* is the allocation such that each agent is assigned to her first-choice object. Thus, $\varrho_i(\psi(\succ_i, \succ_{-i})) = 1$, and so $\min_{\succ_{-i}} \varrho_i(\psi(\succ_i, \succ_{-i})) = 1$. Since it is obvious that $\varrho_i(\psi(\succ'_i, \succ_{-i})) \geq 1$ for any $(\succ'_i, \succ_{-i})$, we have $\min_{\succ_{-i}} \varrho_i(\psi(\succ_i, \succ_{-i})) \leq \min_{\succ_{-i}} \varrho_i(\psi(\succ'_i, \succ_{-i}))$, and therefore part (ii) of Definition 1 holds.

■

References

Abdulkadiroğlu, A., Y.-K. Che, P. A. Pathak, A. E. Roth, and O. Tercieux (2020): "Efficiency, Justified Envy, and Incentives in Priority-Based Matching," *American Economic Review: Insights*, 2, 425–442.

Abdulkadiroğlu, A. and T. Sönmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Alva, S. and V. Manjunath (2019): "Strategy-proof Pareto-improvement," *Journal of Economic Theory*, 181, 121–142.

Archbold, T., B. de Keijzer, and C. Ventre (2022): "Non-Obvious Manipulability for Single-Parameter Agents and Bilateral Trade," *arXiv preprint arXiv:2202.06660.*

Azevedo, E. M. and E. Budish (2019): "Strategy-proofness in the large," *The Review of Economic Studies*, 86, 81–116.

Aziz, H. and A. Lam (2021): "Obvious Manipulability of Voting Rules," in *International Conference on Algorithmic Decision Theory*, Springer, 179–193.

Budish, E. and E. Cantillon (2012): "The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," *American Economic Review*, 102, 2237–71.

Budish, E. B. and J. B. Kessler (2017): "Can Agents "Report Their Types"? An Experiment that Changed the Course Allocation Mechanism at Wharton," *Chicago Booth Research Paper*.

Cerrone, C., Y. Hermstrüwer, and O. Kesten (2022): "School choice with consent: an experiment," *MPI Collective Goods Discussion Paper*.

Combe, J., O. Tercieux, and C. Terrier (2022): "The design of teacher assignment: Theory and evidence," *Review of Economic Studies, forthcoming*.

Delacrétaz, D., S. D. Kominers, and A. Teytelboym (2016): "Refugee resettlement," *University of Oxford Department of Economics Working Paper*.

Dur, U., R. G. Hammond, and T. Morrill (2018): "Identifying the harm of manipulable school-choice mechanisms," *American Economic Journal: Economic Policy*, 10, 187–213.

Featherstone, C. R. (2020): "Rank efficiency: Investigating a widespread ordinal welfare criterion," Working paper.

Gale, D. and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 69, 9–15.

Li, J. and P. Dworczak (2021): "Are simple mechanisms optimal when agents are unsophisticated?" in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 685–686.

Li, S. (2017): "Obviously Strategy-Proof Mechanisms," *American Economic Review*, 107, 3257–87.

Nikzad, A. (2022): "Rank-optimal assignments in uniform markets," *Theoretical Economics*, 17, 25–55.

Ortega, J. and T. Klein (2022): "Improving Efficiency and Equality in

School Choice," .

Ortega, J. and E. Segal-Halevi (2019): "Obvious manipulations in cake-cutting," *arXiv preprint arXiv:1908.02988*.

Pathak, P. A. and T. Sönmez (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," 98, 1636–1652.

Roth, A. E. and E. Peranson (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," 89, 748–780.

Roth, A. E. and U. G. Rothblum (1999): "Truncation Strategies in Matching Markets-In Search of Advice for Participants," *Econometrica*, 67, 21–43.

Sethuraman, J. (2022): "A note on the average rank of rank-optimal assignments," Working paper, Columbia University.

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

Troyan, P., D. Delacrétaz, and A. Kloosterman (2020): "Essentially stable matchings," *Games and Economic Behavior*, 120, 370–390.

Troyan, P. and T. Morrill (2020): "Obvious manipulations," *Journal of Economic Theory*, 185, 104970.