Matching with a Status Quo: The Agreeable Core

Peter Doe*

June 12, 2024

Abstract

We provide a framework to unify classic models of two-sided matching with recent models of recontracting. In the classic model, agents from two sides must be matched; in models of recontracting, agents improve a status quo match. We generalize the core (matches not blocked by any coalition) from cooperative game theory to our setting by restricting the set of permissible coalitions to coalitions containing neither or both agents in a status quo match, dubbed "agreeable" coalitions. The *agreeable core* is the set of all weak improvements of the status quo that are not blocked by any agreeable coalition. Our main result is that the agreeable core is nonempty and can be found through a computationally efficient and economically meaningful algorithm: our Propose-Exchange algorithm. The applications of the agreeable core include early decision, out-of-match agreements in the NRMP, matching with minimum constraints, and efficiency in school choice.

1 Introduction

The classic theory of two-sided matching models markets *ex ante* to any agreements made by agents. We introduce a framework to describe two-sided matching markets *ex interim* when some agents are already matched while others are not. Such agreements form the *status quo* and can only be broken by the mutual agreement of both agents in the pair. Standard solutions such as the core often supply outcomes that harm some agents compared to the status quo, violating the guarantee provided by the status quo. In this paper we provide a more general notion of the core that (weakly) improves the status quo yet maintains an appropriate definition of blocked matches.

Numerous markets exist with *ex interim* contracts which are costly to break: early decision agreements in college admissions, multi-year contracts between athletes and sports teams, and the outside-the-match residencies of the NRMP, among many others. Market designers have limited

*pdoe@caltech.edu, Division of the Humanities and Social Sciences, California Institute of Technology

tools in these markets. Depending on the application, the standard practice in the literature is to exclude either agents who hold a contract (as in the NRMP) or agents who do not hold a contract (Combe et al., 2022; Dur and Ünver, 2019).[1] This deficiency has severely limited the applications to markets with regular recruitment cycles (the NRMP or school choice) or markets in which newly arriving agents are undesirable (teacher reassignment). Treating matched and unmatched agents separately lowers the overall efficiency of the mechanism and fails to incorporate the voluntary nature of contracts – an existing contract is not a constraint if both parties agree to cancel it. The theory we present unifies the two models to provide a general model applicable to substantially more matching markets.

The foundation of our analysis is a reinterpretation of the fundamental justification for the core as the outcomes that "cannot be profitably upset by the collusive action of any subset of the participants, *acting by themselves*" (Shapley and Scarf, 1974), emphasis ours. We model the status quo as restricting the set of coalitions to only those that include exactly neither or both partners of a status quo match. These coalitions are dubbed *agreeable* because they can only block one match by recommending another that both partners in a status quo match agree to. If a coalition is not agreeable, then the coalition cannot act by itself – any agent whose status quo partner is in the coalition may veto the action. The *agreeable core* is the set of weak improvements of the status quo that are not blocked by any agreeable coalition. These matches are immune to deviations by coalitions that are not vetoed by some agent.

Our main result is a constructive algorithm that always outputs a match in the agreeable core for strict preferences. No other assumptions are needed on preferences or the status quo. Our *Propose-Exchange* (PE) algorithm is a nuanced hybrid of the Deferred Acceptance (DA, the "Propose" phase) and Top Trading Cycle (TTC, the "Exchange" phase) algorithms. The broad design of the PE is to divide the market into those agents who cannot form a new match freely because their status quo partners cannot find better matches, and those who can. The former group finds a match by a variation of the DA while the latter group finds a match by a variation of the TTC .

The Propose-Exchange algorithm offers interpretive insights into the status quo's effect on the resulting match and its connection with one-sided allocation problems. When no agent is matched in the status quo, the PE emulates the DA, which results in a pairwise stable match. When all agents have a status quo match, the PE algorithm emulates a variation of the TTC, the TO-BE of Combe et al. (2022). In mixed environments when some agents have a status quo match and others

[1] A notable exception is Abdulkadiroğlu and Sönmez (1999), who study a one-sided allocation problem when some agents already own objects.

do not, the PE first attempts to satisfy both partners in a status quo pair individually by releasing one agent from the pair to either make or receive proposals only when the other has received or not (yet) been rejected by some more preferred partner (the propose phase). The PE then allows the remaining agents to exchange status quo contracts in a Pareto efficient manner.

The agreeable core has broad applications beyond markets with a clear set of *ex interim* contracts. School choice with minimum quotas can be implemented through a straightforward assignment of status quo matches (if every school assigned is assigned its minimum quota of students in the status quo match, then none will be matched to fewer students). Similarly, the Boston Public Schools opted for the DA over the more efficient TTC because the TTC allows students to trade all priorities when they believed that "certain priorities – e.g., sibling priority – apply only to students for particular schools and should not be traded away."[2] The PE naturally incorporates both tradable and nontradable priorities through the status quo (see section 7 for details). These and other applications are remarkably simple in our framework.

Like many mechanisms in two-sided matching, the PE algorithm suffers from manipulability. Although both the DA and the TTC are strategyproof for one side of the market, their combination is not. Some agents in the Exchange phase of the algorithm have an incentive to misreport their preferences to affect the match in the Propose phase and induce more agents to participate in the Exchange phase. Our second result shows that manipulability is a universal property of mechanisms whose output is in the agreeable core.

The rest of the paper is organized as follows. The remainder of the introduction highlights the connections between this paper and the classic literature on two-sided matching and the more recent developments on matching with a status quo and matching with minimum constraints. Section section 2 contains the formal model and defines the core. Section section 3 illustrates the emptiness of the core in the presence of status quo matches and the existence of the agreeable core. The intuition built in the example informs our definition of the agreeable core in Section section 4. We answer the question of existence in section 5 by introducing our Propose-Exchange algorithm. Section 6 explores related questions of manipulability and structure, and section 7 closes with a discussion of the agreeable core and its applications.

1.1 Connection to the Literature

This paper develops and connects several literatures on two-sided matching. An exhaustive review of the literature is far beyond the scope of this paper, so we list the only the most closely related work and its connections with this paper.

[2] We could not find the original source of this quote, so we copy it from Morrill (2013).

We integrate the classic model of two-sided matching with recent advances in recontracting. In the classic model, a stable match always exists and can be found in by the DA (Gale and Shapley, 1962). It is well known that the set of pairwise-stable matches corresponds to the core of a related cooperative game (Roth and Sotomayor, 1990). Later research largely discarded the connection with the core in favor of pairwise-stability notions. When considering matching with a status quo (in which the intersection of pairwise stable and individually rational outcomes may be empty), Pereyra (2013); Guillen and Kesten (2012) generalize pairwise-stability by partitioning claims between valid and invalid claims and then removing all valid claims. This may be strongly inefficient (Combe and Schlegel, 2024; Combe et al., 2022), and hence a mechanism with minimal envy is considered (Kwon and Shorrer, 2023). Although efficient, these minimal envy mechanisms are inscrutable to participants: the designer allows some claims but not others only because doing so minimizes some objective. Our paper advances this literature by reconnecting the status quo back to the core, a more interpretable solution. We both minimize envy as in Kwon and Shorrer (2023) but also provide a clear definition of valid and invalid claims as in Pereyra (2013).

