FACULTAD ECONOMÍA Y NEGOCIOS
UNIVERSIDAD DE CHILE
DEPARTAMENTO DE ECONOMÍA

SERIE DE DOCUMENTOS DE TRABAJO

SDT 553

STRATEGIC BEHAVIOR WITHOUT OUTSIDE OPTIONS

Autores:
Camilo J. Sirguiado
Juan Pablo Torres-Martínez

Santiago, Abril de 2024

sdt@econ.uchile.cl
econ.uchile.cl/publicaciones

STRATEGIC BEHAVIOR WITHOUT OUTSIDE OPTIONS

CAMILO J. SIRGUIADO AND JUAN PABLO TORRES-MARTÍNEZ

Abstract. In two-sided one-to-one matching markets, each side of the market has a single stable mechanism that is strategy-proof for its members (Alcalde and Barberà, 1994). When agents may not declare potential partners inadmissible, this uniqueness result only holds for the short side, if there is one. Furthermore, among the stable mechanisms that are strategy-proof for the long side of the market, there is one that is less manipulable by coalitions of its members than the long-side optimal deferred acceptance mechanism. These properties can be extended to scenarios in which a part of the population has outside options.

Keywords: Matching markets - Outside options - Strategy-proofness - Stability
JEL Classification: D47, C78.

1. Introduction

Matching theory has established itself as an active field of research, thanks to its successful applicability to the design of school choice mechanisms, college admission systems, object allocation rules, and kidney exchange platforms.[1] One area of interest within this field has been the study and characterization of *strategy-proof mechanisms*, that is, centralized protocols that induce agents to truthfully report their preferences independently of the actions of others. Strategy-proofness simplifies the interpretation of reported information, and promotes equity by not giving an advantage to more sophisticated agents.[2] Under a strategy-proof mechanism, policymakers should expect social goals to be achieved even when preferences are not observable.

It is well-known that designing strategy-proof mechanisms with good stability or efficiency properties may be impossible. In two-sided one-to-one matching markets, Roth (1982) shows that no stable mechanism is strategy-proof, while Alcalde and Barberà (1994) establish that no Pareto

Date: April, 2024.

We are grateful to Gabriela Denis, Jorge Arenas, Marta Faias, Matteo Triossi, Roberto Jara, Rodrigo Yáñez, and Sebastián Cea-Echenique for useful comments and suggestions.

Camilo J. Sirguiado (cjaramillos@uandes.cl)
School of Business and Economics, Universidad de los Andes, Chile

Juan Pablo Torres-Martínez (juan.torres@fen.uchile.cl)
Department of Economics, Faculty of Economics and Business, University of Chile.

[1]We refer to Echenique, Inmorlica, and Vazirani (2023) for surveys on the applications of matching theory.

[2]Evidently, factors such as mistrust, altruism, or cognitive inability can induce agents to misrepresent preferences even when a strategy-proof mechanism is implemented (see Hassidim, Marciano, Romm, and Shorrer, 2017).

efficient and individually rational mechanism is strategy-proof. However, stability and one-sided strategy-proofness are compatible: the *deferred acceptance mechanism* (Gale and Shapley, 1962) implements the best stable matching for those who make proposals and is strategy-proof for them (cf., Dubins and Freedman, 1981; Roth, 1982). Consequently, this mechanism is a good alternative when a policymaker is concerned about the welfare and incentives of the same side of the market. Nevertheless, there are situations in which one might want to maximize the welfare of one side of the market without incentivizing the other side to misreport preferences. For example, the matching between courses and instructors at a college, or the assignment of tasks to employees of a company. In any of these scenarios, the institution may want to implement a stable mechanism that both maximizes its welfare (i.e., the welfare of the side of the market representing courses/tasks) and incentivizes instructors/employees to honestly report their preferences.

Unfortunately, when potential partners can be declared inadmissible, the only way to incentivize one side of the market to truthfully report preferences is to minimize the welfare of the other side. Indeed, Alcalde and Barberá (1994, Theorem 3) show that deferred acceptance is the only stable mechanism that is strategy-proof for those who make proposals. Since the best stable matching for one side of the market is the worst for the other (Knuth, 1976), deferred acceptance minimizes the welfare of those who receive proposals. Underlying Alcalde and Barberà's uniqueness result is the fact that, when a stable mechanism other than deferred acceptance is implemented, each agent on the side of the market that makes proposals can get the best stable partner by simply declaring the mate under deferred acceptance as the only acceptable one.[3] However, when no one may declare others inadmissible, these manipulation strategies cannot be used, and it is natural to ask whether the uniqueness result still holds.[4]

We study strategic behavior when agents have no outside options and the mechanism designer knows it. Under this information requirement, the lack of outside options—a characteristic of preferences—can be associated with the inability to misrepresent preferences by declaring potential partners unacceptable. Focusing on two-sided one-to-one matching markets, we show that:

- In balanced matching markets, for each side of the market there are many stable mechanisms that are strategy-proof for its members.[5]
- In unbalanced matching markets, only the long side of the market has more than one stable mechanism that is strategy-proof for its members.
- Among the stable mechanisms that are strategy-proof for the long side of the market, there is one that is less manipulable by coalitions of its members than the long-side optimal deferred acceptance mechanism.

[3]The Rural Hospital Theorem (Gale and Sotomayor, 1985) ensures the success of this manipulation strategy, since it implies that those who form a pair when deferred acceptance is implemented will remain paired with someone under any other stable outcome. Therefore, regardless of the stable mechanism that is implemented, agents who declare their mate under deferred acceptance as the only acceptable one are not left alone.

[4]When agents may not declare potential partners inadmissible, Roth's (1982, Theorem 3) impossibility theorem holds unless one side of the market has only two agents (see Remark 2). Furthermore, since individual rationality is trivially satisfied in this context, the impossibility result of Alcalde and Barberà (1994, Proposition 1) never holds: the *serial dictatorship algorithm* is Pareto efficient, individually rational, and strategy-proof (cf., Svensson, 1999).

[5]A matching market is *balanced* when there is an equal number of agents on each side.

These results show how the ability to manipulate a stable mechanism decreases in the absence of outside options. The first two properties fully characterize the scenarios in which Alcalde and Barberà's uniqueness result hold when no one may declare others unacceptable (see Proposition 1). The third property ensures that it is possible to improve the welfare of short-side agents and reduce the manipulability of long-side coalitions without compromising stability (see Proposition 2).

In unbalanced markets, to prove that there is more than one stable and strategy-proof mechanism for the long side of the market, we describe a family of centralized protocols that combine the two versions of deferred acceptance (those obtained by choosing one side of the market to make proposals). Essentially, we find a subdomain of preferences in which the impossibility of declaring others unacceptable prevents agents on the long side of the market from manipulating the outcome of deferred acceptance when the short side makes the proposals. Thus, we construct a mechanism that is strategy-proof for the long side of the market by associating the *short-side optimal* stable matching in that subdomain and the *long-side optimal* stable matching otherwise (see the proof of Proposition 1). For balanced matching markets, a similar strategy allow us to prove that each side of the market has many stable mechanisms that are strategy-proof for its members.

To ensure that there is a single stable mechanism that is strategy-proof for the short side of an unbalanced market, we appeal to the Rural Hospital Theorem (Gale and Sotomayor, 1985). Given a preference profile, this result guarantees that there is an agent on the long side who is left alone in all stable matchings. Thus, when a stable mechanism other than the *short-side optimal* deferred acceptance mechanism is implemented, there are preference profiles at which some agents on the short side can benefit from declaring as first-rated alternatives their best stable partner and someone who is left alone in any stable matchings, in this order (see the proof of Proposition 1).

We also characterize a stable mechanism that improves the welfare of the short side of the market without incentivizing the long side to misreport preferences: match each agent on the short side of the market with his best potential partner; when it is not possible, implement the long-side optimal stable matching. We show that this mechanism is also less manipulable by *coalitions* of the long side of the market than the long-side optimal deferred acceptance mechanism (see Proposition 2).

It is important to note that—regardless of the existence of outside options—our results remain valid when the mechanism designer only knows that everyone considers all potential partners admissible. Indeed, this ensures that an agent may not misrepresent preferences by declaring a potential partner unacceptable, which is the main assumption underlying our findings.[6]

Our analysis can be extended to allow a number of agents to have outside options that they consider better than some potential partners (see Theorems 1, 2, and 3). Among other properties, in this more general framework we show that:

- In balanced markets, Alcalde and Barberà's uniqueness result holds for one side of the market if and only if at most one agent on that side may not declare inadmissibilities. Moreover, the same property holds for the long side in an unbalanced market.

[6] It might be thought that it is sufficient to force agents (implicitly or explicitly) to report all potential partners as admissible. For instance, this happens in school choice systems in which, to prevent anyone from being excluded, it is assumed that all students find neighborhood schools acceptable (cf., Teo, Sethuraman, and Tan, 2001). However, when agents may not be able to report their true preferences, studying strategy-proofness does not make sense.

- If all agents on one side of the market have outside options, then there is more than one stable mechanism that is strategy-proof for the other side if and only if at least two of its members may not declare their potential partners as unacceptable.
- If at least two agents on the long side of the market may not declare their potential partners as unacceptable, then there is a stable mechanism that is strategy-proof for the long side of the market and is less manipulable by coalitions of its members than the long-side optimal deferred acceptance mechanism.

Furthermore, among the stable mechanisms that are strategy-proof for one side of the market, there is always one that is Pareto-superior for the other side (see Theorem 4). In many situations, this optimal mechanism does not implement any of the outcomes of deferred acceptance.

To contextualize our results, consider a company that wants to assign tasks to a group of employees taking into account their preferences. Suppose that for each task there is a team leader who ranks workers according to their ability to perform it; and at least two employees evaluate all chores as acceptable. In this context, when there are as many employees as tasks or some employees can be considered unsuitable for some duties, our results ensure that there is a stable mechanism that improves the welfare of the company in relation to the *employee-optimal stable mechanism* and is still strategy-proof for the workers. Furthermore, this mechanism can be chosen in such a way that the efficiency of task allocation cannot be improved—from the company's perspective—without incentivizing employees to lie about their preferences.

Related literature. To the best of our knowledge, there are no studies that examine the validity of the uniqueness result of Alcalde and Barberá (1994) in contexts where agents may not declare potential partners as unacceptable. Teo, Sethuraman, and Tan (2001) highlight how discarding the strategic option of remaining single may significantly affects the ability of some agents to manipulate a stable mechanism. For balanced marriage markets, they show that the inability to declare potential partners unacceptable prevents a woman from reaching her best stable partner by manipulating the men-optimal deferred acceptance mechanism. Notice that, this never happens when women are the short side of the market, because any man who is left alone in a stable matching can be used to reach the best stable partner (see the proof of our Proposition 1).

