Reallocation with Priorities*

Julien Combe[†] Jan Christoph Schlegel[‡]

August 22, 2023

Abstract

We consider a reallocation problem with priorities where each agent is initially endowed with a house and is willing to exchange it but each house has a priority ordering over the agents of the market. In this setting, it is well known that there is no individually rational and stable mechanism. As a result, the literature has introduced a modified stability notion called μ_0-stability. In contrast to college admission problems, in which priorities are present but there is no initial endowment, we show that the ownership-adjusted Deferred Acceptance mechanism identified in the literature is not the only individually rational, strategy-proof and μ_0-stable mechanism. By introducing a new axiom called the independence of irrelevant agents and using the standard axiom of unanimity, we show that the ownership-adjusted Deferred Acceptance mechanism is the unique mechanism that is individually rational, strategy-proof, μ_0-stable, unanimous and independent of irrelevant agents.

JEL Classification: C78, D47.
Keywords: Matching, Housing Market, Reallocation, Stability, Priorities.

*We thank Umut Dur, Bettina Klaus, Yukio Koriyama, Thayer Morill, Matias Nunez, Olivier Tercieux, seminar participants at North Carolina State University, participants of the 6th Game Theory World Congress and the Dagstuhl Workshop on "Matching under Preferences" for valuable comments. Combe is grateful to Vasiliki Skreta for the support from the ERC Grant No. 682417, Frontiers in Design - Frontiers in Mechanism Design: Methodology and Applications and the grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" Program (LabEx ECODEC ; ANR-11-LABX-0047).

[†]CREST - Ecole polytechnique, France Email: julien.combe@polytechnique.edu

[‡]City, University of London, United Kingdom. Email: jansc@alumni.ethz.ch

1 Introduction

In many applied matching problems, indivisible resources have to be reallocated. In theory, agents are initially endowed with an indivisible object (following the standard terminology in the literature, we call these objects *houses*), monetary transfers are not allowed, and we use agents' preferences for houses and the initial allocation to determine a new allocation. In practice, we often encounter a situation where priorities are defined to discriminate between agents. Example applications include campus housing (Guillen and Kesten, 2012), reassignment of workers to positions (Compte and Jehiel, 2008; Dur and Ünver, 2019), teacher assignment (Pereyra, 2013; Combe, Tercieux, and Terrier, 2016) or school choice with a default option (e.g., a neighborhood school). The problem can also occur if some or all of the houses are initially unallocated and an initial allocation of the unallocated houses is generated by a lottery (Sönmez and Ünver, 2005) or as a second stage of an assignment procedure, where an initial allocation is generated by a matching mechanism.

Ideally, a good reallocation mechanism should satisfy a combination of desirable properties: a minimal requirement for any such mechanism should be *individual rationality* (IR); i.e., each agent should be weakly better off after reallocation. Additionally, the designer would like to achieve incentive compatibility in the sense of *strategy-proofness* (SP), efficiency and some form of fairness. The Top Trading Cycle (TTC) mechanism defined by Shapley and Scarf, 1974 and attributed to David Gale is an IR, SP and Pareto efficient (PE) mechanism, and in fact the only such mechanism (Ma, 1994). Given indivisibilities and the absence of monetary transfers, fairness is generally harder to achieve. For example, minimal fairness requirements such as the equal treatment of equals or envy-freeness will be violated by any reallocation mechanism. However, such solutions completely disregard the priority rankings of the houses. Reallocation problems with priorities can be seen as hybrids between the classical marriage problem (Gale and Shapley, 1962), where priorities exist but there is no initial allocation, and housing market problems (Shapley and Scarf, 1974), where there is an initial allocation but no priorities are defined. Then, fairness can be understood in the sense that there is no justified envy; i.e., no agent should prefer a house allocated to a lower priority agent to his allotment.

With initial endowments and priorities, it is well known that there is no matching that is both IR and stable in the sense that no agent has justified envy. To ensure the compatibility between the two notions of IR and stability, the concept of stability has been relaxed to exclude blocking pairs caused by a house that is assigned to its initial owner. With this relaxed notion, which is called μ_0-stability,[1] a simple variation of the Deferred Acceptance (DA) mechanism has been identified: it starts to simply rank the initial owners at the top of the priority ordering of their initial house and runs the standard DA mechanism over these modified priorities. This mechanism, which is called DA^*,[2] is IR, SP and μ_0-stable.

Our purpose is to give a normative justification for the use of the DA^* mechanism by providing an axiomatic characterization of it. In a model without initial endowments, the classic *Deferred Acceptance* (DA) mechanism is the unique IR (individual rationality is now understood in the sense that applicants obtain an assignment weakly preferred to being unmatched), SP and *stable* mechanism (see Alcalde and Barberà (1994), Theorem 3). However, the DA^* mechanism in the case of initial endowments is *not* the only mechanism that is IR, SP and μ_0-stable. For example, the trivial mechanism that assigns each agent his initial house is IR, SP and trivially μ_0-stable since blocking pairs are not considered when each agent is assigned his initial house. To rule out this trivial mechanism, one may require some limited form of efficiency such as *unanimity*: a mechanism should assign each agent his top choice whenever that is possible. We show that there are mechanisms other than DA^* that are IR, SP, μ_0-stable and unanimous. However, these mechanisms are problematic in the sense that the assignment can depend on "irrelevant" preference information. More specifically, an agent's change of ranking of a house can influence the allocation of the other houses, even though the allocation of the former house has not changed. We introduce a new axiom called *Independence of Irrelevant Agents* (IIAg) that rules out the possibility of irrelevant ranking information from an agent influencing the allocation. We discuss how this axiom is related to but different to the standard axiom of Irrelevance of Independent Alternatives (IIA) as well as the monotonicity and the non-bossiness axioms from the social choice literature. In particular, we introduce a natural priority adjusted non-bossiness axiom called *Minimal Bossiness* and show that it implies IIAg. Our main result shows that in a reallocation problem with priorities, DA^* is the unique IR, SP, μ_0-stable, unanimous and IIAg mechanism. This result also implies that DA^* is the unique, IR, SP, μ_0-stable, unanimous and MB mechanism.

[1] This terminology is borrowed from Compte and Jehiel (2008). Pereyra (2013) called such matchings *acceptable matchings* and Guillen and Kesten (2012) simply called them *fair matchings*.

[2] This terminology is borrowed from Combe, Tercieux, and Terrier (2016). Pereyra (2013) called this mechanism the *teacher proposing Deferred Acceptance* algorithm. We used the former terminology to highlight that this mechanism differs from the standard DA run over the primitive priorities since the latter is not IR.

Related literature. We build upon the classical housing market setting of Shapley and Scarf (1974) where all of the agents are initially assigned to houses and are willing to exchange them. Our characterization result can be trivially extended to the case where there are initially vacant houses and unassigned agents, which was introduced by Abdulkadiroglu and Sonmez (1999). We discuss this extension in the conclusion. Moreover, we add the feature that each house now has a priority ordering over the agents, which makes the model closer to the standard marriage market of Gale and Shapley (1962). The problem of allocating objects with priorities has been extensively studied in the literature starting with Balinski and Sonmez (1999) and extended later by the school choice literature (Abdulkadiroglu and Sonmez, 2003). The reallocation problem with priorities can be seen as a hybrid of the two extensions.

Guillen and Kesten (2012) were the first to notice that the NH4 mechanism used for off-campus housing reallocation at MIT is equivalent to DA^*. In their framework, houses have a common priority over agents. These authors performed an experiment to compare DA^* with the TTC mechanism and found that the participation under DA^* is significantly higher. Still, these authors' model differs from ours since we allow houses to have different priority orderings over agents. Our goal is also different, as we seek to provide a characterization of DA^*.

Pereyra (2013) also studied DA^* in the context of teacher assignments. Our model can be seen as a one-to-one version of his, i.e., where each school has only one initially assigned teacher. This author's focus is on the relaxation of the stability notion in the presence of IR. He defines a matching as acceptable if it is IR, and the only justified envies are the ones where a teacher prefers a school and has a higher priority than an initial teacher of that school who is assigned to it (we call these matchings μ_0-stable). This author's main result shows that an acceptable matching minimizes the remaining blocking pairs in the sense of inclusion if and only if it is the matching produced by the DA^* mechanism. This property can be seen as characterizing the DA^* mechanism. In the same vein, an alternative to our characterization is to require the mechanism to always return a μ_0-*optimal stable* matching, i.e., a μ_0-stable matching that every agent prefers to any other μ_0-stable matching. In the standard setting without endowment, it is well known that this property alone is enough to characterize DA, and it would be the same for DA^* in our setting. Our characterization provides another approach that uses standard axioms in the literature and allows us to clearly use the properties of IR, μ_0-stable and strategy-proof mechanisms. This update is important for two reasons. First, μ_0-optimal stability is not independent of the strategy-proofness axiom since it implies this axiom. For this reason, the characterization of Alcalde and Barberà (1994) does not rely on this stability and proves that with the weaker axiom of strategy-proofness, DA is still the only stable and SP mechanism. In our setting with an initial endowment and priorities, the set of IR, SP and μ_0-stable mechanisms is not a singleton anymore, hence, understanding the structure of such mechanisms is important. Our characterization and the related examples we provide help to reveal how one can build different mechanisms in this class. Second, in many applications, constrained efficiency is not the main policy objective. Policy makers may want to trade off the welfare of agents with other objectives, such as distributional objectives or the welfare of entities outside the model that are encoded into the priorities of the objects. For instance, this scenario occurs if one considers teachers' assignments, which are discussed in Combe, Tercieux, and Terrier (2016), or tuition exchanges as in Dur and Ünver (2019). Thus, a policymaker may be willing to only consider IR, SP and μ_0-stable mechanisms but select from among them a mechanism that respects other desiderata. Our results help to clearly identify which necessary properties to trade off when one picks different mechanisms in this class.