Research in school choice has made extensive use of both the DA and TTC. Abdulkadiroğlu and Sönmez (2003) suggests the Deferred Acceptance (DA) algorithm from Gale and Shapley (1962) or the Top Trading Cycles (TTC) algorithm from Shapley and Scarf (1974) as desirable and implementable solutions. Both algorithms run in polynomial time, are relatively easy to describe, and are strategyproof. The DA is fair while the TTC is efficient. A plethora of researchers seek to combine the two algorithms to gain improve efficiency by allowing certain priority violations (Abdulkadirog, 2011; Dur et al., 2019; Kesten, 2006; Kwon and Shorrer, 2023; Reny, 2022; Troyan et al., 2020; Morrill, 2013; Dur and Morrill, 2017). Papers in this vein typically define a set of properties of a mechanism (such as the allowable priority violations, efficiency, strategyproofness, etc.), and then present a satisfactory algorithm, typically a variation of the DA or TTC (usually the unique such one). Our work complements this approach by an algorithm derived from first principles rather than with specific objectives in mind.

A connected branch of matching theory develops methods for matching with minimum quotas. Schools are modeled as having both a maximum capacity for students but also a minimum required quota of students. One approach is to allow for wasted seats but not envy (Fragiadakis and Troyan, 2017). A separate approach uses an auxiliary "master list" (Ueda et al., 2012) or "precedence list" (Fragiadakis et al., 2016; Hamada et al., 2017) as a means to break ties: if two students wish to take an empty seat but the minimum quota requires that only one may do so, the list determines which worker can. The algorithms described in both approaches typically either sacrifice efficiency (based on the DA) or fairness (based on the TTC), and both require that all agents are mutually acceptable. We develop both approaches by endogenizing the master list into the status quo and not requiring any assumptions on preferences. Although a master list is natural in some applications, whether a master list or the status quo is more appropriate depends on the application.

Surprisingly, no authors have connected matching with minimum quotas and the matching with a status quo. We combine these subfields with the observation that, if the status quo provides a guarantee for both workers *and firms*, then minimum quotas are the special case when every firm is assigned workers equal to its minimum quota in the status quo. The status quo provides a different justification for why some blocking pairs are allowable but others are not, one which we think applies well to school choice.

Finally, the paper closest in spirit to ours is Abdulkadiroğlu and Sönmez (1999), "House Allocation with Existing Tenants." Their model is one-sided, and they show that a hybrid of the Serial Dictatorship algorithm and the TTC algorithm provides an efficient improvement over the status quo. We present a two-sided model with a hybrid algorithm between the DA and the TTC. Although our models are different, our approach is remarkably similar to theirs.

2 Model

In this section we present the one-to-one matching model. In section 7 we discuss extending the theory to the many-to-one model of school choice.

The set of agents A is partitioned into the set of workers W and the set of firms F. For clarity of exposition we use masculine pronouns for workers and feminine pronouns for firms. Every worker $w \in W$ has an linear preference order $\succsim_w$ over $F \cup \{w\}$, and every $f \in F$ has an associated linear preference order $\succsim_f$ over $W \cup \{f\}$ (no indifferences are allowed).[3] A preference for oneself is a preference to be unmatched: if a prefers a to b this means that a prefers to remain unmatched than to match to b.

A *match* is a function $\mu : A \to A$ such that the following hold:

1. if $w \in W$ then $\mu(w) \in F \cup \{w\}$; and
2. if $f \in F$ then $\mu(f) \in W \cup \{f\}$; and
3. $\mu(\mu(a)) = a$.

[3] We considered introducing indifferences in the model to better fit the school choice application, combined with optimization techniques from Erdil and Ergin (2008), to find a "student optimal assignment." However, as noted by Balbuzanov and Kotowski (2019), in this setting the weak core is too big and the strong core is empty. Adapting the exclusion core seems to our setting resolves this issue, but at the cost of clarity and with little gain: even with an appropriate definition of the core, we have not found an efficient way to adapt Erdil and Ergin (2008) to our model.

The first two require that the match is two-sided: every worker matches to a firm (or is unmatched) and every firm matches to a worker (or is unmatched). The third requires that every agent is matched to the agent matched to him or her. If $\mu(a) = a$ then a is *μ-unmatched*; otherwise, $\mu(a)$ is the *μ-partner* of a. We write $\mu \succsim_X \mu'$ to mean $\mu(x) \succsim_x \mu'(x)$ for all $x \in X$.

There is a *status quo* match μ_0. The status quo match limits the set of matches we consider to the set of matches we consider to those satisfying the following:

Definition 1. Match μ is *individually rational* if $\mu \succsim_A \mu_0$.

The interpretation is that if agent a prefers $\mu_0(a)$ to $\mu(a)$, then a can demand the status quo match μ_0.

A *coalition* $C \subseteq A$ is a subset of agents who may form a match among themselves. Let $\mu(C) \equiv \{\mu(a) : a \in C\}$. Note that if $\mu(C) \subseteq C$, then $\mu(C) = C$. If a coalition weakly prefers a match μ' to μ and μ' only matches agents in C to agents in C, then C may block μ; formally,

Definition 2. Coalition C *blocks* μ *through* μ' if $\mu' \succsim_C \mu$, $\mu'(a) \succ_a \mu(a)$ for at least one $a \in C$, and $\mu'(C) = C$.

The *core* is the set of all individually rational matches not blocked by any coalition through any match.[4]

3 A Motivating Example

The following example illustrates the main idea of the paper. It demonstrates that the intersection of the core and the individually rational matches may be empty but that certain selections from the set of individually rational matches are more appealing than others.

Example 1 (Regretted Contracts:)**.** Two cities A and B are seeking contractors for construction projects, an *apartment complex* for city A and a *bridge* for city B. There are two qualified contractors 1 and 2, both of whom have signed contracts with cities A and B, respectively. In the language of our model, these contracts are the status quo match, the set of contracts to which any agent can appeal, denoted μ_0. Contracts may be canceled only by the mutual agreement of the parties involved. Therefore, any outcome must guarantee that all agents are weakly better off than under the status quo match.

[4] Formally, this is the *strong core* because we consider all *weak* blocks (allowing some coalition members to be indifferent between μ' and μ). In two-sided matching without indifferences all weak blocks are strong blocks. Because the coalitions we will consider later will usually contain agents who do not change partners, we use the strong core as it is smaller.