For school choice problems in which institutions and students may not declare potential partners unacceptable, the results of Kesten (2010, Proposition 1) and Kesten and Kurino (2019, Corollary 3) characterize the existence of mechanisms that, from the perspective of students, are strategy-proof and Pareto dominate the student-optimal deferred acceptance mechanism.[7] They show that this type of mechanism exists if and only if the school system has more students than vacancies. We complement this property, as our Proposition 1 and Theorem 4 imply that there is a (stable) mechanism that Pareto dominates the student-optimal deferred acceptance mechanism *from the perspective of schools*, and remains strategy-proof for students, if and only if there are as many students as vacancies in the schools system (see Section 5).

[7]From the perspective of students, when schools may be declared unacceptable, no strategy-proof mechanism Pareto dominates the student-optimal deferred acceptance mechanism (see Abdulkadiroglu, Pathak, and Roth, 2009; Erdil, 2014).

The rest of the paper is organized as follows. Section 2 describes our framework. Sections 3 and 4 analyze the validity of Alcalde and Barberá (1994, Theorem 3) when (some) agents may not report inadmissibilities. Section 5 shows that each side of the market has an optimal stable mechanism among those that are strategy-proof for the other side. Remarks on topics for future research are included in Section 6. Some proofs are left to the Appendix.

2. Model

We study matching markets in which the mechanism designer knows that agents have no outside options. Hence, no one may misrepresent preferences by declaring potential partners as inadmissible.

Let $[M, W, (\succ_i)_{i\in M\cup W}]$ be a two-sided one-to-one matching market in which the population is divided into two finite sets, M and W, with at least two agents each. Given $H \in \{M, W\}$, each agent $h \in H$ has a complete, transitive, and strict preference $\succ_h$ defined on H^c, where $M^c \equiv W$ and $W^c \equiv M$. Let $\mathcal{P}$ be the set of preference profiles $\succ = (\succ_i)_{i\in M\cup W}$ satisfying the conditions above.

A *matching* is a function $\mu : M \cup W \to M \cup W$ determining a partner for each agent in $M \cup W$. That is, $\mu(h) \in H^c \cup \{h\}$ and $\mu(\mu(h)) = h$ for each $H \in \{M, W\}$ and $h \in H$. Let $\mathcal{M}$ be the set of matchings between M and W. A matching μ is *stable* when no pair of agents can block it, in the sense that there is no pair $(m, w) \in M \times W$ such that $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. Since there are no outside options, no one is interested in blocking a matching to be alone.

A mechanism is a centralized protocol that associates a matching to each preference profile. Given a mechanism $\Omega : \mathcal{P} \to \mathcal{M}$ and a side of the market $H \in \{M, W\}$, consider the following properties:

- Ω is *stable* when for each $\succ \in \mathcal{P}$ the matching $\Omega[\succ]$ is stable in $[M, W, \succ]$.
- Ω is *strategy-proof for H* when for any preference profiles $\succ, \succ' \in \mathcal{P}$ there is no agent $h \in H$ such that $\Omega[\succ'_h, \succ_{-h}](h) \succ_h \Omega[\succ](h)$, where $\succ_{-h} = (\succ_i)_{i\neq h}$.[8]
- Ω is *manipulable by a coalition $C \subseteq H$ at a preference profile* $\succ \in \mathcal{P}$ when there exists $\widetilde{\succ} \in \mathcal{P}$ such that the following conditions hold:
  - Each agent $h \in C$ considers $\Omega[\widetilde{\succ}_C, \succ_{-C}](h)$ at least as preferred as $\Omega[\succ](h)$,
  - For some $h \in C$, $\Omega[\widetilde{\succ}_C, \succ_{-C}](h) \succ_h \Omega[\succ](h)$,

  where $\widetilde{\succ}_C = (\widetilde{\succ}_h)_{h\in C}$ and $\succ_{-C} = (\succ_h)_{h\notin C}$.
- Ω is *less manipulable by H-coalitions* than $\widetilde{\Omega} : \mathcal{P} \to \mathcal{M}$ as long as:
  - For each $C \subseteq H$ and $\succ \in \mathcal{P}$, if Ω is manipulable by C at $\succ$, then $\widetilde{\Omega}$ is too.
  - There are $C \subseteq H$ and $\succ \in \mathcal{P}$ such that, $\widetilde{\Omega}$ is manipulable by C at $\succ$ and Ω is not.
- Ω is *group strategy-proof for H* when it is not manipulable by any coalition of agents in H.

Let $\text{DA}_H : \mathcal{P} \to \mathcal{M}$ be the stable mechanism that associates to each preference profile in $\mathcal{P}$ the outcome of the *deferred acceptance algorithm* when agents in $H \in \{M, W\}$ make proposals. Some classic properties of DA_H will be key to prove our main results:

- $\text{DA}_H[\succ]$ is the best stable matching of $[M, W, \succ]$ for agents in H (Gale and Shapley, 1962).
- $\text{DA}_H : \mathcal{P} \to \mathcal{M}$ is strategy-proof for H (Dubins and Freedman, 1981; Roth, 1982).

[8]Therefore, Ω is strategy-proof for H when truth-telling is a dominant strategy for agents in H in the game in which every $i \in M \cup W$ reports preferences $\succ_i$ and $\Omega[(\succ_i)_{i\in M\cup W}]$ is implemented.

We will also appeal to the *Rural Hospital Theorem* (Gale and Sotomayor, 1985): those who are single in a stable matching of $[M, W, \succ]$ remain single in any other stable outcome.

Assuming that agents have the ability to declare potential partners as unacceptable, Alcalde and Barberà (1994, Theorem 3) guarantee that DA_H is the only stable mechanism that is strategy-proof for the members of $H \in \{M, W\}$. We will show that this result may not hold in contexts in which the mechanism designer knows that no one has outside options.

3. One-sided strategy-proofness and stability

In this section, we prove that in a two-sided one-to-one matching market in which the mechanism designer knows that agents have no outside options, Alcalde and Barberà's uniqueness result only holds for the short side of the market, if there is one.

Notice that, if an agent $m \in M$ may not declare potential partners as inadmissible, he loses the ability to manipulate a stable mechanism other than DA_M by simply declaring $\mathrm{DA}_M[\succ](m)$ as the *only* acceptable partner. This manipulation strategy is a key ingredient in the classic proof of Alcalde and Barberà's uniqueness result. Therefore, the assumption that the mechanism designer knows that there are no outside options may compromise the validity of this result.

However, when M is the short side of the market, it is possible to show that DA_M remains the only stable mechanism that is strategy-proof for M. More precisely, if $|M| < |W|$, then for each preference profile $\succ \in \mathcal{P}$ there exists an agent $w^* \in W$ that stays alone in all stable matchings of $[M, W, \succ]$. Hence, when a stable mechanism other than DA_M is implemented, each $m \in M$ can report $\mathrm{DA}_M[\succ](m)$ and w^* as the best alternatives, in this order. This manipulation strategy ensures that $\mathrm{DA}_M[\succ]$ remains stable under the new preferences and, as a consequence of the Rural Hospital Theorem, w^* stands alone. In addition, m never forms a stable pair with someone less preferred to w^*, because w^* always considers him acceptable. Hence, by misrepresenting preferences in this way, m manage to pair with $\mathrm{DA}_M[\succ](m)$—the best partner he would get in a stable matching. Therefore, no stable mechanism other than DA_M is strategy-proof for M.

Nevertheless, when M is the long side of the market, the scarcity of potential partners in W will allow the existence of many stable mechanisms defined on $\mathcal{P}$ that are strategy-proof for M. Essentially, we will manage to obtain a family of mechanisms with these properties by properly combining DA_M and DA_W. To gain some intuition about this claim, consider a preference profile in which agents in W have different best potential partners. In this scenario, agents in W do not compete when the DA_W is implemented, and the inability of agents in M to declare others unacceptable makes it impossible for them to improve through misrepresenting preferences (cf., Teo, Sethuraman, and Tan, 2001). The next result formalizes the ideas described above.

Proposition 1. *In a two-sided one-to-one matching market between M and W, there is more than one stable mechanism defined in $\mathcal{P}$ that is strategy-proof for M if and only if $|M| \geq |W|$.*

Proof. ($\Longleftarrow$) Suppose that $|M| \geq |W|$. Let $\mathcal{F}^*$ be the non-empty set of injective maps $f : W \to M$. For each $\mathcal{F} \subseteq \mathcal{F}^*$, let $\mathcal{P}_{\mathcal{F}} = \{\succ \in \mathcal{P} : \exists f \in \mathcal{F},\ b_w(\succ) = f(w),\ \forall w \in W\}$, where $b_w(\succ)$ is the best potential partner of w under $\succ$. That is, $\mathcal{P}_{\mathcal{F}}$ are the preference profiles in which the best potential partners of agents in W are determined by a function in $\mathcal{F}$. In particular, $b_w(\succ) \neq b_{w'}(\succ)$ for all $w, w' \in W$ and $\succ \in \mathcal{P}_{\mathcal{F}}$. Consider the stable mechanism $\Omega_{\mathcal{F}} : \mathcal{P} \to \mathcal{M}$ characterized by

$$\Omega_{\mathcal{F}}[\succ] = \begin{cases} \text{DA}_W[\succ], & \text{when } \succ \in \mathcal{P}_{\mathcal{F}}, \\ \text{DA}_M[\succ], & \text{when } \succ \in \mathcal{P} \setminus \mathcal{P}_{\mathcal{F}}. \end{cases}$$

We claim that $\Omega_{\mathcal{F}}$ is strategy-proof for M. By contradiction, suppose that there is $m \in M$ such that $\Omega_{\mathcal{F}}[\succ'_m, \succ_{-m}](m) \succ_m \Omega_{\mathcal{F}}[\succ](m)$ for some $\succ, \succ' \in \mathcal{P}$. Notice that, $\succ \in \mathcal{P}_{\mathcal{F}}$ if and only if $(\succ'_m, \succ_{-m}) \in \mathcal{P}_{\mathcal{F}}$. If $\succ$ and $(\succ'_m, \succ_{-m})$ belong to $\mathcal{P}_{\mathcal{F}}$, as no one in M has an outside option, it follows that $\Omega_{\mathcal{F}}[\succ'_m, \succ_{-m}] = \text{DA}_W[\succ'_m, \succ_{-m}] = \text{DA}_W[\succ] = \Omega_{\mathcal{F}}[\succ]$, a contradiction. Since DA_M is strategy-proof for M, if both $\succ$ and $(\succ'_m, \succ_{-m})$ belong to $\mathcal{P} \setminus \mathcal{P}_{\mathcal{F}}$, then we have that $\Omega_{\mathcal{F}}[\succ](m) = \text{DA}_M[\succ](m) \succeq_m \text{DA}_M[\succ'_m, \succ_{-m}](m) = \Omega_{\mathcal{F}}[\succ'_m, \succ_{-m}](m)$, a contradiction.