In the standard school choice setting, Balinski and Sonmez (1999) and Morrill (2013) provided a characterizations of DA whenever the priority profile is fixed. As we will discuss in the conclusion, our result can be interpreted as a characterization of DA in a setting without outside options and a restriction on the priority domains: agents do not have the outside option of being unmatched and the priority of houses are unanimous, i.e. there exists a matching assigning each house to its top priority agent. These two restrictions are motivated by the particular setting and applications of reassignment with priorities. Kesten and Kurino (2017) also considered the absence of outside options in a school choice framework to revisit the non-existence of strategy-proof mechanisms which Pareto-dominate the DA mechanism (Abdulkadiroğlu, Pathak, and Roth, 2009). Even though they show that the non-existence is still valid without outside option, they provide a maximal preference domain for the agents which ensures the existence of such mechanism. The relevance of outside options depends on the particular application one has in mind. In the context of teacher assignment, Combe, Tercieux, and Terrier (2016) report that newly recruited teachers have to rank all the schools since they are required to have an assignment at the end of the assignment procedure. Recently, Akbarpour, Kapor, Neilson, Van Dijk, and Zimmerman (2020) show that, in school choice, students always prefer manipulable over strategy-proof mechanisms if and only if they have an outside option and provide empirical evidences of such result. Priority restrictions have also been studied in the literature. For instance, Ergin (2002a) showed that, on a subdomain of priorities called "acyclic priorities", the DA mechanism always returns a Pareto-efficient matching. Up to our knowledge, we do not know any paper studying a joint restriction on both outside options and priorities.

In Section 2, we introduce the reallocation problem with priorities and the standard axioms of the literature. Then, in Section 3, we note that DA^* is not the only IR, μ_0-stable and SP mechanism motivating the introduction of the standard axiom of unanimity. We then provide an example to show that the DA^* mechanism is not the only IR, SP, μ_0-stable and unanimous mechanism which motivates the introduction of our new axiom of Independence of Irrelevant Agents. We finish the section by providing our main characterization result for DA^* and its proof. Section 4 discusses the relation our new axiom to the literature and provides additional results. Section 5 concludes and discusses extensions and further research.

2 Model and Definitions

A reallocation problem with priorities first starts with a standard housing market problem as proposed by Shapley and Scarf (1974). Let I be a finite set of **agents**, and let H be a finite set of **houses** such that $|I| = |H|$. Agents have strict **preferences** over houses that are modeled by a linear order over H.[3] We denote by $\mathcal{P}$ the set of all profiles of strict preferences $P = (P_i)_{i \in I}$ such that for each $i \in I$, P_i is a linear order over H. Following Guillen and Kesten (2012), the main departure from the standard housing market problem is that each house h has a strict **priority ordering** $\succ_h$ over agents, which is a linear order over I. We use standard notions: for a set of houses $H' \subset H$ and a preference profile P, we denote by $P|_{H'}$ the profile of linear orders over the subset H' implied by P. For a set of agents $I' \subset I$ and a preference profile P, we let $P_{I'}$ be the restriction of P to the agents in I'. For a preference profile P and an agent i, P_{-i} will be a shorthand notation for $P_{I \setminus \{i\}}$. [4]

[3] A linear order over H is a binary relation P_i that is *antisymmetric* (for each $h, h' \in H$ with $h \neq h'$ if $h \, P_i \, h'$ then we do not have $h \, P_i \, h'$), *transitive* (for each $h, h', h'' \in H$, if $h \, P_i \, h'$ and $h' \, P_i \, h''$, then $h \, P_i \, h''$), and *complete* (for each $h, h' \in H$ with $h \neq h'$, $h \, P_i \, h'$ or $h' \, P_i \, h$). We write $h \, R_i \, h'$ if $h \, P_i \, h'$ or $h = h'$. Hence, given $h, h' \in H$, $h \, P_i \, h'$ means that h is *strictly preferred* to h'; $h \, R_i \, h'$ means that h is *weakly preferred* to h' where indifference between different houses is not possible.

[4] Our model follows the standard housing market setting, by assuming that there are as many houses as agents (more generally, there could be more houses than agents) and being unassigned is not an option. The absence of an outside option is assumed in Shapley and Scarf (1974) and the following literature working on characterizations such as Ma (1994) or the more recent ones such as Pycia and Ünver (2017). This assumption also has an impact for the existing characterization results in allocation problems without initial ownership. For instance, under this assumption, Svensson (1999) shows that Serial Dictatorship mechanisms are the only Pareto-efficient, strategy-

A **matching** μ is a bijection between I and H so that $\mu(i)$ is the house assigned to agent i. We will also interchangeably use μ_i and $\mu(i)$. We abuse notations and denote by $\mu(h)$ the agent assigned to house h. Denote the set of all matchings by $\mathcal{M}$. We assume that there is an **initial matching** $\mu_0 \in \mathcal{M}$ that we may want to improve upon through reallocation. A (reallocation) **mechanism** φ assigns a matching to each preference profile, i.e., it is a mapping $\varphi : \mathcal{P} \to \mathcal{M}$. $\varphi_i(P)$ (resp. $\varphi_h(P)$) denotes the house (resp. the agent) assigned to agent i (resp. house h). We are interested in designing mechanisms that have certain desirable properties. In the context of reallocation, the existing rights should be respected by ensuring that every agent is as least as well off as under their initial assignment. Formally,

Axiom (Individual Rationality). *A mechanism φ is individually rational (IR) with respect to the initial matching μ_0 if for each $P \in \mathcal{P}$ and $i \in I$ we have*

$$\varphi_i(P)\, R_i\, \mu_0(i).$$ [5]

For a matching μ, a pair $(i, h) \in I \times H$ is called a blocking pair of μ if $hP_i\mu_i$ and $i \succ_h \mu(h)$. A matching is $\succ$-stable if it does not have any blocking pair under the priority profile $\succ$.[6] A matching is $\succ$-optimal stable if it is the most preferred $\succ$-stable matching for individuals. Individual rationality can in general be in conflict with the priorities so that there could be no matching that is both IR and stable. However, we can require a relaxed notion of stability:

Axiom (μ_0-Stability). *A matching μ is μ_0-stable with respect to preferences R and priorities $\succ$ if for each $i \in I$ and $h \in H$ we have the following: if $h\, P_i\, \mu(i)$ and*

proof and neutral mechanisms. Recently, Pycia and Unver (2021) have provided the full class of such mechanisms when the outside options are introduced.

In addition, we impose that each agent is initially assigned a house. In the conclusion, we discuss the implications of relaxing each of the two restrictions, the absence of an outside option and the assumptions of an initial assignment for each agent.

[5] In the literature of allocation with priorities but without initial assignment, a matching is IR if each agent prefers his match to the outside option, usually denoted by $\emptyset$. As discussed in Footnote 4, we follow the housing market literature by assuming that such outside option is not available. Here, the initial house serves as outside option. This IR definition is the one used in this literature, see for instance Ma (1995).

[6] Implicitly, stability also depends on the preference profile P but we omit it from the notations since it will be clear from the context. Without an initial assignment, the standard definition of stability also includes *non-wastefulness* which imposes that if an agent prefers a house to its assigned one, then this house must be occupied. In our housing market context, because we do not have any outside option, all matchings trivially respect this condition so that we omit it in the definition.

$i \succ_h \mu(h)$, *then* $\mu_0(\mu(h)) = h$. *Mechanism* φ *is* μ_0*-stable if it assigns to each profile* $P \in \mathcal{P}$ *a* μ_0*-stable matching.*

In other words, μ_0-stability only allows blocking pairs if they are caused by an agent staying at his initial house. For a priority profile $\succ$, let $\succ^*$ be the priority profile where each initial agent is moved at the top of the priority ranking of his initial house. The following lemma clarifies the link between μ_0-stable matching and $\succ^*$-stable matching for IR matchings. This will be clear once we describe the Ownership-Adjusted Deferred Acceptance mechanism below.

Lemma 1. *A matching* μ *is IR and* μ_0*-stable if and only if it is* $\succ^*$*-stable.*

Proof. We provide the proof using our notations and framework but the result is equivalent to Lemma 1 in Pereyra (2013).