1	2
$\tilde{1}$	2^*
$\boldsymbol{A}^*$	$\tilde{A}$
$\vdots$	$\boldsymbol{B}$

A	B
$\tilde{2}$	$\tilde{B}^*$
$\mathbf{1}^*$	$\mathbf{2}$
$\vdots$	$\vdots$

Figure 1: We depict the preferences in Regretted Contracts. The header of the column is the agent and the column lists preferences from most to least preferred. The entries in bold indicate the status quo match. The $*$ and $\sim$ indicate μ^* and $\tilde{\mu}$, respectively. For example, the preferences in the 2 column are read as $2 \succ_2 A \succ_2 B$, and $\mu^*(2) = 2$ and $\tilde{\mu}(2) = A$.

Both contractors 1 and 2 fall upon hard times. Contractor 1's chief apartment engineer is now employef by contractor 2, and contractor 2's chief bridge engineer has resigned. Contractor 1 desires no contract and contractor 2 would rather have no contract to the contract with A to the contract with B.[5] City A still wants the apartment building built and prefers contractor 2 now, but city B would rather not have any contracts now that the bridge engineer has resigned. The preferences are summarized in figure 1, with the status quo match in bold and the two Pareto efficient matches that improve on the status quo match are marked by the $\sim$ and superscript $*$. Consider the core. Because 1 and 2 both prefer to be unmatched, any match in the core must leave both 1 and 2 unmatched (otherwise, the coalition of $\{1\}$ or $\{2\}$ blocks the match). However, this implies that A is unmatched, a violation of the status quo match μ_0. It follows that no match in the core is a (weak) Pareto improvement of the status quo match and so the standard cooperative game theoretic solution provides no insights.

The failure of the core to provide a match that is individually rational arises from the blocking coalitions allowed. Allowing every subset of agents to block is too permissive and ignores the status quo match μ_0. Indeed, the core is usually justified by arguing that agents in the blocking coalition could form contracts among only themselves, which allows for coalitions such as $\{1\}$ or $\{B\}$. For example, consider match μ^*. The coalition $\{1\}$ blocks μ^*, but would city A willingly cancel her contract with contractor 1? City A's approval depends on whom she matches with after block. As

[5]We assure the concerned reader that it is not critical to our example that some contracts fail the standard notion of individual rationality (that is, both contractors 1 and 2 would rather be unmatched than fulfill their obligations). A larger example would demonstrate the same concepts at the cost of expository conciseness.

contractor 2 does not prefer city A to his match under μ^*, contractor 1 cannot secure a (weakly) preferred partner for city A. No coalition C such that $\mu_0(C) = C$ blocks μ^*.

In contrast to μ^*, consider $\tilde{\mu}$. Contractor 2 would like to contract with himself, but needs the approval of city B. To guarantee that city B weakly prefers the blocking match to $\tilde{\mu}$, contractor 2 must include both city B in the blocking coalition. Because $\mu_0(\{2, B\}) = \{2, B\}$, no other agent needs to approve of the block.

Our solution, the *agreeable core*, requires that every agent in a blocking coalition guarantees his or her partner in the status quo match a partner he or she weakly prefers to the current match. Equivalently, we do not consider coalitions that contain an agent but not his or her partner in the status quo match. Such coalitions are *agreeable* because agents in a status quo match mutually agree to the coalition. In this example, the coalition $\{2, B\}$ is agreeable and blocks $\tilde{\mu}$. No agreeable coalition blocks μ^*, and hence μ^* is the (unique) element in the agreeable core.

4 The Agreeable Core

Example 1 demonstrates that the core may be empty. The nonexistence of a match that is both individually rational and unblocked by every coalition of agents motivates restricting either the matches a coalition can block through or the coalitions considered. The choice is nontrivial and hinges upon the interpretation of the status quo match.

If the matches that a coalition can block through are restricted, then the natural requirement is that any coalition can block but only through an individually rational match μ. The interpretation is that the status quo match is inviolable *ex post*. In order to block a match, a coalition needs only to suggest an individually rational match; as long as all agents are weakly better off than at μ_0, no agent can complain about his or her partner. In example 1, however, no individually rational match is unblocked in this sense: we showed that $\tilde{\mu}$ is blocked by coalition $\{2, 3, B\}$ with the individually rational match μ^*, while μ^* is blocked by coalition $\{1\}$ with the individually rational match $\tilde{\mu}$. Hence, this notion may be empty.

The alternative is to restrict the set of coalitions but not matches they can block through. The interpretation is that the status quo match is not only inviolable *ex post* but also that any new contract formed by an agent requires the *ex ante* approval of his or her μ_0-partner. We consider only coalitions meeting the following criterion:

Definition 3. A coalition C is *agreeable* if $\mu_0(C) = C$.

A coalition C is agreeable if any contract in μ_0 does not contain both an agent in C and an agent not in C. By restricting our attention to agreeable coalitions, we require that every agent in a blocking coalition of μ guarantees his or her μ_0-partner a weakly better partner at match μ' than at μ. To guarantee such an improvement, the μ_0-partner's partner at μ' must also be included in the coalition, which implies that the μ_0-partner's μ'-partner must also be included in the coalition, and so on. Definition 4 formalizes this idea.

Definition 4. The *agreeable core* is the set of individually rational matches not blocked by any agreeable coalition.

The agreeable core puts a strong requirement on blocking coalitions: every agent in the coalition *and their μ_0-partners* must be made weakly better off. Our interpretation is that if some agent a is harmed by a block and his or her μ_0-partner is in the blocking coalition, then a can veto the block by refusing to cancel the status quo contract. The important nuance is that the harmed agent can veto μ' even if he or she prefers μ' to μ_0.

The veto power inherent in the agreeable core allows one member of a status quo match to dictate the matches his or her partner can form. The picture to have in mind is both agents in a status quo match simultaneously searching for better matches. They both agree to cancel their status quo match *simultaneous to both confirming new partners.* Because the match of one partner influences who is willing to match with the other, both must agree not only to cancel their status quo match but also approve of the other's new match. By only considering agreeable coalitions, we allow agents to veto a blocking coalition before the coalition acts.

We find the following justification for the agreeable core helpful in explaining the agreeable core and how we allow agents veto blocking coalitions *ex ante*. For a given status quo μ_0, agents are considering forming the individually rational match μ. Before μ is realized among the agents (say, before the agents cancel their status quo agreements and form the μ agreements), a coalition considers enforcing some match μ' among themselves. If some agent a is in the coalition but $\mu_0(a)$ is not in the coalition, then $\mu_0(a)$ may refuse to permit a to form μ' unless $\mu_0(a)$ is certain he or she will prefer μ' to μ. Hence, $\mu_0(\mu)$ must also be in the coalition.