We claim that $\Omega_{\mathcal{F}} \neq \text{DA}_M$. Suppose that $M = \{m_1, \ldots, m_r\}$ and $W = \{w_1, \ldots, w_s\}$, where $r \geq s \geq 2$. Given a function $f \in \mathcal{F}$, let $\succ \in \mathcal{P}_{\mathcal{F}}$ be such that $(b_w(\succ))_{w \in W} = (f(w))_{w \in W}$ and

- For $i, j \in \{1, 2\}$ with $i \neq j$, w_i is the best potential partner of $b_{w_j}(\succ) \in M$.
- For $i \in \{3, \ldots, s\}$, agents $b_{w_i}(\succ)$ and w_i consider each other the best alternative.
- The agents $m_1, \ldots, m_s$ are the top alternatives for each $w \in W$.

It is not difficult to verify that

$$\begin{aligned} \text{DA}_M[\succ] &= \{(b_{w_1}(\succ), w_2), (b_{w_2}(\succ), w_1), (b_{w_3}(\succ), w_3), \ldots, (b_{w_s}(\succ), w_s), m_{s+1}, \ldots, m_r\}, \\ \text{DA}_W[\succ] &= \{(b_{w_1}(\succ), w_1), (b_{w_2}(\succ), w_2), (b_{w_3}(\succ), w_3), \ldots, (b_{w_s}(\succ), w_s), m_{s+1}, \ldots, m_r\}. \end{aligned}$$

Hence, $\Omega_{\mathcal{F}}[\succ] = \text{DA}_W[\succ] \neq \text{DA}_M[\succ]$. Therefore, when $|M| \geq |W|$ there is more than one stable mechanism defined in $\mathcal{P}$ that is strategy-proof for M.[9]

($\Longrightarrow$) Suppose that $|M| < |W|$. We want to prove that there is a single stable mechanism defined in $\mathcal{P}$ that is strategy-proof for M. By contradiction, assume that there exists $\Omega : \mathcal{P} \to \mathcal{M}$, stable and strategy-proof for M, satisfying $\Omega[\succ] \neq \text{DA}_M[\succ]$ for some $\succ \in \mathcal{P}$. Since DA_M is the M-optimal stable mechanism (Gale and Shapley, 1962; Theorem 2), it follows that $\text{DA}_M[\succ](m) \succ_m \Omega[\succ](m)$ for some $m \in M$. In particular, $\text{DA}_M[\succ](m) \in W$. Furthermore, $|M| < |W|$ implies that there exists $w^* \in W$ such that $\text{DA}_M[\succ](w^*) = w^*$. Let $\succ'_m$ be a preference profile in which $\text{DA}_M[\succ](m)$ and w^* are the best alternatives for m, in this order. Since $\text{DA}_M[\succ]$ is stable under $(\succ'_m, \succ_{-m})$, the Rural Hospital Theorem (Gale and Sotomayor, 1985, Theorem 1) implies that w^* is alone in every stable matching of $[M, W, (\succ'_m, \succ_{-m})]$. Hence, $\Omega[\succ'_m, \succ_{-m}](m) \succ'_m w^*$, because otherwise (m, w^*) would block the stable matching $\Omega[\succ'_m, \succ_{-m}]$ (remember that m is always acceptable for w^*).[10] As a consequence, $\Omega[\succ'_m, \succ_{-m}](m) = \text{DA}_M[\succ](m)$, which implies that $\Omega[\succ'_m, \succ_{-m}](m) \succ_m \Omega[\succ](m)$. This contradicts the strategy-proofness of Ω. □

[9]For each $f \in \mathcal{F}$ we found $\succ \in \mathcal{P}_{\mathcal{F}}$ such that the best potential partners of agents in W are determined by f and $\Omega_{\mathcal{F}}[\succ] \neq \text{DA}_M[\succ]$. Hence, the mechanisms $\Omega_{\mathcal{F}}$ and $\Omega_{\mathcal{F}'}$ are different as long as $\mathcal{F} \neq \mathcal{F}'$. This implies that there are at least $2^{|\mathcal{F}|} = 2^{|M|!/(|M|-|W|)!}$ stable mechanisms that are strategy-proof for M.

[10]This is the only step in the proof of Proposition 1 in which we use that agents in W may not declare their potential partners unacceptable.

Assuming that $|M| \geq |W|$, let $\mathcal{P}' \subseteq \mathcal{P}$ be the non-empty collection of preference profiles for which agents in W have different best potential partners.

Proposition 2. *In a two-sided one-to-one matching market in which $|M| \geq |W| \geq 3$, the stable mechanism $\Psi : \mathcal{P} \to \mathcal{M}$ characterized by*

$$\Psi[\succ] = \begin{cases} \mathrm{DA}_W[\succ], & \text{when } \succ \in \mathcal{P}', \\ \mathrm{DA}_M[\succ], & \text{when } \succ \in \mathcal{P} \setminus \mathcal{P}', \end{cases}$$

is strategy-proof for M and less manipulable by M-coalitions than DA_M*.*

Proof. The arguments made in the proof of Proposition 1 imply that Ψ is strategy-proof for M, because it coincides with the mechanism $\Omega_{\mathcal{F}^*}$. Suppose that Ψ is manipulable by a coalition $C \subseteq M$ at a preference profile $\succ \in \mathcal{P}$. Hence, there exists $\widetilde{\succ} \in \mathcal{P}$ such that:

- Each agent $m \in C$ considers $\Psi[\widetilde{\succ}_C, \succ_{-C}](m)$ at least as preferred as $\Psi[\succ](m)$.
- For some $m \in C$ we have that $\Psi[\widetilde{\succ}_C, \succ_{-C}](m) \succ_m \Psi[\succ](m)$.

Notice that $\succ \in \mathcal{P}'$ if and only if $(\widetilde{\succ}_C, \succ_{-C}) \in \mathcal{P}'$. If $\succ \in \mathcal{P}'$, as no agent in M may declare a potential partner inadmissible, $\Psi[\widetilde{\succ}_C, \succ_{-C}](m) = \mathrm{DA}_W[\widetilde{\succ}_C, \succ_{-C}](m) = \mathrm{DA}_W[\succ](m) = \Psi[\succ](m)$ for all $m \in C$. This contradicts the fact that C manipulates Ψ at $\succ$. Hence, $\succ \in \mathcal{P} \setminus \mathcal{P}'$ and we have that $\Psi[\widetilde{\succ}_C, \succ_{-C}](m) = \mathrm{DA}_M[\widetilde{\succ}_C, \succ_{-C}](m)$ and $\Psi[\succ](m) = \mathrm{DA}_M[\succ](m)$ for all $m \in C$, which imply that DA_M is manipulable by C at $\succ$. Therefore, if the mechanism Ψ is manipulable by a coalition $C \subseteq M$ at $\succ$, then DA_M is too.

Given that no coalition in M can manipulate Ψ at a preference profile in $\mathcal{P}'$, to conclude the proof it is sufficient to find $C \subseteq M$ and $\succ \in \mathcal{P}'$ such that DA_M is manipulable by C at $\succ'$. Asume that $M = \{m_1, \ldots, m_r\}$ and $W = \{w_1, \ldots, w_s\}$, with $r \geq s \geq 3$. Let $\succ \in \mathcal{P}'$ be a preference profile such that

$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$	$\succ_{m_4}$	$\cdots$	$\succ_{m_s}$	$\succ_{m_{s+1}}$	$\cdots$	$\succ_{m_r}$	$\succ_{w_1}$	$\succ_{w_2}$	$\succ_{w_3}$	$\cdots$	$\succ_{w_s}$
w_2	w_1	w_1	w_4	$\cdots$	w_s	w_1	$\cdots$	w_1	m_1	m_2	m_3	$\cdots$	m_s
w_1	w_2	w_3	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	m_3	m_1	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	m_2	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Notice that $\mathrm{DA}_M[\succ] = \{(m_1, w_1), (m_2, w_2), (m_3, w_3), \ldots, (m_s, w_s), m_{s+1}, \ldots, m_r\}$.

Let $C = \{m_1, m_2, m_3\}$ and $\widetilde{\succ} \equiv (\widetilde{\succ}_{m_3}, \succ_{-m_3})$, where $w_3 \,\widetilde{\succ}_{m_3}\, w_1 \,\widetilde{\succ}_{m_3} \cdots$. Since the matching $\mathrm{DA}_M[\widetilde{\succ}_C, \succ_{-C}]$ is given by $\{(m_1, w_2), (m_2, w_1), (m_3, w_3), \ldots, (m_s, w_s), m_{s+1}, \ldots, m_r\}$, the coalition C manipulates DA_M at $\succ$. Therefore, Ψ is less manipulable by M-coalitions than DA_M. □

In markets without (attractive) outside options in which $|M| \geq |W| \geq 3$, the mechanism Ψ is a good alternative to DA_M: it improves the welfare of the short side of the market, reduces the incentives of groups of agents on the long side to misreport their preferences, and has the same computational complexity as DA_M.