Assume that μ is not $\succ^*$-stable. Then there exists a blocking pair (i, h) where $hP_i\mu(i)$ and $i \succ_h^* \mu(h)$.[7] If $i = \mu_0(h)$, then μ would not be IR. If μ is IR, then $hP_i\mu(i)R_i\mu_0(i)$ so that $h \neq \mu_0(i)$. But by construction of $\succ_h^*$, if i is not the initial owner of h and $i \succ_h^* \mu(h)$, then agent $\mu(h)$ cannot be the initial owner of h either, i.e. $\mu(h) \neq \mu_0(h) \Leftrightarrow h \neq \mu_0(\mu(h))$ so that μ is not μ_0-stable.

Now assume that μ is IR but not μ_0-stable. Then there exists (i, h) s.t. $hP_i\mu(i)$, $i \succ_h \mu(h) := i'$ and $\mu_0(\mu(h)) \neq h$ so that $i' \neq \mu_0(h)$. By construction of $\succ^*$, the latter differs from $\succ$ only by moving initial owners from the priority of their initial house so that if $i \succ_h i'$ and $i' \neq \mu_0(h)$ then $i \succ_h^* i'$ so that μ is not $\succ^*$-stable.

Finally, assume that μ is not IR but is μ_0-stable, then there exists an agent i s.t. $h := \mu_0(i)P_i\mu(i)$ so that $\mu(h) \neq i = \mu_0(h)$. But by construction of $\succ_*$, we have that $i \succ_h^* i'$ for all $i' \neq i$, in particular $i \succ_i^* \mu(h)$ so that we conclude that μ is not $\succ^*$-stable. □

Because of Lemma 1, the set of μ_0-stable IR matchings has the same structure as the set of $\succ^*$-stable matchings. This observation will be used in the proof of our main characterization.

Additionally, we use the incentive compatibility property of strategy-proofness:

Axiom (Strategy-Proofness)**.** *A mechanism* φ *is strategy-proof if for agent* $i \in I$ *and profiles* P, P' *with* $P'_{-i} = P_{-i}$ *we have*

$$\varphi_i(P)\, R_i \varphi_i(P').$$

[7]Note that in our setting without outside option, a matching is a bijection between I and H so that a house is always assigned to an agent so $\mu(h) \in I$.

Deferred Acceptance is the unique IR, stable and strategy-proof mechanism in the model without initial endowments (Alcalde and Barberà, 1994). The mechanism can easily be adapted to respect initial ownership rights by treating owners as if they have top priority in their initial houses. Formally, the **ownership-adjusted Deferred Acceptance mechanism** for preferences P, priorities $\succ$ and initial matching μ_0 proceeds in rounds where in each round the following steps are performed.

1. Each agent i applies to his favorite house according to P_i that has not previously rejected i.

2. Each house h definitely accepts $\mu_0(h)$ if $\mu_0(h)$ has applied to it. Otherwise, the house tentatively accepts the highest priority agent according to $\succ_i$ among the agents that have applied to it and rejects all other applicants.

We omit the priorities in our notations since the context will always be clear, and we denote the final matching by $DA^*(P)$.

3 New axiom and main result

As in the setting without endowments, one might expect that DA* is the only IR, μ_0-stable and strategy-proof mechanism. However, there are many different individually rational, μ_0-stable and strategy-proof mechanisms. For example, one such mechanism is the trivial mechanism that assigns each agent to his initial house independently of the submitted preferences. This is an important departure from the case without endowment. This trivial mechanism is not very satisfactory and quite inefficient. To rule it out, one may require a basic efficiency property. Ideally, a mechanism is individually rational, strategy-proof, respecting of priorities (in the sense of μ_0-stability) and Pareto-efficient. Generally, these properties are incompatible (Ergin, 2002b). Thus, we have to content ourselves with a weaker notion of efficiency. For each $P \in \mathcal{P}$ and $i \in I$, we denote by $top(P_i)$ the highest ranked house according to P_i. We call a profile $P \in \mathcal{P}$ **unanimous** if for $i \neq j$ we have $\text{top}(P_i) \neq \text{top}(P_j)$. A mechanism is unanimous if it assign everyone their top house for unanimous profiles.

Axiom (Unanimity)**.** *A mechanism φ is unanimous if for each unanimous profile P we have $\varphi_i(P) = top(P_i)$.*

By definition of DA*, it is easy to see that it is an unanimous mechanism. One can also easily see that the trivial mechanism mentioned earlier is not unanimous.

But more surprisingly perhaps, there even exist IR, SP, μ_0-stable and unanimous mechanisms that differ from DA^*:

Example 1. *Consider five agents $I = \{a, b, c, d, e\}$, five houses $H = \{h_a, h_b, h_c, h_d, h_e\}$ and an initial matching μ_0 s.t. $\mu_0(k) = h_k$ for $k \in I$. Consider a priority relation $\succ$ such that we have the following:*

$$\begin{array}{llllll} \succ_{h_a}: & a & b & c & d & e \\ \succ_{h_b}: & b & a & c & d & e \\ \succ_{h_c}: & c & a & b & d & e \\ \succ_{h_d}: & d & a & b & c & e \\ \succ_{h_e}: & e & a & b & c & d \end{array}$$

We define φ as an IR, strategy-proof, μ_0-stable and unanimous mechanism that is not DA^* with priorities $\succ$ as follows.

Denote by $\mathcal{P}' \subseteq \mathcal{P}$ the set of preference profiles P such that

$$\begin{array}{lllllll} P_a: & h_b & P_a & h_a & P_a \ldots \\ P_b: & h_c & P_b & h_b & P_b \ldots \\ P_c: & h_b & P_c & h_a & P_c & h_c & P_c \ldots \end{array}$$

Define a matching μ by

$$\mu(a) = h_b, \quad \mu(b) = h_c, \quad \mu(c) = h_a, \quad \mu(d) = h_d, \quad \mu(e) = h_e.$$

$$\varphi(P) := \begin{cases} DA^*(P), & \text{if } P \notin \mathcal{P}', \\ \mu, & \text{if } P \in \mathcal{P}'. \end{cases}$$

The mechanism φ is unanimous since DA^* is unanimous, and profiles in $\mathcal{P}'$ are not unanimous since agents a and c both rank house h_b first at these profiles. Moreover, φ is μ_0-stable, since DA^* is μ_0-stable and μ is a μ_0-stable matching for each $P \in \mathcal{P}'$ (as d and e have lower priority than a, b and c at houses h_a, h_b and h_c). For strategy-proofness, note that by the strategy-proofness of DA^*, only a, b and c can possibly manipulate φ. However, note that for each profile $P \in \mathcal{P}'$ we have $\varphi_i(P) = DA^*_i(P)$ for $i \in \{a, b, c\}$. Thus, strategy-proofness follows from the strategy-proofness of DA^*. Finally, note that $\varphi \neq DA^*$. Indeed, we select a profile $P \in \mathcal{P}'$ such that

$$P_d: \quad h_e \quad P_d \quad h_d \quad P_d \ldots$$
$$P_e: \quad h_d \quad P_e \quad h_e \quad P_e \ldots$$

In this case, $DA^*(R)$ assigns h_e to d and h_d to e, whereas $\varphi(R)$ assigns h_d to d and h_e to e.

The mechanism φ also satisfies other desirable properties that have been discussed in the context of axiomatizations of the Deferred Acceptance mechanism. The mechanism is, for example, *weakly Maskin monotonic* in the sense of Kojima and Manea (2010), and it is weakly Pareto efficient. However, the mechanism has an important and less appealing feature: in the last preference profile that is considered in the example where $\varphi(P) \neq DA^*(P, \succ)$, the mechanism φ does not allow agents d and e to exchange their houses. However, if agent c reports the profile $P'_c : h_a \quad P'_c \quad h_c$, then φ allows d and e to exchange houses under (P'_c, P_{-c}). Therefore, at profile P, mechanism φ forbids the exchange between d and e because of the preference profile of c and his ranking of house h_b. The exchange is forbidden even though this house is not a part of the exchange between d and e, not even indirectly (as would be the case if, for example, h_b would be allocated to a different agent so that the original recipient of h_b could now be assigned h_d or h_e). Thus, the assignment switches based on the preference information of an agent that seems to be *irrelevant* for the assignment of h_d and h_e.

Based on the intuition of the mechanism in Example 1, let us define a new axiom. Fix an agent i, a house h and two profiles P_i and $\tilde{P}_i$. We say that profile $\tilde{P}_i$ **moves** h **in** P_i if the comparison of any pairs of houses which do not include h does not change. Formally, $\tilde{P}_i \neq P_i$ and $\tilde{P}_i|_{H\setminus\{h\}} = P_i|_{H\setminus\{h\}}$.

Axiom (Independence of Irrelevant Agents)**.** *A mechanism φ is Independent of Irrelevant Agents (IIAg) if $\forall i \in I$, $\forall h \in H$, and $P, \tilde{P} \in \mathcal{P}$ such that $\tilde{P}_{-i} = P_{-i}$, $\tilde{P}_i$ moves h in P_i and $\varphi_h(P) \neq i$, we have*

$$\varphi_h(\tilde{P}) = \varphi_h(P) \Rightarrow \varphi(\tilde{P}) = \varphi(P).$$

The axiom states that if an agent does not obtain a house and changes his report by moving that house in his preferences but this change is *irrelevant* for the allocation of that house, i.e., if it does not change the allocation of that house, then the whole allocation must remain the same.