Perhaps surprisingly, the set of matches not blocked by any agreeable coalition is not a subset of the individually rational matches. Our definition of blocking coalition does not allow an agent to demand μ_0, and hence the restriction to individually rational matches is substantive. For a simple example, restrict example 1 to just contractor 1 and city A. The match $\mu(1) = 1$ and $\mu(A) = A$ is not blocked by any coalition but does not Pareto improve μ_0.

We devote section 5 to developing the machinery to prove our main result, namely, that the agreeable core is never empty. In the remainder of this section we briefly touch on several aspects of the agreeable core that do not require our more involved techniques. Section 4.1 shows that the agreeable core is always Pareto efficient, and conversely if μ_0 is Pareto efficient then $\{\mu_0\}$ is the agreeable core. As alluded to in the introduction, our model features several connections with both the classical model of stability (Gale and Shapley, 1962) and more recent models of reassignment (Combe et al., 2022; Pereyra, 2013). In section 4.2 and section 4.3 we develop these connections; as an expository device and a prelude to our algorithm, we highlight the two leading algorithms in two-sided matching – the Deferred Acceptance (DA) and the Top Trading Cycles (TTC) algorithms – and their adaptations used in the literature to guarantee individual rationality.

4.1 Efficiency

In this subsection we investigate the efficiency of the agreeable core. Our first observation is that every match in the agreeable core is not Pareto dominated:[6] if ν Pareto dominates μ, then A (which is always agreeable) blocks μ through ν. Our second observations is a kind of converse: if μ_0 is not Pareto dominated, then μ_0 is in the agreeable core. To see this, suppose (toward a contradiction) that some agreeable coalition C blocks μ_0 through μ. But then because $\mu_0(C) = \mu(C) = C$, we can define μ' that agrees with μ for agents in C and agrees with μ_0 everywhere else. But μ' then Pareto dominates μ_0, a contradiction to the supposition that μ_0 is Pareto efficient.

Remark 1. *Every μ in the agreeable core is Pareto efficient.*[7] *Moreover, μ_0 is Pareto efficient if and only if the μ_0 is the unique element of the agreeable core.*

Remark 1 assures us that the agreeable core satisfies the most common efficiency standard.

4.2 Connection to Stability

Our model parallels the classic theory of stability introduced by Gale and Shapley (1962) except that the classical model does not include a status quo in the primitives. A *blocking pair* of a match is any worker and firm pair such that both prefer each other to their match. A match is *stable* if all agents prefer their match to being unmatched and there are no blocking pairs of the match. It is well-known (Roth and Sotomayor, 1990) that the set of stable matches is the core that we defined previously. Our definition of the agreeable core guarantees that if $\mu_0(a) = a$ for all $a \in A$, then the agreeable core corresponds to the core because every coalition is agreeable. Therefore stability is the special case of the agreeable core when μ_0 leaves all agents unmatched.

Gale and Shapley (1962) gives an efficient algorithm for constructing a stable match: the Deferred Acceptance algorithm (algorithm 1). The DA includes an auxiliary agent $\varnothing$ that indicates if a worker has been matched by the algorithm yet. Initially, the DA assigns every worker to $\varnothing$ and

[6] We say that ν *Pareto dominates* μ if every agent weakly prefers ν to μ and at least one agent strictly prefers ν to μ.

[7] If μ is not Pareto dominated by any ν, then μ is *Pareto efficient*.

every firm to herself. At every step of the DA, every worker matched to $\varnothing$ proposes to the firm he prefers the most among those he has not proposed to yet (if he would rather be unmatched, he is matched to himself and removed). Every firm then reviews the proposals she receives and her current match and rejects all but her most preferred proposal or match. The process continues until no more workers are matched to $\varnothing$.

Algorithm 1 Deferred Acceptance (DA) algorithm

Notation: when we write $\mu(f) \leftarrow w$, we mean that f is matched to w and if another worker w' was matched to f, then w' is matched to $\varnothing$.

set $\mu_1(w) \leftarrow \varnothing$ for all $w \in W$. ▷ every w needs to be matched
set $\mu_1(f) \leftarrow f$ for all $f \in F$. ▷ every f is initially unmatched
while $\mu^{\mathrm{DA}}(w) = \varnothing$ for some $w \in W$ **do** ▷ continue until every w has been matched
  for every $w \in W$ such that $\mu^{\mathrm{DA}}(w) = \varnothing$ **do** ▷ every unmatched w proposes
    w proposes to his most-preferred firm that he has not yet proposed to; if he would rather be unmatched, instead he proposes to himself and set $\mu^{\mathrm{DA}}(w) \leftarrow w$.
  end for
  for every $f \in F$ **do** ▷ every f rejects all but one proposal
    let w be f's most preferred proposal that she prefers to being unmatched (if any); f rejects all other proposals and set $\mu^{\mathrm{DA}}(f) \leftarrow w$;
  end for
end while
return μ^{DA}

Although guaranteed to produce a match unblocked by any coalition, the DA fails to satisfy individual rationality (see Pereyra 2013; Combe et al. 2022). There are two ways in which individual rationality can fail. First, a worker may strictly prefer his μ_0-partner to his match. Pereyra (2013) resolves this issue by requiring that each firm ranks accepts her μ_0-partner if he proposes to her. This modification guarantees that workers find the outcome individually rational because no worker proposes to a less preferred firm without being rejected by his μ_0-partner.

In our setting firms also have individual rationality constraints. The DA fails to accommodate these because a worker makes proposals (and may be matched to another firm) even though his μ_0-firm has not received a proposal she prefers to the worker. We will see in section 5.2 how to resolve this tension by limiting which workers can propose.

4.3 Connection to Reassignment

As referenced in our literature review, a series of recent papers have considered reassignment problems. The main thrust of this literature is to find a match through a strategyproof mechanism that is both individually rational and maximizes some objective function (Combe et al., 2022). Because the agreeable core is motivated with first principles (the core) rather than with an objective in mind (obtaining a strategyproof mechanism), there are substantial differences in definitions and results. However, both approaches employ the same method: the Top Trading Cycles (TTC) algorithm. The TTC finds a match such that no coalition of workers can reallocate their μ_0-firms among themselves and improve their matches. We define the TTC in algorithm 2.

Algorithm 2 Top Trading Cycles (TTC) algorithm

set $\mu^{\mathrm{TTC}}(a) = a$ for all a. ▷ every a starts unmatched
every agent stands up.
while at least one agent is standing **do** ▷ continue until every a is matched
  every standing worker points to his most-preferred of the standing firms.
  every standing firm points to her most preferred of the standing workers.
  choose an arbitrary cycle $(w_1, f_2, \dots w_{2k-1} \equiv w_1, f_{2k} \equiv f_2)$ such that every agent points to the next agent in the cycle and have all agents in the cycle sit down.
  set $\mu^{\mathrm{TTC}}(w_k) \leftarrow f_{k+1}$. ▷ match w in the cycle to the f he points to
end while
return μ^{TTC}

If some agents are matched by μ_0, then the TTC may not be individually rational. To accommodate this, Combe et al. (2022) and Combe (2023) make the following two modifications. First, a firm must point to her μ_0-worker so long as he is standing. This guarantees that $\mu^{\mathrm{TTC}} \succsim_W \mu_0$. Second, no worker may point to a firm if that firm prefers her μ_0-partner to the worker. This guarantees that $\mu^{\mathrm{TTC}} \succsim_W \mu_0$.