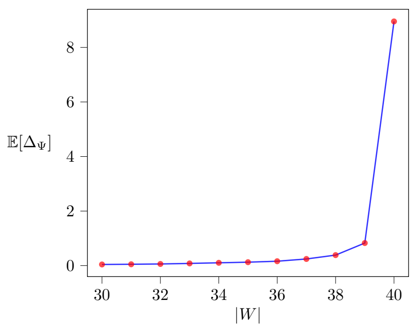
When Ψ is implemented instead of DA_M, the improvement in the welfare of agents in W at a preference profile $\succ \in \mathcal{P}$ can be quantified by a *utilitarian index* that measures the increase in the average ranking of mates:

$$\Delta_\Psi[\succ] \equiv \frac{1}{|W|}\sum_{w\in W} R_{\succ,w}[\text{DA}_M(w)] - \frac{1}{|W|}\sum_{w\in W} R_{\succ,w}[\Psi(w)],$$

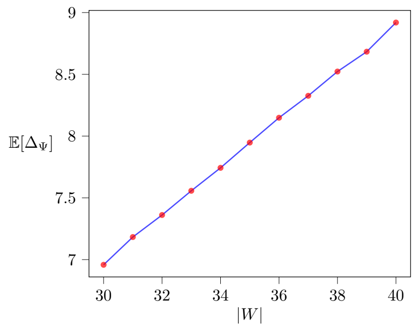
where $R_{\succ,w}[m] = |\{h \in M : h \succeq_w m\}|$ is the ranking of m in the preferences of w. Alternatively, we can use a *Rawlsian index* that measures the increase in the ranking of the worst mate:

$$\Theta_\Psi[\succ] \equiv \max_{w\in W} R_{\succ,w}[\text{DA}_M(w)] - \max_{w\in W} R_{\succ,w}[\Psi(w)].$$

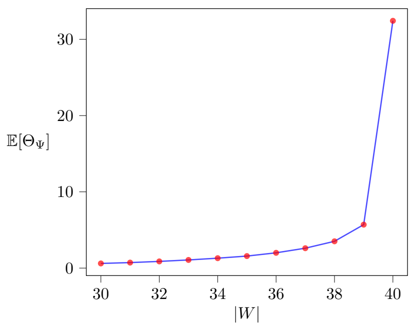
Assuming that all preference profiles in $\mathcal{P}'$ are equiprobable, the following figures describe the evolution of the expected values of these indices, denoted by $\mathbb{E}[\Delta_\Psi]$ and $\mathbb{E}[\Theta_\Psi]$.



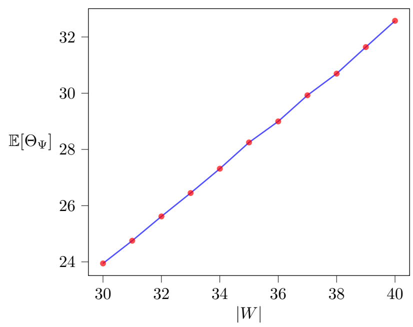
Unbalanced markets ($|M| = 40$)



Balanced markets ($|M| = |W|$)



Unbalanced markets ($|M| = 40$)



Balanced markets ($|M| = |W|$)

Note: $\mathbb{E}[\Delta_\Psi]$ and $\mathbb{E}[\Theta_\Psi]$ have been approximated by averaging 10,000 realizations of $\Delta_\Psi[\succ]$ and $\Theta_\Psi[\succ]$, where preference profiles were randomly generated by a uniform distribution supported on $\mathcal{P}'$.

It follows that the welfare of the agents in W increases monotonically with the population size when Ψ is implemented instead of DA_M. Ashlagi, Kanoria, and Leshno (2017) show that agents on the short side of the market are matched with one of their top choices in any stable matching. This property is what underlies the fact that $\mathbb{E}[\Delta_\Psi]$ and $\mathbb{E}[\Theta_\Psi]$ take low values when $|W| < |M|$.

Notice that Ψ is considerably fairer than DA_M when $|M| = |W|$, because the ranking of the worst partner of agents in W improves significantly when Ψ is implemented instead of DA_M.

4. Extensions

In this section, we discuss the validity of the uniqueness result of Alcalde and Barberà (1994) in scenarios in which some agents may declare potential partners as unacceptable.

In a two-sided one-to-one matching market between M and W, denote by $H_{\otimes}$ the set of agents in the side of the market $H \in \{M, W\}$ that always consider all potential partners acceptable. We assume that the identity of the agents in $H_{\otimes}$ is known by the mechanism designer and, therefore, they never misrepresent preferences by declaring some potential partners inadmissible. Let $\mathcal{Q}(M_{\otimes}, W_{\otimes})$ be the collection of preference profiles $(\succ_i)_{i \in M \cup W}$ such that, given $H \in \{M, W\}$:

- For each $h \in H_{\otimes}$, $\succ_h$ is a complete, transitive, and strict preference defined on H^c.
- For each $h \in H \setminus H_{\otimes}$, $\succ_h$ is a complete, transitive, and strict preference defined on $H^c \cup \{h\}$.

Notice that, $\mathcal{P} = \mathcal{Q}(M, W) \subseteq \mathcal{Q}(M_{\otimes}, W_{\otimes})$ for all $M_{\otimes} \subseteq M$ and $W_{\otimes} \subseteq W$. Given a preference profile $\succ \in \mathcal{Q}(M_{\otimes}, W_{\otimes})$, a matching $\mu \in \mathcal{M}$ is *stable* under $\succ$ when the following properties hold:

- *Individual rationality*: there is no $i \in (M \setminus M_{\otimes}) \cup (W \setminus W_{\otimes})$ such that $i \succ_i \mu(i)$.
- There is no $(m, w) \in M \times W$ such that $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$.

If $\mathcal{Q} \equiv \mathcal{Q}(M_{\otimes}, W_{\otimes})$, denote by $\mathbb{S}_H(\mathcal{Q})$ the non-empty set of stable mechanisms $\Omega : \mathcal{Q} \to \mathcal{M}$ that are strategy-proof for agents in $H \in \{M, W\}$. In terms of our notation, when everyone may declare others unacceptable ($M_{\otimes} = W_{\otimes} = \emptyset$), the uniqueness result of Alcalde and Barberà (1994) shows that $|\mathbb{S}_M(\mathcal{Q})| = |\mathbb{S}_W(\mathcal{Q})| = 1$.

The following result extends Proposition 1 to scenarios in which some agents may misreport their preferences by appealing to outside options.

Theorem 1. *When $W_{\otimes} = W$ and $|M_{\otimes}| \geq 2$, $|\mathbb{S}_M(\mathcal{Q})| > 1$ if and only if $|M| \geq |W|$.*

Among other properties, the next result determines sufficient conditions for the existence of many stable mechanisms that are strategy-proof for the *short side* of the market, a situation that never arises when no one has outside options. Essentially, if all agents on one side of the market may declare potential partners as inadmissible, then there are multiple stable mechanisms that are strategy-proof for the other side if and only if at least two of its members have no outside option.

Theorem 2. *When either $W_{\otimes} = \emptyset$ or $|M| \geq |W|$, $|\mathbb{S}_M(\mathcal{Q})| > 1$ if and only if $|M_{\otimes}| \geq 2$.*

To illustrate this result, consider a college that wants to implement a matching between introductory courses of economics (W) and instructors (M). Suppose that there are at least two instructors who consider all courses acceptable ($|M_{\otimes}| \geq 2$). For each course, instructors are ranked based on their ability to teach it. In this context, when there are as many instructors as courses ($|M| \geq |W|$) or instructors can be considered unsuitable to teach some of them ($W_{\otimes} = \emptyset$), the Theorem 2 guarantees that a stable mechanism exists that improves the welfare of the college in relation to the *instructor-optimal stable mechanism* and remains strategy-proof for the teachers.

Remark 1. To guarantee that $|\mathbb{S}_M(\mathcal{Q})| > 1$ it is necessary that at least two agents in M must be unable to declare their potential partners as unacceptable. Indeed, within the preference domains $\mathcal{Q}(M_\otimes, W_\otimes)$, it follows from Theorem 2 that a *maximal domain* in which Alcalde and Barberá's uniqueness result does not hold for M is characterized by $|M_\otimes| = 2$ and $W_\otimes = \emptyset$. □

Assuming that $|M| \geq |W|$, let $\Phi : \mathcal{Q}(M_\otimes, W_\otimes) \to \mathcal{M}$ be the stable mechanism satisfying

$$\Phi[\succ] = \begin{cases} \mathrm{DA}_W[\succ], & \text{when } \succ \in \mathcal{Q}', \\ \mathrm{DA}_M[\succ], & \text{when } \succ \notin \mathcal{Q}', \end{cases}$$

where $\mathcal{Q}'$ are the preference profiles in which agents in W have different best potential partners.

The following result extends Proposition 2 to scenarios in which some agents may declare potential partners inadmissible.

Theorem 3. *When* $|M| \geq |W| \geq 3$, $\Phi : \mathcal{Q}(M_\otimes, W_\otimes) \to \mathcal{M}$ *satisfies the following properties:*

(i) Φ *is strategy-proof for* M *if and only if* $M_\otimes = M$.
(ii) Φ *is less manipulable by* M*-coalitions than* DA_M *whenever* $M_\otimes = M$.
(iii) *Like the mechanism* DA_M, Φ *is not group strategy-proof for* M.

When $(M_\otimes, W_\otimes) = (M, W)$, the stable mechanisms Φ and Ψ coincide. Hence, the result of Theorem 3 implies that the preference domain $\mathcal{P}$ cannot be expanded to allow some agents in M to have outside options without losing the one-sided strategy-proofness of Ψ.

Despite the simplicity and the good properties of Φ, there are stable mechanisms that Pareto dominate it from the perspective of W and remain strategy-proof for M (see Theorem 4(iv)).

5. On W-optimal mechanisms in $\mathbb{S}_M(\mathcal{Q})$

In this section, we prove the existence of a stable mechanism that is optimal for all agents in W among those in which agents in M do not have incentives to misrepresent preferences. Given a side of the market $H \in \{M, W\}$, consider the partial order $\geqslant_H$ defined over the family of mechanisms $\Omega : \mathcal{Q} \to \mathcal{M}$, where $\mathcal{Q} \equiv \mathcal{Q}(M_\otimes, W_\otimes)$, and characterized by

$$\Omega_1 \geqslant_H \Omega_2 \qquad \Longleftrightarrow \qquad \Omega_1[\succ](h) \succeq_h \Omega_2[\succ](h), \quad \forall h \in H, \forall \succ \in \mathcal{Q}.$$

Hence, $\Omega_1 \geqslant_H \Omega_2$ if and only if, regardless of preferences in $\mathcal{Q}$, all agents in H weakly prefer the outcome of Ω_1 to that of Ω_2. As usual, $\Omega_1 >_H \Omega_2$ indicates that $\Omega_1 \geqslant_H \Omega_2$ and $\Omega_1 \neq \Omega_2$. Notice that, $\Omega_1 >_H \Omega_2$ if and only if Ω_1 is Pareto-superior to Ω_2 for agents in H.

A mechanism $\Omega_W : \mathcal{Q} \to \mathcal{M}$ is W*-optimal in* $\mathbb{S}_M(\mathcal{Q})$ when it is the greatest element of $\mathbb{S}_M(\mathcal{Q})$ under the partial order $\geqslant_W$. That is, from the point of view of agents in W, the mechanism Ω_W is Pareto-superior to any other mechanism in $\mathbb{S}_M(\mathcal{Q})$. Since agents in W compete with each other, the existence of such a mechanism seems non-trivial.