Remark 1. *We briefly comment on the construction of IR SP and μ_0-stable mechanisms. As mentioned, the constant mechanism trivially respects the three axioms.*

Example 1 also suggests an alternative: one can select a subset of agents (d and e in Example 1) and always match them to their initial assignment while running DA^* for the remaining agents. One can check that in Example 1, this would indeed generate an IR, SP and μ_0-stable mechanism. However, such procedure does not always work. Indeed, the set of agents being kept at their initial matching can still generate blocking pairs with other houses so that this procedure might generate non μ_0-stable matchings.[8] *This shows that the construction of IR, SP and μ_0-stable mechanisms heavily depends on the priority structure*[9] *which makes it challenging to give an explicit procedure and characterization of such mechanisms.*

We discuss the relevance of IIAg and its connection to the literature in Section 4 below.

Before stating our main theorem, we discuss the combination of axioms. We are interested in defining reasonable mechanisms that are IR, SP and μ_0-stable. As we have seen, one trivial solution is to consider the constant mechanism that always retain agents at their initial allocation. Other constant mechanisms can be constructed where only a subset of agents are kept at their initial allocation. However, to respect μ_0-stability, the choice of such subset depends on the priority profile.[10] Unanimity allows us to rule out such (semi) constant mechanisms. Intuitively, one way to construct a new IR, SP, μ_0-stable and unanimous mechanism is to force some group of agents to stay at their initial allocation under some preference profiles when they would otherwise move under DA^*, but to let them move under other preference profiles, typically the unanimous ones.[11] Intuitively, to maintain strategy-proofness for these agents, the decision of whether to hold them at their initial allocation cannot depend on their reported preferences. Thus, this decision must be taken by

[8]In Example 1, one can easily modify the priorities to make this procedure fail by ranking either d or e first in one of the houses h_a, h_b or h_c.

[9]One can easily show that the procedure mentioned above works only when the priority structure ranks the all the agents being kept at their initial matching below all other agents.

[10]For instance, in Example 1, the mechanism φ that always keep say agent a at his initial house h_a and runs DA^* for the other agents produces is IR and SP mechanism but not μ_0-stable. Indeed, one can easily construct a profile of preferences where say agent c is assigned to house h_b under φ and where agent a would prefer house h_b to h_a. In that case, agent a would form a blocking pair with h_b, violating μ_0-stability.

[11]Of course, one can also select another μ_0-stable matching for these agents instead of keeping them at their initial allocation. However, under some profile, strategy-proofness would force the same mechanism to hold some of these agents to their initial allocations. Indeed, if for some agent i and profile P, we have $DA_i^*(P)P_i\varphi_i(P)P_i\mu_0(i)$, then by reporting the profile P_i' that only ranks $DA_i^*(P)$ above $\mu_0(i)$, strategy-proofness would lead to $DA_i^*(P_i', P_{-i})P_i\varphi_i(P_i', P_{-i}) = \mu_0(i)$. In a way, our proof below will work with such a minimal example.

using the change in the preferences from another "irrelevant" agent, which is exactly illustrated by our Example 1.

Thus, one may wonder what mechanism is left once we rule out such group variations based on irrelevant agents. The answer to this question is exactly our main result.

Theorem 1. *A mechanism is IR, SP, μ_0-stable, unanimous and independent of irrelevant agents if and only if it is the DA^* mechanism.*

Proof. In the following, we fix a priority profile $\succ$ so that it is omitted from the notations. As mentioned before, we denote by $\succ^*$ the priority profile obtained from $\succ$ by moving the initial agent of each house to the top of its priority ranking.

First, it is standard to show that DA^* is IR, SP, μ_0-stable and unanimous.[12] We show that DA^* satisfies IIAg in Propositions 1 and 2 in Section 4 below.

For the other direction, let φ be an IR, SP, μ_0-stable, unanimous, and IIAg mechanism. Assume that $\varphi \neq DA^*$. In the following, for each profile P, we denote by

$$M(P) := \sum_{i \in I} |\{h : h \, P_i \, \mu_0(i)\}|$$

the number of houses ranked above the initial assignment. In addition, we denote by

$$N(P) := \sum_{i \in I} |\{h : h \, P_i \, DA_i^*(P)\}|$$

the number of houses ranked above the DA^* assignment at profile P. Let $\mathcal{Q} = \{P \in \mathcal{P} : \varphi(P) \neq DA^*(P)\}$. By assumption, because $\varphi \neq DA^*$, $\mathcal{Q}$ is non empty. Now, let $P \in \mathcal{Q}$ be a profile such that

1. for each $P' \in \mathcal{Q}$, we have $M(P) \leq M(P')$;

2. for each $P' \in \mathcal{Q}$ s.t. $M(P) = M(P')$, we have $N(P) \leq N(P')$.

Note that because $\mathcal{Q}$ is non-empty and finite, such profile P must exist. Let $\mu := DA^*(P)$ and $\nu := \varphi(P)$.

We prove the result through a sequence of claims. The first claim states that for the profile P, any agent who does not attain his DA^*-outcome under φ only ranks his DA^*-outcome above his initial assignment. This claim only requires the axioms of individual rationality, μ_0-stability and strategy-proofness.

[12]See Guillen and Kesten (2012), Pereyra (2013) or Compte and Jehiel (2008).

Claim 1. *For each $j \in N$ with $\mu(j) \neq \nu(j)$, only $\mu(j)$ is ranked above the initial match, i.e.,*

$$P_j : \mu(j)\, P_j\, \mu_0(j) = \nu(j) \ldots$$

Proof. Suppose otherwise and let $P'_j : \mu(j)\, P'_j\, \mu_0(j) \ldots$ and $P' := (P'_j, P_{-j})$. By the strategy-proofness of DA^*, we have $DA^*_j(P') = \mu(j)$. Since $DA^*(P)$ is a standard DA run over the priorities $\succ^*$, it returns the $\succ^*$-optimal stable matching, i.e., $DA^*(P)$ returns the most preferred $\succ^*$-stable matching for the agents which, by Lemma 1, is equivalent to the most preferred IR and μ_0-stable matching for the agents. In the following, we will simply refer to the matching in the previous argument as the μ_0-optimal stable matching. By μ^0-optimality of μ we have $\mu(j) R_j \nu(j)$. By the strategy-proofness of φ, we have $\varphi_j(P') = \mu_0(j) \neq \mu(j) = DA^*_j(P')$. However, this equation would imply $P' \in \mathcal{Q}$ and $M(P') < M(P)$, a contradiction with the assumption onthe choice of P. Thus, j only ranks $\mu(j)$ above his initial match. □

If $N(P) = 0$, then P is a unanimous profile, and by the unanimity of φ we have $\varphi(P) = DA^*(P)$. Thus, we may assume $N(P) > 0$. Since $N(P) > 0$, there is an agent $i \in I$ and house $h \in H$ with $h\, P_i\, \mu(i)$. In the following, we assume w.l.o.g. that i is the agent with $h\, P_i\, \mu(i)$ who has the highest priority for h among those agents.

Claim 2. *Let $j \in I$ be the agent such that $\mu(j) = h$. Then, $\mu(j) \neq \nu(j)$.*

Proof. Suppose $\mu(j) = h = \nu(j)$. Since $hP_i\mu(i)$, we know that there there is at least one agent k (for example, $k = i$) with $hP_k\mu_0(k)$. Among such agents, let k be the agent with the lowest priority at h. By the individual rationality of μ, we have $\mu_0(k) \neq h$. Consider the profile $\tilde{P}_k$ where agent k moves h below $\mu_0(k)$, i.e., $\tilde{P}_k|_{H\setminus\{h\}} = \tilde{P}_k|_{H\setminus\{h\}}$ and $\mu_0(k)\, \tilde{P}_k\, h$. Let $\tilde{P} = (\tilde{P}_k, P_{-k})$. Note that $M(\tilde{P}) < M(P)$.

We will now show that $DA^*(\tilde{P}) = DA^*(P)$. First, note that $\tilde{P}$ is a *monotonic transformation* of P at $DA^*(P)$, i.e., for each $h' \in H$ and $\ell \in I$, we have $h'\tilde{P}_\ell DA^*_\ell(P) \Rightarrow h' P_\ell DA^*_\ell(P)$. Note that the profile $\tilde{P}_k$ as we defined it for agent k downgrades house h below $\mu_0(k)$ so that it is indeed a monotonic transformation at $DA^*_k(P)$. As shown by Kojima and Manea (2010), the DA mechanism, and thus DA^*, is weakly Maskin monotonic. It means that all agents weakly prefer, according to their preferences, the matching returned by DA^* after a monotonic transformation of such preferences. Thus, it must be the case that for all agents, $DA^*(\tilde{P})$ is weakly preferred to $DA^*(P)$ at profile P. Assume that $DA^*(\tilde{P}) \neq DA^*(P)$. Then, the matching $DA^*(\tilde{P})$ Pareto dominates the matching $DA^*(P)$ at profile P. Moreover, the assignment of house h has changed and there is an agent $k' := DA^*_h(\tilde{P}) \neq j = \mu(h)$ who was rejected in favor of agent k at house h under $DA^*(P)$.[13] Thus, $k \succ_h k'$ and $h P_{k'} DA^*_{k'}(P) R_{k'} \mu_0(k')$, contradicting the assumption that k has the lowest priority at h among the agents who strictly prefer h to their initial house. Thus, we have $DA^*(\tilde{P}) = DA^*(P)$.