In our setting, however, these modifications are not enough. As we saw in section 4.2, the agreeable core equals with the set of stable matches when all agents are μ_0-unmatched. At least in this case firms must be given power to decide between the workers pointing to them, as in the DA. In section section 5.3 we incorporate this by limiting which workers and firms participate in the TTC.

5 A Proof of Existence: The Propose-Exchange Algorithm

In this section we present a computationally efficient and economically meaningful algorithm that always produces a match μ_2 (defined through this section) in the agreeable core. It directly implies that the agreeable core exists and provides some insight into its structure. Our main result is the following:

Theorem 1. *μ_2 is in the agreeable core.*

The proof (and definition of μ_2) occupies the remainder of this section. All omitted proofs are contained in appendix A.

Our algorithm is the *Propose-Exchange* algorithm (PE) and is composed of two phases. The first phase resembles the Deferred Acceptance (DA) algorithm and eliminates any block by a coalition that either includes an agent who is unmatched in the status quo or who becomes unmatched by the block. The second phase resembles the Top Trading Cycles (TTC) algorithm and eliminates all blocks that involve reshuffling status quo partners among themselves. For readers unfamiliar with the DA and the TTC, we refer the reader to our introductions in section 4.2 and section 4.3, respectively.

5.1 A Graph-Theoretic Depiction

Despite our parsimonious definition of the agreeable core, so far testing whether μ is in the agreeable core requires checking whether any coalition can block μ through any μ', which is only feasible in small examples. Our main result from this subsection is a characterization of blocking coalitions in terms of alternating paths in an (edgewise) colored graph, which is computationally efficient. We use the language of graph theory to formalize our ideas, and throughout we denote by μ both the function μ and the set $\{\{a, b\} : \mu(a) = b\}$; if $\mu(a) = a$, then we include the element $\{a\} \in \mu$.

An (edgewise 2-colored) *graph* G is a triplet (V, E_0, E_1) where V is a set of *vertices* and E_0 and E_1 are sets of unordered pairs of vertices called *edges* (possibly including an edge from a vertex to itself, called a *loop*). Both E_0 and E_1 may contain an edge between the same set of nodes; we consider these edges distinct.[8]

We consider graphs where the vertices are agents, the first set of edges is μ_0 (connecting status quo partners) and the second set of edges connects agents who weakly prefer each other to some given match μ. Accordingly, let $I(\mu)$ be the set edges (possibly including loops) between agents that strictly prefer each other to their μ-partners; formally $I(\mu) = \{\{a, b\} : b \succ_a \mu(a) \text{ and } a \succ_b \mu(b)\}$.

[8] This means that G is a *multi*graph.

Our main graph of interest is $(A, \mu_0, \mu \cup I(\mu))$. That is, the vertices are agents, the first set of edges connects status quo partners, and the second set of edges connects all pairs that weakly prefer each other over their μ-partners.

A (simple) *path* in (V, E_0, E_1) is a vector of edges $P = (e_1, \ldots, e_n)$ without repeats such that $e_k \cap e_{k+1} \neq \emptyset$ for $1 \leq k < n$. Recall that an edge may appear twice (in both E_0 and E_1), and so a path may consist of an edge between two nodes twice, once from E_0 and once from E_1. We say a *vertex is in a path* if the path contains an edge that contains the vertex. We sometimes abuse notation and write P for the vertices in P.

A path P is *complete* if every vertex contained in the path is contained in exactly two edges of the path. A path is *alternating* if no pair of consecutive edges are from the same set of edges (E_0 or E_1). For an arbitrary complete and alternating path P in $(A, \mu_0, \mu \cup I(\mu))$, we define μ^P for $a \in P$ such that $\{a, \mu^P(a)\}$ is the edge in P containing a that is from $\mu \cup I(\mu)$. That is, μ^P matches $a \in P$ to the agent whom a shares an edge from $\mu \cup I(\mu)$ in P with and is arbitrary otherwise. By lemma A.2 in the appendix, every agent in P is contained in one edge from μ_0 and one edge is from $\mu \cup I(\mu)$, so μ^P is well defined and $\mu^P(P) = P$.

Our main result of this subsection is that a path that is complete, alternating, and contains an edge from $I(\mu)$ corresponds to an agreeable blocking coalition in $(A, \mu_0, \mu \cup I(\mu))$. We formalize this as follows:

Definition 5. Path P is a *blocking path of* μ if P is a complete and alternating path in $(A, \mu_0, \mu \cup I(\mu))$ that contains at least one edge from $I(\mu)$.

A blocking path of μ is aptly named as it corresponds to a blocking coalition of μ.

Proposition 1. *Match μ is in the agreeable core if and only if μ admits no blocking paths. Moreover, if P is a blocking path of μ then P blocks μ through μ^P.*

Proposition 1 provides a test that is linear in the number of edges to see if μ is in the agreeable core.[9]

In the following subsections we find it useful to partition paths between those that form cycles and those that do not:

Definition 6. Let $P = (e_1, \ldots, e_n)$. If $e_1 \cap e_n = \emptyset$, then P is *cyclic*; otherwise, P is *linear*.

As the name suggests, cyclic paths start with an agent and then return to that agent. In $(A, \mu_0, \mu \cup I(\mu))$, a cyclic, complete, and alternating path corresponds to agents (who are μ_0-matched) trading their μ_0-partners among themselves. Linear paths that are also complete and alternating

[9] A depth first search initiated from every edge in $I(\mu)$ is sufficient.

start with a loop and end with a loop, forming a line in the graph. In $(A, \mu_0, \mu \cup I(\mu))$, a linear, complete, and alternating path corresponds to agents trading their μ_0-firms among themselves, except that two agents are unmatched by one or both sets of edges. The Propose-Exchange algorithm works by first producing a match μ_1 that admits no linear blocking paths, then finding a series of Pareto improvements of μ_1 to produce a match μ_2 that has no cyclic blocking paths.

5.2 The Propose Phase

The first phase of our algorithm outputs a match μ_1 by systematically removing all linear blocking paths from $(C, \mu_0, \mu \cup I(\mu))$. A linear blocking path P in $(A, \mu_0, \mu \cup I(\mu))$ corresponds to a series of trades, but the agents at either end of the path are either μ_0-unmatched or μ^P-unmatched. These may be thought of as a cycle that includes the "unmatched" agent.