Theorem 4. *In a two-sided one-to-one matching market between M and W, there always exists a mechanism $\Omega_W : \mathcal{Q} \to \mathcal{M}$ that is W-optimal in $\mathbb{S}_M(\mathcal{Q})$. Moreover, the following properties hold:*

(i) *Each agent in $M \setminus M_\otimes$ has the same partner in all mechanisms in $\mathbb{S}_M(\mathcal{Q})$.*

(ii) *When agents have no outside options, we have that:*
- $\Omega_W \neq \mathrm{DA}_M$ *if and only if* $|M| \geq |W|$.
- $\Omega_W \neq \mathrm{DA}_W$ *if and only if* $|W| \geq 3$.

(iii) *If $|M| \geq |W| \geq 3$, then $\Omega_W \neq \mathrm{DA}_W$ and $\Omega_W \neq \mathrm{DA}_M$ if and only if $|M_\otimes| \geq 2$.*

(iv) *If $\min\{|M|, |W|\} \geq 5$ and $|M_\otimes| \geq 2$, there are preference profiles $\succ \in \mathcal{Q}$ such that*

$$\Omega_W[\succ] \notin \{\mathrm{DA}_M[\succ], \mathrm{DA}_W[\succ]\}$$

as long as either $W_\otimes = \emptyset$ or $|M| \geq |W|$.

To prove the existence of a W-optimal mechanism in $\mathbb{S}_M(\mathcal{Q})$, we will show that the set of stable and one-side strategy-proof *mechanisms* is a lattice under the same binary operations for which Knuth (1976) shows that the set of stable *matchings* has this property (see the Appendix).

Theorem 4(i) shows that agents without outside options are the only ones that can be negatively affected when a mechanism $\Omega \in \mathbb{S}_M(\mathcal{Q})$ different from DA_M is implemented.[11] Theorem 4(ii)-(iii) determines necessary and sufficient condition to ensure that the optimal mechanism Ω_W does not coincide with any version of deferred acceptance. In particular, when M is the long side of the market and W has at least three agents, Ω_W always differs from DA_W and $\Omega_W = \mathrm{DA}_M$ if and only if at most one agent in M has no outside options. When $|M_\otimes| \geq 2$ and each side of the market has more than four members, Theorem 4(iv) determines sufficient conditions to ensure the existence of preference profiles at which the optimal mechanism Ω_W differs from both DA_M and DA_W.

Remark 2. In two-sided one-to-one matching markets, Roth (1982, Theorem 3) shows that no stable mechanism is strategy-proof. Without outside options, this impossibility result holds if and only if there are at least three agents on each side of the market.

Indeed, independently of the existence of outside options, the original proof of Roth's theorem guarantees that $\min\{|M|, |W|\} \geq 3$ is sufficient to ensure that stability and strategy-proofness are incompatible. The necessity of $\min\{|M|, |W|\} \geq 3$ follows from Theorem 4(ii) and the fact that DA_H is strategy-proof for the agents in H. □

In markets without outside options, Kesten (2010, Proposition 1) and Kesten and Kurino (2019, Corollary 3) show the existence of a mechanism that is strategy-proof for the long side of the market and Pareto-dominate the long-sided optimal deferred acceptance mechanism. Our last result uses Theorem 4(ii) to complement this property:

[11] The advantages of having an outside option can reach the point of inducing agents to always prefer a manipulable mechanism over a strategy-proof one (see Akbarpour, Kapor, Neilson, Van Dijk, Zimmerman, 2022).

Corollary. *Assume that agents have no outside options and* $|M| > |W|$*. For each* $H \in \{M, W\}$ *there exists a mechanism* $\Omega : \mathcal{P} \to \mathcal{M}$ *that is strategy-proof for* M *and satisfies* $\Omega >_H \mathrm{DA}_M$*.*

6. Concluding remarks

In two-sided one-to-one matching markets, assuming that some agents may not misrepresent preferences by declaring potential partners unacceptable, we have found necessary and sufficient conditions to ensure the multiplicity of stable mechanisms that are strategy-proof for one side of the market (see Theorems 1 and 2). We also characterized a stable mechanism that is strategy-proof for the long side of the market and less manipulable by coalitions of its members than the long-side optimal deferred acceptance mechanism (see Theorem 3). Furthermore, among the stable mechanisms that are strategy-proof for one side of the market, we have shown that there is one that is Pareto superior for the other side (see Theorem 4). In many situations, which we describe in terms of the relative size of the sides of the market, this optimal mechanism does not coincide with any version of deferred acceptance.

A natural extension of our results is to the context of *matching markets with contracts*, that is, many-to-one matching markets between *hospitals* and *doctors* in which side payments are allowed. Hatfield and Milgrom (2005) show that many of the mechanism design properties of one-to-one matching markets remain valid in this context as long as doctors are *substitutes* and the *law of aggregate demand* holds.[12] Moreover, Sakai (2011) and Hirata and Kasuya (2017) show that Alcalde and Barberà's uniqueness result hold when agents may declare others unacceptable. Although this is a topic for future research, our results should be easily adaptable to this more general framework.

Appendix: omitted proofs

The Theorems 1 and 2 will be a consequence of the following result:

Lemma 1. *In a two-sided one-to-one matching market between* M *and* W*, we have that:*

(i) *If* $|M_\otimes| < 2$*, then* DA_M *is the only mechanism in* $\mathbb{S}_M(\mathcal{Q})$*.*
(ii) *If* $|M| < |W|$*,* $|M_\otimes| \geq 2$*, and* $|W_\otimes| \leq |M|$*, then* $|\mathbb{S}_M(\mathcal{Q})| > 1$*.*
(iii) *If* $|M| \geq |W|$ *and* $|M_\otimes| \geq 2$*, then* $|\mathbb{S}_M(\mathcal{Q})| > 1$*.*
(iv) *If* $|M| < |W|$ *and* $W_\otimes = W$*, then* DA_M *is the only mechanism in* $\mathbb{S}_M(\mathcal{Q})$*.*

Proof. **(i)** Suppose that $M_\otimes = \{\overline{m}\}$.[13] By contradiction, assume that there is $\Omega : \mathcal{Q} \to \mathcal{M}$ different from DA_M that is stable and strategy-proof for M. Let $\succ \in \mathcal{Q}$ such that $\Omega[\succ] \neq \mathrm{DA}_M[\succ]$. Since agents in M consider the matching $\mathrm{DA}_M[\succ]$ the best stable outcome in $[M, W, \succ]$, for each $m \in M$ we have $\mathrm{DA}_M[\succ](m) \succ_m \Omega[\succ](m)$ or $\mathrm{DA}_M[\succ](m) = \Omega[\succ](m)$. If there exists $m \neq \overline{m}$ such that

[12] Doctors are substitutes when any contract that is selected from a set of alternatives continues to be chosen when some of those alternatives are no longer available. Hospital's preferences satisfy the law of aggregate demand whenever the number of contracts chosen does not decrease as the set of alternatives expands.

[13] When $M_\otimes = \emptyset$, Alcalde and Barberà (1994, Theorem 3) guarantee that DA_M is the only mechanism in $\mathbb{S}_M(\mathcal{Q})$.

$\text{DA}_M[\succ](m) \succ_m \Omega[\succ](m)$, then $\text{DA}_M[\succ](m) \in W$. Hence, m can manipulate the mechanism Ω by reporting any preference $\succ'_m$ such that $\text{DA}_M[\succ](m) \succ'_m m \succ'_m \cdots$. Indeed, since $\text{DA}_M[\succ]$ is stable in $[M, W, (\succ'_m, \succ_{-m})]$, it follows from the Rural Hospital Theorem that $\Omega[\succ'_m, \succ_{-m}](m) = \text{DA}_M[\succ](m)$, which implies that $\Omega[\succ'_m, \succ_{-m}] \succ_m \Omega[\succ](m)$. This contradicts the strategy-proofness of Ω. Therefore, $\overline{w} \equiv \text{DA}_M[\succ](\overline{m}) \succ_{\overline{m}} \Omega[\succ](\overline{m})$. Furthermore, as $\Omega[\succ](m) = \text{DA}_M[\succ](m)$ for all $m \neq \overline{m}$, we have that $\Omega[\succ](\overline{w}) = \overline{w}$.[14] Therefore, $(\overline{m}, \overline{w})$ blocks the matching $\Omega[\succ]$ under $\succ$. A contradiction.

(ii) We want to prove that $|\mathbb{S}_M(\mathcal{Q})| > 1$ whenever $|M| < |W|$, $|M_\otimes| \geq 2$, and $|W_\otimes| \leq |M|$.

Suppose that $M = \{m_1, \ldots, m_r\}$ and $W = \{w_1, \ldots, w_s\}$, where $2 \leq r < s$. Also, $\{m_1, m_2\} \subseteq M_\otimes$ and $W_\otimes \subseteq \{w_1, \ldots, w_r\}$. Let $\widetilde{\succ} \in \mathcal{Q}$ be a preference profile that satisfies the following conditions:

(1) Given $i \in \{1, 2\}$, m_i considers w_i the best potential partner.
(2) Given $i, j \in \{1, 2\}$ with $i \neq j$, w_i considers m_j her best potential partner and m_i admissible.
(3) For each $i \in \{3, \ldots, r\}$, agents m_i y w_i consider each other the best alternative.
(4) Each agent $w \in \{w_{r+1}, \ldots, w_s\}$ considers m_1 and m_2 unacceptable.

It follows that the market $[M, W, \widetilde{\succ}]$ has only two stable matchings:

$$\begin{aligned} \text{DA}_M[\widetilde{\succ}] &= \{(m_1, w_1), (m_2, w_2), (m_3, w_3), \ldots, (m_r, w_r), w_{r+1}, \ldots, w_s\}, \\ \text{DA}_W[\widetilde{\succ}] &= \{(m_1, w_2), (m_2, w_1), (m_3, w_3), \ldots, (m_r, w_r), w_{r+1}, \ldots, w_s\}. \end{aligned}$$

Define $\overline{M} = M \setminus \{m_1, m_2\}$ and $K = \left\{\succ \in \mathcal{Q} : (\succ_i)_{i \in \overline{M} \cup W} = (\widetilde{\succ}_i)_{i \in \overline{M} \cup W}\right\}$. Let $\Omega : \mathcal{Q} \to \mathcal{M}$ be the mechanism such that

$$\Omega[\succ] = \begin{cases} \text{DA}_W[\succ], & \text{when } \succ \in K, \\ \text{DA}_M[\succ], & \text{when } \succ \notin K. \end{cases}$$

The mechanism Ω is different from DA_M, because $\Omega[\widetilde{\succ}] = \text{DA}_W[\widetilde{\succ}] \neq \text{DA}_M[\widetilde{\succ}]$.