Now note that, by construction of the profile $\tilde{P}_k$, we have $\tilde{P}_k \neq P_k$ and $\tilde{P}_k|_{H\setminus\{h\}} = \tilde{P}_k|_{H\setminus\{h\}}$ so that $\tilde{P}_k$ moves h in P_k. So using IIAg, we either have $\varphi_h(\tilde{P}) \neq \varphi_h(P)$ or $\varphi(\tilde{P}) = \varphi(P)$. In both cases, $\varphi(\tilde{P}) \neq DA^*(P) = DA^*(\tilde{P})$ and therefore $\tilde{P} \in \mathcal{Q}$. Since $M(\tilde{P}) < M(P)$ this is a contradiction with the assumption on the choice of P. □

Claim 3. *For each $k \neq i$, we have $\mu_0(k)\, R_k\, \mu(i)$, i.e., $\mu(i)$ is unacceptable or the endowment for all other agents.*

Proof. Suppose there is an agent k with $\mu(i)\, P_k\, \mu_0(k)$. Choose k to be the agent with the lowest priority at $\mu(i)$ among such agents. Let $\tilde{P}_k$ be a profile such that $\tilde{P}_k|_{H\setminus\{\mu(i)\}} = P_k|_{H\setminus\{\mu(i)\}}$ and $\mu_0(k)\, \tilde{P}_k\, \mu(i)$. This profile is a monotonic transformation of P_k. Note that $M(\tilde{P}) < M(P)$. By Claim 1 applied to agent i and by the assumption that $h\, P_i\, \mu(i)$, we have $\mu(i) = \nu(i)$. Using the same argument that we used in Claim 2, we have $DA^*(\tilde{P}) = DA^*(P)$. As in Claim 2, we have that $\tilde{P}_k$ moves h in P_k so that, because φ respects IIAg, applied to agent k and house $\mu(i)$, either $\varphi(\tilde{P}) = \varphi(P)$ or $\varphi_i(\tilde{P}) \neq \varphi_i(P) = \nu(i) = \mu(i)$. In both cases, we have $\varphi(\tilde{P}) \neq DA^*(\tilde{P}) = DA^*(P)$ and therefore $\tilde{P} \in \mathcal{Q}$. Since $M(\tilde{P}) < M(P)$ this is a contradiction with the assumption on the choice of P. □

Now, consider the preference

$$P'_j : \mu(i)\, R'_j\, \mu_0(j)\, \ldots,$$

and let $P' := (P'_j, P_{-j})$.

Claim 4. *Define the matching μ' as follows:*

$$\mu'(i) = h, \quad \mu'(j) = \mu(i), \quad \mu'(k) = \mu(k) \text{ for } k \neq i, j.$$

Then, the matching μ' is μ_0-stable and individually rational under P'.

[13] Using the terminology of Kesten (2010), agent k was *an interrupter* at house h. Agent k is an interrupter at house h if while running $DA^*(P)$, he has been temporarily accepted at house h at Step t and later rejected at $t' > t$ and there has been an agent k' who has been rejected by house h at a Step $\ell \in \{t, t+1, \ldots, t'-1\}$.

Proof. Individual rationality for $k \neq i, j$ follows from the individual rationality of μ. Individual rationality for j follows by the definition of P'_j. Individual rationality for i follows by the assumption that $\mu'(i) = h\, P_i\, \mu(i)$ and by the individual rationality of μ.

Next, we show μ_0-stability. First, consider agent i. Agent i and $\mu(i)$ do not block μ' because $\mu'(i) = h\, P_i\, \mu(i)$. Moreover, for $h' \notin \{\mu(i), h\}$ we have $h' = \mu(k) = \mu'(k)$ for an agent $k \neq i, j$. If i and h' block μ', then $h'\, P_i\, \mu'(i) = h\, P_i\, \mu(i)$ and both i and h' would also block μ under P, contradicting the μ_0-stability of μ under P. Thus, there is no blocking pair involving i. Because agent j obtains his top choice in μ', he cannot be involved in a blocking pair. Finally, we consider $k \neq i, j$. By Claim 3, we have $\mu_0(k)\, R'_k\, \mu(i)$. Moreover, by the individual rationality of μ, we have $\mu(k)\, R'_k\, \mu_0(k)$. Thus, $\mu'(k) = \mu(k)\, R'_k\, \mu_0(k)\, R'_k\, \mu(i)$ and k and $\mu(i)$ do not block μ'. By assumption, i has highest priority for h among those agents who rank h strictly above their assignment under μ. Thus, if $h\, P'_k\, \mu(k) = \mu(k)$, then either $\mu_0(k) = h$ or $i \succ_h k$. The first possibility contradicts the individual rationality of μ under R. In the second case, k and h do not block μ'. Thus, k and h do not block μ'. Finally, k does not block μ' with a house $h' \neq h, \mu(i)$ because otherwise k and h' would block μ under R. □

By the construction of P'_j, $M(P') \leq M(P)$. Remember that because of Lemma 1, we know that DA^* returns the most preferred IR and μ_0-stable matching. As μ' is μ_0-stable and individually rational for P', it is Pareto-dominated by $DA^*(P')$. Therefore, we have

$$N(P') \leq \sum_{k \in I} |\{h' : h'\, P'_k\, \mu'(k)\}| < \sum_{k \in I} |\{h' : h'\, P_k\, \mu(k)\}| = N(P),$$

where the inequality in the middle is strict because $\mu'(i) = h\, P_i\, \mu(i)$. Thus, $DA^*(P') = \varphi(P')$ and $\varphi_j(P') = \mu(i)$. Next, let

$$\tilde{P}_j : \, h\, \tilde{P}_j \mu(i)\, \tilde{R}_j\, \mu_0(j).$$

By strategy-proofness applied to P' and $(\tilde{P}_j, P_{-j})$, we have $\varphi_j(\tilde{P}_j, P_{-j}) \in \{\mu(i), h\}$. By strategy-proofness applied to P and $(\tilde{P}_j, P_{-j})$ and by Claim 2, we have $\varphi_j(\tilde{P}_j, P_{-j}) \neq h$. Thus, $\varphi_j(\tilde{P}_j, P_{-j}) = \mu(i)$. Furthermore, note that $DA^*_i(\tilde{P}_j, P_{-j}) = \mu(i) = DA^*_i(P)$. Since $\varphi(\tilde{P}_j, P_{-j})$ is μ_0-stable at $(\tilde{P}_j, P_{-j})$, it is Pareto-dominated by $DA^*(\tilde{P}_j, P_{-j})$. Consequently, as $\phi_i(\tilde{P}_j, P_{-j}) \neq \mu(i)$, we have

$$\mu(i) = DA^*_i(\tilde{P}_j, P_{-j})\, P_i\, \varphi_i(\tilde{P}_j, P_{-j}),$$

in particular, $\mu^0(i) \neq \mu(i)$. Now, suppose i reports

$$\tilde{P}_i : \mu(i)\, \tilde{P}_i\, \mu_0(i) \ldots$$

By strategy-proofness, for $(\tilde{P}_j, P_{-j})$ and $\tilde{P} := (\tilde{P}_i, \tilde{P}_j, P_{-i,j})$, we have $\varphi_i(\tilde{P}) = \mu_0(i)$. By construction, $M(\tilde{P}) \leq M(P)$. Moreover, by the construction of $\tilde{P}$, μ is μ_0-stable and individually rational under $\tilde{P}$. Since $h\, P_i\, \mu(i)$ but $\mu(i)\, \tilde{P}_i\, h$, this fact implies

$$N(\tilde{P}) \leq \sum_{k \in I} |\{h : h\, \tilde{P}_i\, \mu(k)\}| < \sum_{k \in I} |\{h : h\, P_i\, \mu(k)\}| = N(P).$$

As a result, $\varphi(\tilde{P}) = DA^*(\tilde{P})$. However, because $\varphi_i(\tilde{P}) = \mu_0(i)$ and μ is a μ_0-stable matching and individually rational matching under $\tilde{P}$, $DA^*_i(\tilde{P}) = \mu(i)\tilde{P}_i\mu_0(i) = \varphi_i(\tilde{P})$. Therefore, we have a contradiction.

We conclude that $\varphi = DA^*$. □

We conclude this section by showing that the axioms used in Theorem 1 are independent.

Dropping IR. The standard DA (without modifying the priority structure) is an SP, μ_0-stable, unanimous and IIAg mechanism.

Dropping unanimity. The trivial mechanism that assigns every agent to his or her initial house is IR, SP, μ_0-stable and IIAg.

Dropping μ_0-stability TTC is an IR, SP, and IIAg mechanism.

Dropping IIAg. The mechanism in Example 1 is IR, SP, μ_0-stable and unanimous.