The Propose algorithm is a variation of the Deferred Acceptance algorithm (DA). The DA is designed for markets where all agents are unmatched under μ_0 and is defined in algorithm 1. We noted in section 4.2 that the DA may fail individual rationality for both workers and firms. The following modification from Guillen and Kesten (2012) guarantees individual rationality for the workers: if w proposes to $\mu_0(w)$, then $\mu_0(w)$ *must* accept w and reject all other proposals.

However, in our setting firms also have μ_0 as a guarantee. We incorporate this into the Propose algorithm by only allowing a worker to make a proposal once his μ_0-firm has received a more preferred proposal. The Propose algorithm is defined in algorithm 3 and we indicate our addition in *italics*.

Our first observation is that μ_1 is individually rational. If w strictly prefers μ_0 to μ_1, then w would have proposed to μ_0 (and not been rejected). Again, if $\mu_0(f)$ is matched by μ_1 to a firm other than f, then f received a proposal she prefers to $\mu_0(f)$ and hence she prefers μ_1 to μ_0. We then show that at the end of the Propose algorithm, no blocking path of μ_1 is linear.

Lemma 1. *μ_1 admits no linear blocking paths.*

Our proof leverages that a linear blocking path P in $(A, \mu_0, \mu \cup I(\mu))$ can be rewritten to always begin with either a worker who is μ_0-unmatched and hence proposes or a firm who is μ^P-unmatched (and hence her μ_0-worker makes a proposal). Because the start and finish of the path are connected by workers who (weakly) prefer the firm they receive in the block, we can show that every worker in the path must have had the opportunity to propose. We then show that the path must terminate with either a worker who is μ_0-matched or a firm who is μ_0-unmatched, neither of which would reject the proposal made through the path. We conclude by showing that every firm accepts the proposal from her μ^P-partner, which contradicts that $\mu \neq \mu^P$.

Algorithm 3 Propose Phase algorithm

Notation: when we write $\mu(f) \leftarrow w$, we mean that f is matched to w and if another worker w' was matched to f, then w' is matched to $\varnothing$.

set $\mu_1(w) \leftarrow \varnothing$ for all $w \in W$. ▷ every w needs to be matched

set $\mu_1(f) \leftarrow f$ for all $f \in F$. ▷ every f is initially unmatched

if *$w \succ_{\mu_0(w)} \mu_0(w)$* ***then*** *set $\mu_1(w) \leftarrow \mu_0(w)$.* ▷ *$w$ is initially matched to $\mu_0(w)$*

while $\mu_1(w) = \varnothing$ for some $w \in W$ **do** ▷ continue until every w has been matched

  for every $w \in W$ such that $\mu_1(w) = \varnothing$ **do** ▷ every unmatched w proposes

    w proposes to his most-preferred firm that he has not yet proposed to; if he would rather be unmatched, instead he proposes to himself and set $\mu_1(w) \leftarrow w$.

  end for

  for every $f \in F$ **do** ▷ every f rejects all but one proposal

    if f receives a proposal from her μ_0-partner **then** f rejects all other current and future proposals and set $\mu_1(f) \leftarrow \mu_0(f)$. ▷ f accepts $\mu_0(f)$

    else let w be f's most preferred proposal that she prefers to being unmatched *and to her μ_0-worker* (if any); f rejects all other proposals and set $\mu_1(f) \leftarrow w$.

  end for

end while

return μ_1

5.3 The Exchange Phase

In the second phase of the algorithm, we eliminate all cyclic blocking paths. Cyclic blocking paths have special significance in $(A, \mu_0, \mu \cup I(\mu))$. A cyclic blocking path corresponds to workers and their μ_0-firms rearranging their status quo matches among themselves. No agent in a cyclic path is unmatched by either μ or μ_0. A cyclic blocking path represents an inefficient allocation for C: the coalition could have rearranged their status quo matches among themselves and obtained a better match.

The Exchange algorithm is an adaptation of the Top Trading Cycles (TTC) algorithm to find these cycles and remove them. The difficulty with using solely the TTC in our setting is that the TTC does not give firms the ability to select *between* workers. Although firm's preferences limit the set of acceptable workers, which worker is matched to the firm ultimately depends on the worker the firm is required to point at. If only some workers or firms are matched by μ_0, then the firm's lack of choice leads to violations of the agreeable core.

We resolve this by only applying the TTC to workers and firms who did not both find better partners through the Propose algorithm. Recall that the agreeable core gives agents veto rights over their μ_0-partner's potential matches. In our Exchange algorithm we leverage this by uniformly giving workers the power to determine their μ_0-firm's partner, subject to individual rationality constraints. The Exchange algorithm is defined in algorithm 4 and we indicate our addition in *italics*.

Algorithm 4 Exchange Phase algorithm

---

set $\mu_2(a) \leftarrow \mu_1(a)$ *for all* a. ▷ *start with* μ_1

every w *such that* $\mu_1(w) = f$ *stands up with* $\mu_0(w)$ *and all other agents sit down.*

▷ *status quo partners stand up if* μ_1 *did not improve them*

while at least one worker is standing **do** ▷ continue until every a is matched

every standing worker points to his most-preferred of the standing firms who prefer him to her μ_0-worker.

every standing firm points to her μ_0-worker.

choose an arbitrary cycle $(w_1, f_2, \ldots w_{2k-1} \equiv w_1, f_{2k} \equiv f_2)$ such that every agent points to the next agent in the cycle and have all agents in the cycle sit down.

set $\mu_2(w_k) \leftarrow f_{k+1}$. ▷ match w in the cycle to the f he points to

end while

return μ_2

Our first observation is that the Exchange algorithm makes no agents worse off than under μ_1.

Workers only point to firms they prefer to μ_0, and by our simplification of workers' preferences, firms can only be pointed at by workers they prefer to μ_0. The result is that at the end of the Exchange algorithm, μ_2 admits no cyclic blocking paths.

Lemma 2. *μ_2 admits no cyclic blocking paths.*

Our proof leverages that if w strictly prefers f to $\mu_2(w)$, then f must sit down at least one step *before* w. A cyclic blocking path then implies that the firms in the path sit down on average strictly before the workers in the path sit down. However, because every worker's μ_0-firm is in the path and they sit down in the same step, it must be that the firms in the path sit down on average in the same step as the workers in the path sit down. This contradiction rules out cyclic blocking paths.

5.4 Existence

We are now ready to prove our main result.