We claim that Ω is strategy-proof for M. By contradiction, suppose that there are $m \in M$ and $\succ, \succ' \in \mathcal{Q}$ such that $\Omega[\succ'_m, \succ_{-m}](m) \succ_m \Omega[\succ](m)$. In this context, there are two relevant cases:

- Suppose that $\succ \in K$. Since each agent in $\overline{M}$ is matched with his best alternative in $\text{DA}_W[\succ]$, it follows that $m \in \{m_1, m_2\}$. Hence, $(\succ'_m, \succ_{-m}) \in K$. This implies that $(\succ_w)_{w \in W} = (\widetilde{\succ}_w)_{w \in W}$ and condition (4) ensures that the deferred acceptance algorithm DA_W finishes with the pairs that are formed at the first step when it is applied to either $(\succ'_m, \succ_{-m})$ or $\succ$. We conclude that $\Omega[\succ'_m, \succ_{-m}] = \text{DA}_W[\succ'_m, \succ_{-m}] = \text{DA}_W[\succ] = \Omega[\succ]$. A contradiction.
- Suppose that $\succ \notin K$. Since DA_M is strategy-proof for M in the preference domain $\mathcal{Q}$, we have that $(\succ'_m, \succ_{-m}) \in K$. Therefore, $(\succ_w)_{w \in W} = (\widetilde{\succ}_w)_{w \in W}$ and $m = m_i$ for some $i \in \{3, \ldots, r\}$, which imply that $w_i = \text{DA}_W[(\succ'_m, \succ_{-m})](m) = \Omega[(\succ'_m, \succ_{-m})](m) \succ_m \Omega[\succ](m) = \text{DA}_M[(\succ)](m)$. Therefore, although w_i considers m the best alternative under $\succ_{w_i}$, she rejects his proposal when deferred acceptance is applied to $\succ$. A contradiction.

We conclude that there is more than one stable mechanisms defined in $\mathcal{Q}$ that is strategy-proof for M.

(iii) Let $M = \{m_1, \ldots, m_r\}$, $W = \{w_1, \ldots, w_s\}$, and $\{m_1, m_2\} \subseteq M_\otimes$. Suppose that $r \geq s$ and consider a preference profile $\succ^* \in \mathcal{P}$ characterized by:

- Given $i \in \{1, 2\}$, m_i considers w_i the best potential partner.
- Given $i, j \in \{1, 2\}$ with $i \neq j$, w_i considers m_j her best potential partner.
- Given $i \in \{3, \ldots, s\}$, agents m_i and w_i consider each other the best alternative.

[14] This argument also contradicts the Rural Hospital Theorem, because $\overline{w} = \text{DA}_M[\succ](\overline{m})$ is left alone in $\Omega[\succ]$.

- The agents $m_1, \dots, m_s$ are the top alternatives for each $w \in W$.

It follows that the two-sided market $[M, W. \succ^*]$ has only two stable matchings:

$$\begin{aligned} \mathrm{DA}_M[\succ^*] &= \{(m_1,w_1),(m_2,w_2),(m_3,w_3),\dots,(m_s,w_s),m_{s+1},\dots,m_r\}, \\ \mathrm{DA}_W[\succ^*] &= \{(m_1,w_2),(m_2,w_1),(m_3,w_3),\dots,(m_s,w_s),m_{s+1},\dots,m_r\}. \end{aligned}$$

Let $\overline{M} = \{m_3, \dots, m_s\}$ and $K = \left\{\succ \in \mathcal{P} : (\succ_i)_{i \in \overline{M} \cup W} = (\succ_i^*)_{i \in \overline{M} \cup W}\right\}$.

Consider the mechanism $\Omega : \mathcal{P} \to \mathcal{M}$ defined by

$$\Omega[\succ] = \begin{cases} \mathrm{DA}_W[\succ], & \text{when } \succ \in K, \\ \mathrm{DA}_M[\succ], & \text{when } \succ \notin K. \end{cases}$$

Since $\Omega[\succ^*] = \mathrm{DA}_W[\succ^*] \neq \mathrm{DA}_M[\succ^*]$, the stable mechanisms Ω and DA_M are different. We claim that Ω is strategy-proof for M. By contradiction, suppose that there exists $m \in M$ and $\succ, \succ' \in \mathcal{P}$ such that $\Omega[\succ'_m, \succ_{-m}](m) \succ_m \Omega[\succ](m)$. There are two relevant cases to analyze:

- Suppose that $\succ \in K$. Since each agent in $\overline{M}$ is matched with his best alternative in $\mathrm{DA}_W[\succ]$, it follows that $m \notin \overline{M}$. Hence, $(\succ'_m, \succ_{-m}) \in K$. Moreover, as m_1 and m_2 have no outside options and $(\succ_w)_{w \in W} = (\succ^*_w)_{w \in W}$, DA_W finishes with the pairs formed at the first step when it is applied to either $(\succ'_m, \succ_{-m})$ or $\succ$. Thus, $\Omega[\succ'_m, \succ_{-m}] = \mathrm{DA}_W[\succ'_m, \succ_{-m}] = \mathrm{DA}_W[\succ] = \Omega[\succ]$, a contradiction.
- Suppose that $\succ \notin K$. Since the mechanism DA_M is strategy-proof for M, $(\succ'_m, \succ_{-m}) \in K$. Therefore, $(\succ_w)_{w \in W} = (\succ^*_w)_{w \in W}$ and $m = m_i$ for some $i \in \{3, \dots, s\}$, which guarantee that $w_i = \mathrm{DA}_W[(\succ'_m, \succ_{-m})](m) = \Omega[(\succ'_m, \succ_{-m})](m) \succ_m \Omega[\succ](m) = \mathrm{DA}_M[(\succ)](m)$. It follows that, although w_i considers m the best alternative under $\succ_{w_i}$, she rejects his proposal when deferred acceptance is applied to $\succ$. A contradiction.

(iv) The proof of Proposition 1 works to show this property (see footnote 10). □

The Theorem 1 is a direct consequence of properties (iii) and (iv) of Lemma 1, while Theorem 2 follows from properties (i), (ii), and (iii) of Lemma 1.

Proof of Theorem 3. Asume that $M = \{m_1, \dots, m_r\}$ and $W = \{w_1, \dots, w_s\}$, with $r \geq s \geq 3$.

(i) Suppose that the mechanism Φ is strategy-proof for M. By contradiction, asume that $m_1 \in M \setminus M_\otimes$. Let $\succ \in \mathcal{Q}'$ be a preference profile satisfying:

$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$	$\cdots$	$\succ_{m_s}$	$\succ_{m_{s+1}}$	$\cdots$	$\succ_{m_r}$	$\succ_{w_1}$	$\succ_{w_2}$	$\succ_{w_3}$	$\cdots$	$\succ_{w_s}$
w_2	w_1	w_3	$\cdots$	w_s	w_1	$\cdots$	w_1	m_1	m_2	m_3	$\cdots$	m_s
w_1	w_2	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	m_2	m_1	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

If $\succ'_{m_1}$ is a preference relation in which the only acceptable partner is w_2, then

$$\Phi[\succ'_{m_1}, \succ_{-m_1}](m_1) = w_2 \succ_{m_1} w_1 = \Phi[\succ](m_1),$$

which contradicts the strategy-proofness of Φ.

When $M_\otimes = M$, the same arguments of the proof of Proposition 1 can be applied to show that Φ is strategy-proof for M. In fact, since the absence of outside options for agents in W plays no role in this proof, Φ coincides with the mechanism $\Omega_{\mathcal{F}}$ defined in the proof of Proposition 1 when we consider $\mathcal{F} = \mathcal{F}^*$ and $\mathcal{P}_{\mathcal{F}}$ is changed by $\mathcal{Q}' \equiv \{\succ \in \mathcal{Q}(M_\otimes, W_\otimes) : \exists f \in \mathcal{F},\ b_w(\succ) = f(w),\ \forall w \in W\}$.

(ii) The proof of Proposition 2 works to show this property (it is sufficient to change Ψ to Φ and $\mathcal{P}'$ to $\mathcal{Q}'$).

(iii) For each $C \subseteq M$ and $\succ, \widetilde{\succ} \in \mathcal{Q} \equiv \mathcal{Q}(M_\otimes, W_\otimes)$, we have that $\succ \in \mathcal{Q} \setminus \mathcal{Q}'$ if and only if $(\widetilde{\succ}_C, \succ_{-C}) \in \mathcal{Q} \setminus \mathcal{Q}'$. As a consequence, to show that Φ is not group strategy-proof for M, it is sufficient to find a preference profile in $\mathcal{Q} \setminus \mathcal{Q}'$ at which DA_M is manipulable by a coalition in M.