Dropping SP. Consider three agents $I = \{a, b, c\}$, three houses $H = \{h_a, h_b, h_c\}$ and an initial matching μ_0 s.t. $\mu_0(k) = h_k$ for $k \in I$. Consider the priority relation $\succ$ such that

$$\begin{array}{llll} \succ_{h_a}: & a & c & b \\ \succ_{h_b}: & b & a & c \\ \succ_{h_c}: & c & b & a \end{array}$$

Now, let $\mathcal{P}^* \subset \mathcal{R}$ be the set of preference profiles R such that

$$\begin{array}{llllllll} P_a: & h_b & P_a & \dots \\ P_b: & h_c & P_b & \dots \\ P_c: & h_b & P_c & h_a & P_c & h_c & P_c \end{array}$$

Let φ be the mechanism defined as follows:

$$\varphi(P) = \begin{cases} \mu_0 & \text{if } P \in \mathcal{P}^* \\ DA^*(P) & \text{if } P \notin \mathcal{P}^* \end{cases}$$

It is easy to see that φ is IR and μ_0-stable. Since all of the profiles in $\mathcal{P}^*$ are not unanimous profiles and since φ is the DA^* mechanism outside $\mathcal{P}^*$, φ is a unanimous mechanism. It is also easy to see that φ is not an SP mechanism. At any preference profile $P \in \mathcal{P}$, agent c can manipulate φ by reporting profile $P'_c : h_a, h_c, h_b$.

Here, we show that φ is IIAg. Since we know that DA^* is IIAg, we only need to check two cases. First, when an agent starts at a profile in $\mathcal{P}^*$ and a deviation as defined in the definition of IIAg would lead to a profile outside of $\mathcal{P}^*$. Second, when an agent starts at a profile outside of $\mathcal{P}^*$ and a deviation as defined in the definition of IIAg would lead to a profile inside of $\mathcal{P}^*$. To begin, we take a profile $P \in \mathcal{P}^*$ and fix agent a.

- Assume that $P_a : h_b, h_a, h_c$. If she moves down h_b, then either the matching stays at μ_0 or b and c exchange their houses so that IIAg is trivially respected. If she moves up h_c so that $P'_a : h_c, h_b, h_a$, then note that the matching of $DA^*(P')$ assigns a to h_b, b to h_c and c to h_a. Therefore, IIAg is respected again.

- Assume that $P_a : h_b, h_c, h_a$. We have seen that moving down h_b or moving up h_c so that $P'_a : h_c, h_b, h_a$ forces all of the agents to exchange their houses for IIAg to be respected. If she moves down h_b so that $P'_a : h_c, h_a, h_b$, then since $b \succ_{h_c} a$, $DA^*(P')$ allocates b in h_c and c in h_b and IIAg is still satisfied.

Now, we fix agent b.

- Assume that $P_b : h_c, h_b, h_a$. By moving down h_c, either the matching stays at μ_0 or a and c exchange their houses so that IIAg is respected. If she moves up h_a so that $P'_b : h_a, h_b, h_c$, then because $c \succ_{h_a} b$, $DA^*(P')$ allocates b in h_c and c in h_b, IIAg is still satisfied.

- Assume that $P_b : h_c, h_a, h_b$. By moving down h_c or moving up h_a such that $P_b' : h_a, h_c, h_b$, because $a \succ_{h_b} c$ and $c \succ_{h_a} b$, $DA^*(P')$ would assign a to h_b, b to h_c and c to h_a. Consequently, IIAg would be trivially respected. By reporting $P_b' : h_a, h_b, h_c$, then b would stay at his initial house h_b and so IIAg would be respected whether a and c exchange their houses or not.

Consider agent c.

- Moving down h_b or moving up h_a would make everyone exchange their houses with a assigned to h_b, b assigned to h_c and c assigned to h_a for IIAg to be respected.

- Moving down h_a so that $P_c' : h_b, h_c, h_a$ would make c stay at his initial house since $a \succ_{h_b} c$. In that case, IIAg is trivially respected independently of whether a and b exchange their houses.

We conclude that by starting from any profile $P \in \mathcal{P}^*$, IIAg is respected. Now, we start with a profile $P \notin \mathcal{P}^*$ and fix agent a. We will make a change to P_a' as defined in the definition of IIAg so that the new profile $P' \in \mathcal{R}^*$. In particular, at the initial profile P, $P_c = h_b, h_a, h_c$, agent b ranks house h_c first in P_b and agent a does not rank house h_b first.

- Assume that $P_a : h_a, ...$ Then, a stays at his house h_a. By moving down h_a or moving up h_b, $P_a' : h_b, \ldots, \varphi(P') = \mu_0$. So if b and c were also staying at their initial houses under $\varphi(P)$, then IIAg would be trivially satisfied. If b and c were exchanging their houses, the report P_a' would make the assignment of all houses except h_a change so that IIAg would again be satisfied.

- Assume that $P_a : h_c, h_a, h_b$. In that case, since $b \succ_c a$, then $\varphi(P) = DA^*(P)$ does not assign a to h_c. Hence, a stays at hid initial house h_a. Then, the same argument as above applies.

- Assume that $P_a : h_c, h_b, h_a$. In that case, one can check that $\varphi(P) = DA^*(P)$ assigns a to h_b, b to h_c and c to h_a. Since any change of preference profile from P_a to P_a' by agent a so that $R' \in \mathcal{P}^*$ would lead to $\varphi(P') = \mu_0$, again IIAg is trivially satisfied.

Now, consider agent b and start at a profile $P \notin \mathcal{P}^*$ where $P_c = h_b, h_a, h_c$, agent a ranks h_b first and agent b does not rank h_c first.

- Assume that $P_b : h_b, ...$ Then, b stays at his initial house h_b under $\varphi(P)$. Thus, a similar argument to the above applies and IIAg is respected.

- Assume that $P_b : h_a, h_b, h_c$. One can check that $\varphi(P) = DA^*(P)$ makes b stay at his initial house h_b so that IIAg is respected if the profile moves to $P' \in \mathcal{P}^*$.
- Assume that $P_b : h_a, h_c, h_b$. Again, one can check that all of the agents exchange their houses under DA^*, which implies that IIAg is trivially respected.

Lastly, consider agent c and start at a profile $P \notin \mathcal{P}^*$ where agent a ranks h_b first and agent b ranks h_c first.

- Assume that $P_c : h_c, ...$ Then, c stays at his initial house h_c under $\varphi(P)$. By a similar argument to the above, IIAg is respected.
- Assume that $P_c : h_a, h_c, h_b$ or $P_c : h_a, h_b, h_c$. In that case, all of the agents exchange their houses under $\varphi(P) = DA^*(P)$ and a is assigned to h_b, b is assigned to h_c and c is assigned to h_a. In that case, any change of preferences from P_a to P'_a so that the new profile $P' \in \mathcal{R}^*$ leads to $\varphi(P) = \mu_0$. Consequently, IIAg is trivially satisfied.

We conclude that φ is an IR, μ_0-stable, unanimous and IIAg mechanism that is not strategy-proof.

4 Discussion of IIAg and additional results

Next we discuss the relation of the IIAg axiom to other classic axioms and afterwards discuss which classical matching mechanisms (other than DA) satisfy it.

IIA, Monotonicity and IIAg. The *Independence of Irrelevant Alternatives* (IIA) axiom was introduced by Arrow, 1951 in the context of social welfare functions which must produce a ranking over alternatives based on the profile of rankings of agents. Our setting is closer to the one of social choice functions which have to return an unique alternative based on individuals' rankings. In this context, the natural counterpart of IIA is the invariance axiom famously called Maskin monotonicity.[14] In standard social choice theory, a mechanism is (Maskin) monotonic if whenever a social choice μ is chosen by the mechanism and the preference profile of the agents is P, then μ continues to be chosen at any profile P' where all of the agents rank

[14]The condition was initially called Strong Positive Association (SPA) by Muller and Satterthwaite (1977) where they discuss how it can be seen as the natural counterpart of IIA for social choice functions. Maskin (1999) used the same condition in the context of implementation calling it monotonicity and pointing out its equivalence with SPA.

μ weakly higher than under profile P.[15] Our IIAg axiom restricts social choice in a similar way. However, there are two important differences. We are not in a general social choice model but in a private goods setting where agents have preferences only over their final allocated houses and are indifferent between different matchings at which they obtain the same house. Furthermore, we allow an agent to arbitrarily change the ranking of house h, not only improve it. We also have an additional condition that checks whether the assignment of the house has changed once it has been upgraded. In the social choice context, the Maskin monotonicity axiom states that once the choice has been made, the pieces of information of the houses ranked above the choice by each agent (which have therefore been disregarded by the mechanism) should not impact the initial choice that was made. Our IIAg axiom states that whenever an agent change the ranking of a house that he is not assigned to and that this change is irrelevant to determining the allocation of that house, then it should not impact the overall assignment. Thus, the choice rule must remain consistent to the matching it selected before the change of preferences of that agent.