Proof of theorem 1:

Suppose (toward a contradiction) that μ_2 is not in the agreeable core. Then by proposition 1 the graph $(A, \mu_0, \mu_2 \cup I(\mu_2))$ contains a blocking path P. By lemma 2, P is linear. But P is also blocking path in $(A, \mu_0, \mu_1 \cup I(\mu_1))$ because $\mu_2 \cup I(\mu_2) \subseteq \mu_1 \cup I(\mu_1)$ and $I(\mu_2) \subseteq I(\mu_1)$. By lemma 1, P is not linear. This is a contradiction, which proves the claim. $\square$

The importance of the Propose-Exchange algorithm in our proof cannot be understated. However, the algorithm has practical implications because it is also computationally efficient. The Propose phase runs in polynomial time because each worker can make at most $|F| + 1$ proposals. Similarly, one cycle is removed in every iteration of the Exchange phase, and at most $|F|$ cycles can be removed. An efficient algorithm is necessary for implementing the agreeable core in applications.

6 Additional Results

In this section we address two additional questions, primarily with negative results. First, we consider the incentive properties of the PE algorithm. Second, we study the structure of the agreeable core.

6.1 Manipulability

In many applications (such as school choice), market designers prefer mechanisms that incentivize one side (typically workers) to truthfully report their preferences. A *matching mechanism* maps reported preferences $\succsim$ to a match $\psi(\succsim)$ and it is *manipulable* if there exists a worker[10] w, a preference profile $\succsim$, and a preference $\succsim'_w$ such that the worker prefers $\psi(\succsim_{-w}, \succsim'_w)$ to $\psi(\succsim)$. If the mechanism is not manipulable, then workers do not need to strategize and their reports can be taken at face value for statistical purposes. Neither the DA nor the TTC are manipulable. Because the PE is a combination of the DA and the TTC, one may hope that the PE is also not manipulable by workers. This is not the case, and moreover is a universal difficulty with the agreeable core:

Theorem 2. *If $\psi(\succsim)$ is in the agreeable core for all $\succsim$, then ψ is manipulable.*

6.2 Structure

In this subsection, we discuss the difficulty with importing structural results from the classic literature on stable matches into our framework. Although the set of stable matches has a well-understood structure (which we summarize in the following paragraph), the agreeable core is not as tame. The hurdle in the analysis comes from the Exchange phase. To the best of our knowledge, there are no results from the literature that apply to the agreeable core when every agent is μ_0-matched and prefers their μ_0-match to being unmatched.

We briefly summarize the main structural results on the set of stable matches. First, a *lattice* is a partially ordered set $(L, \geq)$ such that any two elements of L have a unique *least upper bound*, called the *join* of x and y, and a unique *greatest lower bound*, called the *meet* of x and y. That is, there is a unique $x \vee y$ such that if $z \geq x$ and $z \geq y$ then $z \geq x \vee y$, and there is a unique $x \wedge y$ such that if $x \geq z$ and $y \geq x$ then $x \wedge y \geq z$. A key result[11] is that the set of stable matches forms a lattice with the partial order $\succsim_W$. The join of two stable matches is the match that gives every worker w his more preferred partner from $\{\mu(w), \nu(w)\}$ and every f her less preferred partner from $\{\mu(f), \nu(f)\}$; the meet is given symmetrically. This implies that there is a conflict of interest between the workers and the firms: if every worker weakly prefers a stable μ to a stable ν, then every firm weakly prefers ν to μ. Moreover, there is a *worker optimal* stable match and a *firm optimal* stable match.

We show that the agreeable core fails to be a lattice through an example. Let $\mu_0(1) = A$, $\mu_0(2) = B$, and $\mu_0(3) = C$, and preferences are given as in figure 2. Both the spair 2 and B and the pair 3 and C prefer to participate in a cycle with the pair 1 and A, but 1 and A have opposing preferences over the two possible cycles. Worker 1 prefers firm C and firm A prefers worker 2, and so either cycle may be in the agreeable core. The agreeable core consists uniquely of the $*$ match

[10] We only consider one side of the market (workers) because these kinds of mechanisms are typically always manipulable by at least one side of the market.

[11] Donald Knuth attributes this to John H. Conway.

1	2	3
C^*	$\tilde{A}$	A^*
$\tilde{B}$	$\boldsymbol{B}^*$	$\tilde{\boldsymbol{C}}$
$\boldsymbol{A}$	⋮	⋮
⋮	⋮	⋮

A	B	C
$\tilde{2}$	$\tilde{1}$	1^*
3^*	$\mathbf{2}^*$	$\tilde{\mathbf{3}}$
$\mathbf{1}$	⋮	⋮
⋮	⋮	⋮

Figure 2: μ_0 is given in bold. In the $*$ match workers 1 and 2 swap firms A and B, while in the $\sim$ match workers 1 and 3 swap firms A and C.

and the $\sim$ match, a pair which is not ordered by $\succsim_W$. In this example there is no worker optimal match.

Despite the impossibility of recovering a complete lattice over the agreeable core as in the classic model of stability, we show that a (decidedly narrower) result continues to hold. Given that the lattice structure failed in the example because two competing cycles exist in the agreeable core, an astute reader may conjecture that the lattice structure continues to hold for workers and firms who do not lie in such cycles. Suggestively, say that a is a *free agent* in μ if a lies on a linear, complete, and alternating path of (A, μ_0, μ). Our first lemma justifies our terminology:

Lemma 3. *If μ is in the agreeable core, then there are no blocking pairs among free agents in μ. Moreover, every free agent a in μ weakly prefers $\mu(a)$ to being unmatched.*

The proof of lemma 3 shows that these agents are "free" to form blocking pairs because each can satisfy a sequence formed by alternating edges from μ_0 and μ. Free agents resemble the agents in the classic model.

However, an obstacle arises because the free agents depend on μ; that is, a may be a free agent in μ but not in ν. What we can show is that, if μ and ν share the same set of free agents and they agree on the agents who are not free, then $\mu \vee \nu$ is in the agreeable core. Toward that end, we say that μ and ν are *structurally similar* if they have the same set of free agents and $\mu(a) = \nu(a)$ for every agent which is not free. Our second lemma shows that structurally similar matches in the agreeable core play nicely with the join and meet operators defined previously:

Lemma 4. *Let μ and ν be structurally similar matches in the agreeable core. Then $\mu \vee \nu$ is a match. The same holds for $\mu \wedge \nu$.*

Notably, $\mu \vee \nu$ may not be structurally similar to μ and ν.[12] The (possible) structural differences between $\mu \vee \nu$ and μ force us to discard any hope of obtaining a lattice-like result. However, the join and meet operators still produce matches in the agreeable core:

Theorem 3. *Let μ and ν be structurally similar matches in the agreeable core. Then $\mu \vee \nu$ and $\mu \wedge \nu$ are both in the agreeable core.*

The conflict of interest continues to hold for structurally similar matches. That is, if μ and ν are in the agreeable core *and are structurally similar*, then if every worker weakly prefers μ to ν, then every firm weakly prefers ν to μ.

Returning to figure 2 we see that no agents at either match are free agents. Conversely, in the classic matching framework, $\mu_0(a) = a$ for every agent and thus every agent is free. Every match is then structurally similar and hence our theorem 3 generalizes standard results.