Let $\succ^* \in \mathcal{Q} \setminus \mathcal{Q}'$ be a preference profile satisfying

$\succ^*_{m_1}$	$\succ^*_{m_2}$	$\succ^*_{m_3}$	$\succ^*_{m_4}$	$\cdots$	$\succ^*_{m_s}$	$\succ^*_{m_{s+1}}$	$\cdots$	$\succ^*_{m_r}$	$\succ^*_{w_1}$	$\succ^*_{w_2}$	$\succ^*_{w_3}$	$\cdots$	$\succ^*_{w_{s-1}}$	$\succ^*_{w_s}$
w_2	w_1	w_1	w_4	$\cdots$	w_s	w_1	$\cdots$	w_1	m_1	m_2	m_3	$\cdots$	m_{s-1}	m_1
w_1	w_2	w_3	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	m_3	m_1	$\vdots$	$\vdots$	$\vdots$	m_s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	m_2	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

It follows that $\mathrm{DA}_M[\succ^*] = \{(m_1, w_1), (m_2, w_2), (m_3, w_3), \ldots, (m_s, w_s), m_{s+1}, \ldots, m_r\}$. Consider the coalition $C = \{m_1, m_2, m_3\}$ and $\widetilde{\succ}^* \in \mathcal{Q} \setminus \mathcal{Q}'$ that coincides with $\succ^*$ except in the preference relation of agent m_3 which is characterized by $w_3 \,\widetilde{\succ}^*_{m_3}\, w_1 \widetilde{\succ}^*_{m_3} \cdots$. Since the matching $\mathrm{DA}_M[\widetilde{\succ}^*_C, \succ^*_{-C}]$ is given by $\{(m_1, w_2), (m_2, w_1), (m_3, w_3), \ldots, (m_s, w_s), m_{s+1}, \ldots, m_r\}$, the coalition C manipulates DA_M at $\succ^*$. We conclude that Φ is not group strategy-proof for M. □

Proof of Theorem 4. Given $\Omega_1, \Omega_2 \in \mathbb{S}_M(\mathcal{Q})$, let $\vee$ and $\wedge$ be the binary operations such that, for each agent $m \in M$ and preference profile $\succ \in \mathcal{Q}$, we have that

$$\Omega_1 \vee \Omega_2[\succ](m) = \max_{\succ_m}\{\Omega_1[\succ](m),\, \Omega_2[\succ](m)\}, \qquad \Omega_1 \wedge \Omega_2[\succ](m) = \min_{\succ_m}\{\Omega_1[\succ](m),\, \Omega_2[\succ](m)\},$$

where $\max_{\succ_m}\{w, w'\} = w$ if and only if either $w \succ_m w'$ or $w = w'$, and $\min_{\succ_m}\{w, w'\} = w$ if and only if either $w' \succ_m w$ or $w = w'$. Since the stable matchings of $[M, W, \succ]$ form a lattice under the binary operations $\vee$ and $\wedge$ (Knuth, 1976; Theorem 7, attributed to J. H. Conway), the mechanisms $\Omega_1 \vee \Omega_2$ and $\Omega_1 \wedge \Omega_2$ are well-defined and stable. We claim that $\Omega_1 \vee \Omega_2$ and $\Omega_1 \wedge \Omega_2$ belong to $\mathbb{S}_M(\mathcal{Q})$. Indeed, by contradiction, suppose that there exist $m \in M$ and $\succ, \succ' \in \mathcal{Q}$ such that $\Omega_1 \vee \Omega_2[\succ'_m, \succ_{-m}](m) \succ_m \Omega_1 \vee \Omega_2[\succ](m)$. It follows from the definition of $\vee$ that $\Omega_1 \vee \Omega_2[\succ'_m, \succ_{-m}](m) = \Omega_i[\succ'_m, \succ_{-m}](m)$ for some $i \in \{1, 2\}$, which implies that $\Omega_i[\succ'_m, \succ_{-m}](m) \succ_m \Omega_i[\succ](m)$. This contradicts the fact that $\Omega_i \in \mathbb{S}_M(\mathcal{Q})$. Analogously, if $\Omega_1 \wedge \Omega_2[\succ'_m, \succ_{-m}](m) \succ_m \Omega_1 \wedge \Omega_2[\succ](m)$, then $\Omega_1 \wedge \Omega_2[\succ](m) = \Omega_j[\succ](m)$ for some $j \in \{1, 2\}$. Hence, $\Omega_j[\succ'_m, \succ_{-m}](m) \succ_m \Omega_j[\succ](m)$, which contradicts the fact that $\Omega_j \in \mathbb{S}_M(\mathcal{Q})$.

Therefore, $(\mathbb{S}_M(\mathcal{Q}), \vee, \wedge)$ is a lattice. Notice that, $\Omega_1 \geqslant_M \Omega_2$ if and only if $\Omega_1 = \Omega_1 \vee \Omega_2$. Since agents from different sides of the market have opposed preferences for stable matchings (Knuth, 1976; Corollary 1), it follows that $\Omega_1 \geqslant_W \Omega_2$ if and only if $\Omega_2 \geqslant_M \Omega_1$. This relationship implies that a mechanism is W-optimal in $\mathbb{S}_M(\mathcal{Q})$ if and only if it is the least element of $\mathbb{S}_M(\mathcal{Q})$ under the partial order $\geqslant_M$. Since every finite lattice has a least element, we conclude that there is $\Omega_W : \mathcal{Q} \to \mathcal{M}$ that is W-optimal in $\mathbb{S}_M(\mathcal{Q})$.

(i) By contradiction, suppose that there exists $m \in M \setminus M_\otimes$ such that $\Omega[\succ](m) \neq \mathrm{DA}_M[\succ](m)$ for some preference profile $\succ \in \mathcal{Q}$ and mechanism $\Omega \in \mathbb{S}_M(\mathcal{Q})$. Since $\mathrm{DA}_M[\succ]$ is the best stable matching of $[M, W, \succ]$ for agents in M, $w \equiv \mathrm{DA}_M[\succ](m) \succ_m \Omega[\succ](m)$. Let $\succ'_m$ be a preference relation defined on $W \cup \{m\}$ such that $w \succ'_m m \succ'_m \cdots$ (remember that m may declare others unacceptable). Since $\mathrm{DA}_M[\succ]$ is stable under $(\succ'_m, \succ_{-m})$ and w is the only acceptable partner of the agent m under $\succ'_m$, the Rural Hospital Theorem implies that $\Omega(\succ'_m, \succ_{-m})(m) = w$. Hence, $\Omega(\succ'_m, \succ_{-m})(m) \succ_m \Omega[\succ](m)$, which contradicts the strategy-proofness of Ω.

(ii) Suppose that agents have no outside options (i.e., $M_\otimes = M$ and $W_\otimes = W$). In this case, the preference domain $\mathcal{Q} \equiv \mathcal{Q}(M_\otimes, W_\otimes)$ coincides with $\mathcal{P}$.

- For any $\succ \in \mathcal{P}$, $\text{DA}_M[\succ]$ is the best stable matching for agents in M. Hence, $\Omega >_W \text{DA}_M$ for all $\Omega \in \mathbb{S}_M(\mathcal{P}) \setminus \{\text{DA}_M\}$, which implies that $\Omega_W \neq \text{DA}_M$ if and only if $|\mathbb{S}_M(\mathcal{P})| > 1$. Therefore, it follows form Proposition 1 that $\Omega_W \neq \text{DA}_M$ if and only if $|M| \geq |W|$.
- We claim that no agent can manipulate the mechanism $\text{DA}_W : \mathcal{P} \to \mathcal{M}$ when $|W| = 2$. Indeed, Dubins and Freedman (1981, Theorem 9) and Roth (1982, Theorem 5) ensure that DA_W is strategy-proof for W. Moreover, given $m \in M$ and $\succ \in \mathcal{P}$ such that $w_i \succ_m w_j$, we have that:
  - If $\text{DA}_W[\succ](m) = w_i$, then m has no incentives to manipulate DA_W at $\succ$.
  - If $\text{DA}_W[\succ](m) = w_j$, then m does not receive a proposal from w_i when DA_W is implemented. Moreover, $\text{DA}_W[\succ'_m, \succ_{-m}](m) = \text{DA}_W[\succ](m)$ when $\succ'_m$ is characterized by $w_j \succ'_m w_i$. Since the mechanism designer knows that m does not have an outside option, he cannot improve his situation by misrepresenting preferences.
  - If $\text{DA}_W[\succ](m) = m$, then m does not receive any proposal when DA_W is implemented. This will remain the case regardless of the preferences that he decides to report. Therefore, m cannot improve his situation by misrepresenting preferences.

  Hence, DA_W is stable and strategy-proof in the preference domain $\mathcal{P}$ when $|W| = 2$. Since DA_W implements the best stable matching for agents in W, it follows that $\Omega_W = \text{DA}_W$ when $|W| = 2$. Equivalently, $\Omega_W \neq \text{DA}_W$ implies that $|W| > 2$.

  We claim that DA_W is not strategy-proof for M when $|W| > 2$. Indeed, let $\succ \in \mathcal{P}$ be a preference profile such that, for some $m_1, m_2 \in M$ and $w_1, w_2, w_3 \in W$ we have that:

$$w_1 \succ_{m_1} w_2 \succ_{m_1} w_3 \succ_{m_1} \cdots, \qquad w_2 \succ_{m_2} w_1 \succ_{m_2} w_3 \succ_{m_2} \cdots,$$

$$m_2 \succ_{w_1} m_1 \succ_{w_1} \cdots, \qquad m_1 \succ_{w_2} m_2 \succ_{w_2} \cdots, \qquad m_1 \succ_{w_3} m_2 \succ_{w_3} \cdots.$$

  Then, $\text{DA}_W[\succ](m_1) = w_2$. However, when m_1 reports preferences $w_1 \succ'_{m_1} w_3 \succ'_{m_1} w_2 \succ'_{m_1} \cdots$, we have that $\text{DA}_W[\succ'_{m_1}, \succ_{-m_1}](m_1) = w_1$. Thus, DA_W is not strategy-proof for M.

(iii) Suppose that $|M| \geq |W| \geq 3$. The proof of Theorem 3 in Roth (1982) ensures that $\mathbb{S}_M(\mathcal{Q}) \cap \mathbb{S}_W(\mathcal{Q}) = \emptyset$. Hence, Ω_W differs from DA_W. Since $|M| \geq |W|$, it follows from properties (i) and (iii) of Lemma 1 that Ω_W differs from DA_M if and only if $M_\otimes \geq 2$.

(iv) Suppose that $M = \{m_1, \dots, m_r\}$, $\{m_3, m_4\} \subseteq M_\otimes$, and $W = \{w_1, \dots, w_s\}$, where $\min\{r, s\} \geq 5$.

Case I: $|M| \geq |W|$.

Since $r \geq s \geq 5$, let $\succ^* \in \mathcal{Q} \equiv \mathcal{Q}(M_\otimes, W_\otimes)$ be any preference profile satisfying:

$\succ^*_{m_1}$	$\succ^*_{m_2}$	$\succ^*_{m_3}$	$\succ^*_{m_4}$	$\succ^*_{m_5}$	$\cdots$	$\succ^*_{m_s}$	$\succ^*_{w_1}$	$\succ^*_{w_2}$	$\succ^*_{w_3}$	$\succ^*_{w_4}$	$\succ^*_{w_5}$	$\cdots$	$\succ^*_{w_s}$
w_2	w_1	w_4	w_3	w_5	$\cdots$	w_s	m_1	m_2	m_3	m_4	m_1	$\cdots$	m_1
w_1	w_2	w_3	w_4	$\vdots$	$\vdots$	$\vdots$	m_2	m_1	m_4	m_3	m_5	$\cdots$	m_s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

We want to prove that $\Omega_W[\succ^*] \notin \{\text{DA}_M[\succ^*], \text{DA}_W[\succ^*]\}$.