Non-Bossiness and IIAg. Our IIAg axiom shares some similarities with another important independence axiom: *non-bossiness*. The latter states that whenever an agent changes his preference report from P_i to P'_i, if his assignment stays the same after this change (i.e. $\varphi_i(P) = \varphi_i(P'_i, P_{-i})$), then the whole matching should stay the same (i.e. $\varphi(P) = \varphi(P'_i, P_{-i})$). For an extensive discussion of the axiom, one can refer to Thomson (2016). Non-bossiness is a strong property which is often incompatible with μ^0-stability. However, we can relax the axiom so that it becomes compatible with μ^0-stablility. A minimal relaxation of non-bossiness that guarantees this, needs to make sure that the axiom applies whenever the matchings chosen under the two profiles P and (P'_i, P_{-i}) are μ^0-stable under the other profile.[16] More precisely, we define:

Axiom (Minimal Bossiness)**.** *A mechanism φ is Minimally Bossy (MB) if $\forall i \in I$,*

[15] In standard social choice, agents have complete linear orderings over the set of *alternatives*, which in our context would be equivalent to matchings.

[16] If a matching chosen under one of the two profiles fails to be μ^0-stable under the other profile, we cannot require non-bossiness to hold at these profiles, since imposing it would require us to select a μ^0-unstable matching at one profile. A stronger, asymmetric version of the axiom would require that non-bossiness holds whenever the matching selected at P is μ^0-stable under (P'_i, P_{-i}), but not necessarily vice versa. This stronger version of the axiom, however, would be violated by DA^* or other unanimous and μ^0-stable mechanisms, as it would require, for example, that the same matching is selected even if the set of μ^0-stable matchings enlarges through the preference change of agent i.

$P, \tilde{P} \in \mathcal{P}$ with $\tilde{P}_{-i} = P_{-i}$ such that $\varphi(\tilde{P})$ is μ^0-stable under P and $\varphi(P)$ is μ^0-stable under $\tilde{P}$:

$$\varphi_i(\tilde{P}) = \varphi_i(P) \Rightarrow \varphi(\tilde{P}) = \varphi(P).$$

It is straightforward to see that for strategy-proof mechanisms, MB implies IIAg.

Proposition 1. *Let φ be a μ_0-stable, SP and MB mechanism. Then it satisfies IIAg.*

Proof. Consider $i \in I$, $h \in H$, and $P, \tilde{P} \in \mathcal{P}$ such that $\tilde{P}_{-i} = P_{-i}$, $\tilde{P}_i$ moves h in P_i and $\varphi_h(P) \neq i$. By strategy-proofness of φ we have $\varphi_i(\tilde{P}) = h$ or $\varphi_i(\tilde{P}) = \varphi_i(P)$. In the first case, $\varphi_h(\tilde{P}) \neq \varphi_h(P)$ and IIAg holds. In the second case, by minimal bossiness we have $\varphi(\tilde{P}) = \varphi(P)$ or $\varphi(P)$ is not μ_0-stable under $\tilde{P}$ or $\varphi(\tilde{P})$ is not μ_0-stable under P. In the first subcase, IIAg holds. In the second subcase, h and i block $\varphi(P)$ under $\tilde{P}$. In that case, if $\varphi_h(\tilde{P}) = \varphi_h(P)$ then, as $\varphi_i(\tilde{P}) = \varphi_i(P)$, i and h also block $\varphi(\tilde{P})$ under $\tilde{P}$ which contradicts the μ_0-stability of φ. If $\varphi_h(\tilde{P}) \neq \varphi_h(P)$, then IIAg holds. In the third subcase, i and h block $\varphi(\tilde{P})$ under P. In that case, if $\varphi_h(\tilde{P}) = \varphi_h(P)$ then, as $\varphi_i(\tilde{P}) = \varphi_i(P)$, i and h also block $\varphi(P)$ under P which contradicts the μ_0-stability of φ. If $\varphi_h(\tilde{P}) \neq \varphi_h(P)$, then IIAg holds. □

Moreover, one can verify that DA^* satisfies minimal-bossiness:

Proposition 2. *DA^* is minimally bossy.*

Proof. Consider $i \in I$ and $P, \tilde{P} \in \mathcal{P}$ such that $\tilde{P}_{-i} = P_{-i}$. Suppose that $DA^*_i(\tilde{P}) = DA^*_i(P)$, that $DA^*(\tilde{P})$ is μ_0-stable under P and that $DA^*(P)$ is μ_0-stable under $\tilde{P}$. Then, remember that, by Lemma 1, DA^* returns the $\succ^*$- stable matching among μ_0-stable and IR matchings, $DA^*(\tilde{P})$ weakly Pareto dominates $DA^*(P)$ under $\tilde{P}$ and $DA^*(P)$ weakly Pareto dominates $DA^*(\tilde{P})$ under P. Since $\tilde{P}_{-i} = P_{-i}$, this implies that all agents $j \neq i$ receive the same assignment under $DA^*(P)$ and $DA^*(\tilde{P})$. Thus, $DA^*(\tilde{P}) = DA^*(P)$. □

Using Proposition 1 together with Theorem 1, we can immediately deduce the following characterization:

Corollary 1. *A mechanism is IR, SP, μ_0-stable, unanimous and MB if and only if it is the DA^* mechanism.*

We work with IIAg though for two reasons. IIAg is an axiom with normative appeal even in a context without priorities whereas minimal bossiness is only well-defined in conjunction with μ_0-stability. Second minimal-bossiness is a (much) more demanding notion that makes the characterization result arguably normatively less convincing. The following example illustrates this:

Example 2. *Consider three agents $I = \{a, b, c\}$, three houses $H = \{h_a, h_b, h_c\}$ and an initial matching μ_0 s.t. $\mu_0(k) = h_k$ for $k \in I$. Consider a priority relation $\succ$ such that we have the following:*

$$\begin{array}{llll} \succ_{h_a}: & a & b & c \\ \succ_{h_b}: & b & a & c \\ \succ_{h_c}: & c & a & b \end{array}$$

We define φ as an IR, strategy-proof, μ_0-stable and IIAg mechanism that is not minimally-bossy: Define $\mathcal{P}'$ to be the set of all profiles P where agent a has preferences

$$P_a: \quad h_a \quad P_a \quad h_b \quad P_a \quad h_c.$$

Define a mechanism

$$\varphi(P) := \begin{cases} DA^*(P), & \text{if } P \notin \mathcal{P}', \\ \mu_0, & \text{if } P \in \mathcal{P}'. \end{cases}$$

One readily checks that φ is IR and μ_0-stable. To check that φ is strategy-proof note that agent a always obtains his DA^* assignment under φ and that it does not depend on agents b and c's ranking whether the profile is in $\mathcal{P}'$ or not. The mechanism is not minimally bossy, because at profile P where P_a is as above and

$$\begin{array}{llllll} P_b: & h_c & P_b & h_b & P_b & h_a \\ P_c: & h_b & P_b & h_c & P_b & h_a \end{array}$$

we have $\varphi(P) = \mu_0$ but if agent a changes to

$$P'_a: \quad h_a \quad P'_a \quad h_c \quad P'_a \quad h_b,$$

we have $\varphi(P') = DA^*(P') \neq \mu_0$.

The mechanism is IIAg because for problems with only three agents any strategy-proof mechanism satisfies IIAg: if an agent i moves a house h in her preferences then by strategy-proofness $\varphi_i(\tilde{P}) = \varphi_i(P)$ or $\varphi_i(\tilde{P}) = h$. In the first case, $\varphi_h(\tilde{P}) = j = \varphi_h(P)$ implies that the third agent $k \neq i, j$ also gets assigned the same house. In the second case $\varphi_h(\tilde{P}) = i \neq \varphi_h(P)$ and IIAg holds.

IIAg and other mechanisms. Last, one may wonder whether standard mechanisms of the literature respect the IIAg axiom, the following proposition gives the answer for the three main ones studied in the school choice literature:

Proposition 3. *The Deferred Acceptance and the Top Trading Cycles mechanisms respect the IIAg axiom while the Immediate Acceptance mechanism*[17] *does not.*

Proof. For the DA mechanism, the part of the proof of Theorem 1 below showing that DA^* respects IIAg can be applied to show that DA respects IIAg. For the TTC mechanism, recall that the mechanism is strategy-proof and non-bossy. Consider a profile P and an agent i that under $\tilde{P}_i$ moves house h in P_i for a house h that he was not assigned to under $TTC(P)$. By strategy-proofness at $\tilde{P}_i$, i either obtains house h or his previous assignment, $TTC_i(\tilde{P}) = h$ or $TTC_i(\tilde{P}) = TTC_i(P)$. In the first case, house h is assigned differently. In the second case, by non-bossiness of TTC, we have $TTC(\tilde{P}) = TTC(P)$. Thus, IIAg is satisfied.

For the Immediate Acceptance mechanism, the following example shows that it does not respect IIAg: Consider three agents $I = \{a, b, c\}$ and three houses $H = \{h_a, h_b, h_c\}$. Consider a priority relation $\succ$ such that we have the following:

$$\begin{array}{llll} \succ_{h_a}: & a & c & b \\ \succ_{h_b}: & b & c & a \\ \succ_{h_c}: & c & a & b \end{array}$$

Consider preference profiles P such that

$$\begin{array}{llllll} P_a: & h_a & P_a & \dots & & \\ P_b: & h_c & P_b & h_b & P_b & h_a \\ P_c: & h_a & P_c & h_c & P_b & h_b \end{array}$$

and preferences $\tilde{P}_c$ such that

$$\tilde{P}_c: \quad h_c \quad \tilde{P}_c \quad h_a \quad \tilde{P}_b \quad h_b$$

The immediate acceptance mechanism for R assigns the matching:

$$\mu(a) = h_a, \quad \mu(b) = h_c, \quad \mu(c) = h_b.$$

[17] By Top Trading Cycle, we refer to the housing market original version as proposed by Shapley and Scarf (1974). We do not give the formal definition of Immediate Acceptance which is well known in the school choice literature, also called the Boston mechanism, but the reader can refer to Abdulkadiroglu and Sonmez (2003).