It is not difficult to verify that the market $[M, W, \succ^*]$ has four stable matchings:

$$\begin{aligned}
\mathrm{DA}_M[\succ^*] &= \{(m_1,w_2),(m_2,w_1),(m_3,w_4),(m_4,w_3),(m_5,w_5)\dots,(m_s,w_s),m_{s+1},\dots,m_r\},\\
\mathrm{DA}_W[\succ^*] &= \{(m_1,w_1),(m_2,w_2),(m_3,w_3),(m_4,w_4),(m_5,w_5)\dots,(m_s,w_s),m_{s+1},\dots,m_r\},\\
\mu &= \{(m_1,w_1),(m_2,w_2),(m_3,w_4),(m_4,w_3),(m_5,w_5)\dots,(m_s,w_s),m_{s+1},\dots,m_r\},\\
\eta &= \{(m_1,w_2),(m_2,w_1),(m_3,w_3),(m_4,w_4),(m_5,w_5)\dots,(m_s,w_s),m_{s+1},\dots,m_r\}.
\end{aligned}$$

Let $M^* = \{m_1, m_2\} \cup \{m_5, \dots, m_s\}$ and $K^* = \{\succ \in \mathcal{Q} : (\succ_i)_{i\in M^*\cup W} = (\succ^*_i)_{i\in M^*\cup W}\}$.

Consider the mechanism $\Omega : \mathcal{Q} \to \mathcal{M}$ characterized by

$$\Omega[\succ] = \begin{cases} \eta, & \text{when } \succ \in K^*,\\ \mathrm{DA}_M[\succ], & \text{when } \succ \notin K^*. \end{cases}$$

Since $\{m_3, m_4\} \subseteq M_\otimes$, for any $\succ \in K^*$ the matching η is stable in $[M, W, \succ]$. As a consequence, the mechanism Ω is stable. We claim that Ω is strategy-proof for M. By contradiction, assume that there exists $m \in M$ and $\succ, \succ' \in \mathcal{Q}$ such that $\Omega[\succ'_m, \succ_{-m}](m) \succ_m \Omega[\succ](m)$. There are two relevant cases to analyze:

- Suppose that $\succ \in K^*$. Since η pairs each agent in M^* with his best alternative, $m \notin M^*$. Hence, $(\succ'_m, \succ_{-m}) \in K^*$ and $\Omega[\succ'_m, \succ_{-m}](m) = \eta(m) = \Omega[\succ](m)$. A contradiction.
- Suppose that $\succ \notin K^*$. Since DA_M is strategy-proof for M in the preference domain $\mathcal{Q}$, it follows that $(\succ'_m, \succ_{-m}) \in K^*$. Hence, $\eta(m) \succ_m \mathrm{DA}_M[\succ](m)$. It follows from Gale and Sotomayor (1985, Theorem 4) that η is unstable in $[M, W, (\succ'_m, \succ_{-m})]$. A contradiction.

The fact that Ω belongs to $\mathbb{S}_M(\mathcal{Q})$ implies that $\Omega_W[\succ^*](w) \succeq^*_w \Omega[\succ^*](w)$ for all $w \in W$, which guarantees that $\Omega_W[\succ^*] \in \{\mathrm{DA}_W[\succ^*], \eta\}$. However, if $\Omega_W[\succ^*] = \mathrm{DA}_W[\succ^*]$, then m_1 can manipulate Ω_W. Indeed, if $\succ_{m_1}$ satisfies $w_2 \succ_{m_1} w_5 \succ_{m_1} w_1 \succ_{m_1} \cdots$, then in any stable matching of $[M, W, (\succ_{m_1}, \succ^*_{-m_1})]$ the agent m_1 forms a pair with w_2. Hence, $\Omega_W[\succ_{m_1}, \succ^*_{-m_1}](m_1) = w_2 \succ^*_{m_1} w_1 = \Omega_W[\succ^*](m_1)$, which contradicts the strategy proofness of Ω_W. Therefore, $\Omega_W[\succ^*] \notin \{\mathrm{DA}_M[\succ^*], \mathrm{DA}_W[\succ^*]\}$.

Case II: $|M| < |W|$ and $W_\otimes = \emptyset$.

Since $s > r \geq 5$, let $\succ^\circ \in \mathcal{Q}$ be any preference profile satisfying:

$\succ^\circ_{m_1}$	$\succ^\circ_{m_2}$	$\succ^\circ_{m_3}$	$\succ^\circ_{m_4}$	$\succ^\circ_{m_5}$	$\cdots$	$\succ^\circ_{m_r}$	$\succ^\circ_{w_1}$	$\succ^\circ_{w_2}$	$\succ^\circ_{w_3}$	$\succ^\circ_{w_4}$	$\succ^\circ_{w_5}$	$\cdots$	$\succ^\circ_{w_r}$
w_2	w_1	w_4	w_3	w_5	$\cdots$	w_s	m_1	m_2	m_3	m_4	m_1	$\cdots$	m_1
w_1	w_2	w_3	w_4	$\vdots$	$\vdots$	$\vdots$	m_2	m_1	m_4	m_3	m_5	$\cdots$	m_s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Moreover, assume that agents in $\{w_{r+1}, \dots, w_s\}$ consider m_3 and m_4 inadmissible under $\succ^\circ$.

It is not difficult to verify that the stable matchings of $[M, W, \succ^\circ]$ are given by:

$$\begin{aligned}
\mathrm{DA}_M[\succ^\circ] &= \{(m_1,w_2),(m_2,w_1),(m_3,w_4),(m_4,w_3),(m_5,w_5)\dots,(m_r,w_r),w_{r+1},\dots,w_s\},\\
\mathrm{DA}_W[\succ^\circ] &= \{(m_1,w_1),(m_2,w_2),(m_3,w_3),(m_4,w_4),(m_5,w_5)\dots,(m_r,w_r),w_{r+1},\dots,w_s\},\\
\mu &= \{(m_1,w_1),(m_2,w_2),(m_3,w_4),(m_4,w_3),(m_5,w_5)\dots,(m_r,w_r),w_{r+1},\dots,w_s\},\\
\eta &= \{(m_1,w_2),(m_2,w_1),(m_3,w_3),(m_4,w_4),(m_5,w_5)\dots,(m_r,w_r),w_{r+1},\dots,w_s\}.
\end{aligned}$$

Let $M^\circ = \{m_1, m_2\} \cup \{m_5, \dots, m_r\}$ and $K^\circ = \{\succ \in \mathcal{Q} : (\succ_i)_{i\in M^\circ\cup W} = (\succ^*_i)_{i\in M^\circ\cup W}\}$.

Consider the mechanism $\Omega : \mathcal{Q} \to \mathcal{M}$ such that

$$\Omega[\succ] = \begin{cases} \eta, & \text{when } \succ \in K^\circ,\\ \mathrm{DA}_M[\succ], & \text{when } \succ \notin K^\circ. \end{cases}$$

Since $\{m_3, m_4\} \subseteq M_\otimes$ are inadmissible for $\{w_{r+1}, \dots, w_s\}$, η is stable in $[M, W, \succ]$ for any $\succ \in K^\circ$. Hence, identical arguments to those made in the proof of Case I ensure that $\Omega \in \mathbb{S}_M(\mathcal{Q})$, which implies that $\Omega_W[\succ^\circ] \in \{\text{DA}_W[\succ^\circ], \eta\}$. Furthermore, the same ideas applied in the proof of Case I allow us to show that agent m_1 manipulates Ω_W when $\Omega_W[\succ^\circ] = \text{DA}_W[\succ^\circ]$. Therefore, $\Omega_W[\succ^\circ] \notin \{\text{DA}_M[\succ^\circ], \text{DA}_W[\succ^\circ]\}$. □

References

[1] Abdulkadiroğlu, A., P. Pathak, and A. Roth (2009): "Strategy-proofness versus efficiency in matching with indifferences: redesigning the NYC high school match," *American Economic Review*, 99, 1954-1978.

[2] Akbarpour, M., A. Kapor, C. Neilson, W. Van Dijk, and S. Zimmerman (2022): "Centralized school choice with unequal outside options," *Journal of Public Economics*, 210, 104644.

[3] Alcalde, J., and S. Barberà (1994): "Top dominance and the possibility of strategy-proof stable solutions to matching problems," *Economic Theory*, 4, 417-435.

[4] Ashlagi, I., Y. Kanoria, and J. Leshno (2017): "Unbalanced random matching markets: the stark effect of competition," *Journal of Political Economy*, 125, 69-98.

[5] Dubins, L., and D.A. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *The American Mathematical Monthly*, 88, 485-494.

[6] Echenique, F., N. Immorlica, and V. Vazirani (2023): "Online and matching-based market design," Cambridge University Press.

[7] Erdil, A. (2014): "Strategy-proof stochastic assignment," *Journal of Economic Theory*, 151, 146-162.

[8] Gale, D., and L. Shapley (1962): "College admissions and the stability of marriage," *The American Mathematical Monthly,* 69, 9-15.

[9] Gale, D., and M. Sotomayor (1985): "Some remarks on the stable matching problem," *Discrete Applied Mathematics*, 11, 223-232.

[10] Hassidim, A., D. Marciano, A. Romm, and R. Shorrer (2017): "The mechanism Is truthful, why aren't you?," *American Economic Review: Papers & Proceedings*, 107, 220-224.

[11] Hatfield, J., and P. Milgrom (2005): "Matching with contracts," *American Economic Review*, 95, 913-935.

[12] Hirata, D., and Y. Kasuya (2017): "On stable and strategy-proof rules in matching markets with contracts," *Journal of Economic Theory*, 168, 27-43.

[13] Kesten, O. (2010): "School choice with consent," *The Quarterly Journal of Economics,* 125, 1297-1348.

[14] Kesten, O., and M. Kurino (2019): "Strategy-proof improvements upon deferred acceptance: A maximal domain for possibility," *Games and Economic Behavior,* 117, 120-143.

[15] Knuth, D. (1976): "Mariages stables et leurs relations avec d'autres problèmes combinatoires," Les Presses de l'Université de Montréal, Montréal

[16] Roth, A. (1982): "The economics of matching: stability and incentives," *Mathematics of Operations Research*, 7, 617-628.

[17] Sakai, T. (2011): "A note on strategy-proofness from the doctor side in matching with contracts," *Review of Economic Design*, 15, 337-342.

[18] Svensson, L-G. (1999): "Strategy-proof allocation of indivisible goods," *Social Choice and Welfare*, 16, 557-567

[19] Teo, C., J. Sethuraman, and W. Tan (2001): "Gale-Shapley stable marriage problem revisited: strategic issues and applications," *Management Science*, 47, 1252-1267.