Note that the house h_a is not assigned to agent c and that preferences $\tilde{P}_c$ are obtained from R_c by changing the ranking of h_a without changing the relative ranking of the other two houses. With the changed preferences, the immediate acceptance mechanism assigns the matching:

$$\tilde{\mu}(a) = h_a, \quad \tilde{\mu}(b) = h_b, \quad \tilde{\mu}(c) = h_c.$$

Thus, house h_a is assigned in the same way but the overall matching has changed, violating IIAg. □

5 Conclusion and discussions

We study the problem of reallocation with priorities where agents initially own a house and each house has a priority order over agents. We show that this problem has distinct features from its counterparts, namely, the marriage problem and the housing market problem. In particular, we show that the DA^* mechanism, a natural adaptation of the Deferred Acceptance to this context, is not the only individually rational, strategy-proof and μ_0-stable mechanism. Adding a simple efficiency requirement such as unanimity does not rule out other mechanisms than DA^*. We introduce a new axiom, called irrelevance of independent agents which states that if an agent ranking of a house not assigned to him has no impact on the assignment of that house, then a mechanism should keep the same matching that was selected prior to that change. We discuss how this axiom is related to the standard axiom of irrelevance of independent agents and the axiom of non-bossiness. In particular, we show that a natural stability adapted variable of non-bossiness, that we call Minimal Bossiness, implies our axiom of irrelevance of independent agents.

Extensions and assumptions. It is a natural question whether and how our results generalize to the house allocation model with existing tenants (Abdulkadiroglu and Sonmez, 1999) in which some of the agents could initially be unassigned and some houses could initially be vacant. All the arguments used in the proof of Theorem 1 can be easily extended to this case so that our main characterization still holds in this context.

As discussed in Footnote 4, our model imposes two assumptions: 1) the non-availability of an outside option and 2) the existence of an initial matching. If one relaxes the first one and naturally adapts the definitions of the model accordingly then the proof of Alcalde and Barberà (1994) would still be valid and would show that DA* is the unique IR, strategy-proof and μ_0-stable mechanism. If one keeps the restriction that outside options are not available but assumes that there is no initial matching, our main message remains: DA is not the only stable and strategy-proof mechanism in that setting. Indeed, the ownership adjusted Deferred Acceptance mechanism highlights that IR can be seen as a restriction on the priority profiles of the houses: we only consider **unanimous priority profiles**, i.e. profiles where there exists a matching giving each house its top priority agent. Thus, our characterization perfectly applies to this school choice setting and it can be seen as a characterization of the DA mechanism on the domain of unanimous priority profiles when outside options are not available.

Future directions. Our paper highlights that the problem of reallocation with priorities has distinct differences from its counterparts, namely, the marriage problem and the housing market problem. Thus, there is still significant research to be done to study the specific properties of this problem. In our analysis, we take the priority structure as given and study the properties of the μ_0-stable mechanism DA^*. We follow the path of the school choice literature (Abdulkadiroglu and Sonmez, 2003). In such models, one can consider the priorities of one side of the market, e.g., the schools or the houses, as fixed and provide a charaterization of the DA mechanism, this is what Balinski and Sonmez (1999) or Morrill (2013) have done. When priorities are not considered as preferences per se but as part of the design of the mechanism, one can see the DA mechanism as a class of mechanisms, where there is one for each profile of priorities. Kojima and Manea (2010) were the first to propose two axiomatic characterizations of the DA mechanisms, and they introduced the axioms of *individually rational monotonicity* and *weak Maskin monotonicity*. Whereas these authors' first results showed that the (student proposing) DA with acceptant and substitutable choice functions is the only non-wasteful and individually rational monotonic mechanism, their second characterization uses the axioms of non-wastefulness, population monotonicity, and weak Maskin monotonicity. Later, Ehlers and Klaus (2016) provided two characterizations of DA using a set of standard axioms: unavailable-type-invariance, individual-rationality, weak non-wastefulness, truncation-invariance, strategy-proofness and either population-monotonicity or resource-monotonicity. It would be an interesting line of future research to investigate whether it is possible to endogenize the priority structure of DA^* with a set of axioms. The key difficulty is that axioms such as population-monotonicity or resource-monotonicity impose constraints when one adds only one additional agent or one additional house. In our reallocation setting, agents are all initially assigned to a house so that both a house and an agent would be added to the market, making comparative static results difficult. Moreover, non-wastefulness has no applicability in our setting since all of houses are initially assigned and each reassignment of resources is non-wasteful by definition.

Lastly, as we discussed above, our characterization can also be seen as a result on a school choice model where one jointly restrict the domains of preference and priority profiles. As mentioned in the Introduction, this joint restriction can be motivated by important applications such as the assignment of teachers to schools. This possibility of joint restrictions has not been studied so far and opens many interesting research questions such as for instance the maximal joint domains on which DA is the unique strategy-proof and stable mechanism.

References

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2009): "Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match," *American Economic Review*, 99(5), 1954–78.

Abdulkadiroglu, A., and T. Sonmez (1999): "House Allocation with Existing Tenants," *Journal of Economic Theory*, 88, 233–260.

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

Akbarpour, M., A. Kapor, C. Neilson, W. L. Van Dijk, and S. Zimmerman (2020): *Centralized School Choice with Unequal Outside Options*. Princeton University, Industrial Relations Section.

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

Arrow, K. J. (1951): *Social choice and individual values*. Yale university press.

Balinski, M., and T. Sonmez (1999): "A tale of two mechanisms: student placement," *Journal of Economic Theory*, 84, 73–94.

Combe, J., O. Tercieux, and C. Terrier (2016): "The Design of Teacher Assignment: Theory and Evidence," Discussion paper, forth. Review of Economic Studies.

Compte, O., and P. Jehiel (2008): "Voluntary Participation and Re-assignment in Two-sided Matching," Paris School of Economics, Unpublished mimeo.

Dur, U. M., and M. U. Ünver (2019): "Two-sided matching via balanced exchange," *Journal of Political Economy*, 127(3), 1156–1177.

Ehlers, L., and B. Klaus (2016): "Object allocation via deferred-acceptance: Strategy-proofness and comparative statics," *Games and Economic Behavior*, 97, 128–146.

Ergin, H. (2002a): "Efficient Resource Allocation on the Basis of Priorities," *EMA*, 70, 2489–2498.

Ergin, H. I. (2002b): "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 70, 2489–2497.

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

Guillen, P., and O. Kesten (2012): "Matching Markets with Mixed Ownership: The Case for a Real-life Assignment Mechanism," *International Economic Review*, 53(3), 1027–1046.

Kesten, O. (2010): "School choice with consent," *The Quarterly Journal of Economics*, 125(3), 1297–1348.

Kesten, O., and M. Kurino (2017): "Strategy-proof improvements upon deferred acceptance: A maximal domain for possibility," Discussion paper, working paper.

Kojima, F., and M. Manea (2010): "Incentives in the Probabilistic Serial Mechanism," *Journal of Economic Theory*, 145, 106–123.

Ma, J. (1994): "Strategy-Proofness and the Strict Core in a Market with Indivisibilities," *International Journal of game Theory*, 23, 75–83.

——— (1995): "Stable matchings and rematching-proof equilibria in a two-sided matching market," *Journal of Economic Theory*, 66, 352–369.

Maskin, E. (1999): "Nash equilibrium and welfare optimality," *The Review of Economic Studies*, 66(1), 23–38.

MORRILL, T. (2013): "An alternative characterization of the deferred acceptance algorithm," *International Journal of Game Theory*, 42(1), 19–28.

MULLER, E., AND M. A. SATTERTHWAITE (1977): "The equivalence of strong positive association and strategy-proofness," *Journal of Economic Theory*, 14(2), 412–418.

PEREYRA, J. S. (2013): "A Dynamic School Choice Model," *Games and economic behavior*, 80, 100–114.

PYCIA, M., AND M. UNVER (2021): "Outside options in neutral allocation of discrete resources," Discussion paper, Working paper, Boston College, 2021.[26].

PYCIA, M., AND M. U. ÜNVER (2017): "Incentive compatible allocation and exchange of discrete resources," *Theoretical Economics*, 12(1), 287–329.

SHAPLEY, L., AND H. SCARF (1974): "On Cores and Indivisibility," *Journal of Mathematical Economics*, 1, 22–37.

SÖNMEZ, T., AND M. U. ÜNVER (2005): "House allocation with existing tenants: an equivalence," *Games and Economic Behavior*, 52(1), 153–185.

SVENSSON, L.-G. (1999): "Strategy-proof Allocation of Indivisible Goods," *Social Choice and Welfare*, 16.

THOMSON, W. (2016): "Non-bossiness," *Social Choice and Welfare*, 47(3), 665–696.