7 Discussion

We close our exposition by discussing the nuances of applying the agreeable core in several applications. To simplify the exposition we limited our theory to one-to-one matching. Most applications, however, are many-to-one. One standard way to extend our theory to many-to-one matching in the context of school choice is to duplicate schools according to their capacities and then to break ties in students' preferences in an arbitrary manner. To incorporate the status quo, students assigned to a school by μ_0 must be assigned to a particular duplicate of the school. However, care must be taken when giving students arbitrary preferences over the duplicates in the operation of the Propose phase. The trade-offs between different tie-breaking rules is an open question and the direction we intend to take this theory next.

In applications to school matching, the PE algorithm enables market designers to blend "non-tradable" and "tradable" priorities while incorporating minimum quotas. A student i has a non-tradable priority at school s if s ranks i above every student, $\mu_0(i) = i$, and $\mu_0(s) = s$; i can always attend s, but can never trade his priority at s. A i has a tradable priority if $\mu_0(i) = s$ and s prefers to be unmatched rather than matching with i; i can always choose to attend s because $\mu_0(i) = s$, is never forced to attend s because $s \succsim_s i$, and can trade s during the Exchange phase if i does not find a better partner in the Propose phase. Likewise, a minimum quota can be enforced by setting $\mu_0(i) = s$ and placing i just above being unmatched in s's preference; i will only b required to attend s if there is no way to rearrange students (without harming any student) and still meet

[12] We have an example demonstrating this (available upon request), but it is too lengthy to include because it involves eight workers and eight firms.

the schools minimum quotas.[13]

This paper provides the Propose-Exchange algorithm for finding elements in the agreeable core. Our methods have not allowed us to recover other elements in the agreeable core.[14] The main difficulty is analyzing the agents assigned in the Exchange phase, as seen in section 6.2. We hope that our exposition inspires researchers to find alternative means for finding elements of the agreeable core and to characterize its structure. In school matching, finding a "student optimal" match is a priority, and future work in this direction may be fruitful.

In the NRMP application with early offers, an obvious question is whether taking an early offer (adding a match in μ_0) will negatively impact one or both parties of the contract. The choice of status quo can have unusual effects on the agreeable core. One odd observation is that, even if a worker and firm both prefer to be unmatched rather than be matched to each other, a status quo match between them affects the agreeable core by further restricting the set of agreeable coalitions. Perhaps counter-intuitively, adding a status quo match between a worker and firm may increase the agreeable core by adding elements that both agents strictly prefer to another match in the agreeable core. Understanding how the agreeable core change with respect to the status quo is an important next step for applying this theory to the NRMP and other markets for entry-level professionals.

A Omitted Proofs

Lemma A.1. *μ is individually rational if and only if μ does not admit a blocking path of length 2.*

Proof. For the ($\Rightarrow$) direction: We prove the contrapositive; that is, if μ admits a blocking path of length 2, then μ is not individually rational. Let (e_1, e_2) be a blocking path of μ. Let $e_1 \in \mu_0$ and $e_2 \in I(\mu)$; the argument is symmetric if the order is reversed. Observe that $a \in e_1$ strictly prefers $\mu^P(a)$ to $\mu(a)$. Because $\mu^P(a) = \mu_0(a)$, it follows that a strictly prefers μ_0 to μ. Therefore, μ is not individually rational.

For the ($\Leftarrow$) direction: We prove the contrapositive; that is, if μ is not individually rational, then μ admits a blocking path of length 2. Let μ not be individually rational. Then there is an agent a such that $\mu_0(a) \succ_a \mu(a)$. Then $\{a, \mu_0(a)\} \in I(\mu)$. Therefore $(\{a, \mu_0(a)\}, \{a, \mu_0(a)\})$ with the first edge from μ_0 and the second edge from $I(\mu)$ is a blocking path of length 2 of μ. $\square$

[13]Because schools cannot have indifferences in preferences, this statement applies only to the PE algorithm and not more generally to the agreeable core.

[14]Beyond those obtained by reversing the role of workers and firms in either or both steps of the Propose and Exchange phases.

Lemma A.2. *If P is a complete and alternating path in $(A, \mu_0, \mu \cup I(\mu))$, then every agent contained in P is in exactly one edge from μ_0 and one edge from $\mu \cup I(\mu)$.*

Proof. Let $P = (e_1, \ldots, e_n)$ be a complete and alternating path in $(A, \mu_0, \mu \cup I(\mu))$ and let a be contained in P. If $n = 2$, then the statement is trivial because completeness implies every $a \in P$ is in both e_1 and e_2 and P alternating implies that one of $\{e_1, e_2\}$ is in μ_0 and the other is in $\mu \cup I(\mu)$. Hence, let $n \geq 3$.

Again, if $a \in e_k \cap e_{k+1}$ for $k \geq 1$ then the statement is trivial because completeness implies e_k and e_{k+1} are the only edges in P containing a and P alternating implies that one of $\{e_k, e_{k+1}\}$ is in μ_0 and the other is in $\mu \cup I(\mu)$. Hence, let $a \in e_1 \cap e_n$ and thus P is cyclic. Let a be a worker; the argument is symmetric if a is a firm.

Because there is a bijection[15] between the workers and firms contained in P and every agent in P is contained in two edges of P, n is even. Therefore, if $e_1 \in \mu_0$ then $e_n \in \mu \cup I(\mu)$, and if $e_1 \in \mu \cup I(\mu)$ then $e_n \in \mu_0$. This proves the result. $\square$

Proof of proposition 1:

For the $(\Rightarrow)$ direction: We prove the contrapositive; that is, if μ admits a blocking path, then μ is not in the agreeable core. Let $P = (e_1, \ldots, e_n)$ be a blocking path in $(A, \mu_0, \mu \cup I(\mu))$. Note that $\mu_0(P) = P$ and $\mu^P(P) = P$.

By the definition of $I(\mu)$, it follows that $\mu^P \succsim_P \mu$. Because P is blocking, there is an edge e in P that is also in $I(\mu)$. Hence, both agents in e strictly prefer μ^P to μ. Therefore, P is an agreeable blocking coalition and μ is not in the agreeable core.

For the $(\Leftarrow)$ direction: We prove the contrapositive; that is, if μ is not in the agreeable core then μ admits a blocking path. Let μ be not in the agreeable core. Then either μ is not individually rational, or there exists an agreeable blocking coalition C that blocks μ through μ'. If μ is not individually rational, then by lemma A.1 μ admits a blocking path. We consider the case when there exists an agreeable blocking coalition C that blocks μ through μ'.

Let a_1 be an agent in C such that $\mu'(a_1) \succ_{a_1} \mu(a_1)$; such an agent exists by the definition of a blocking coalition. We will construct a path P from a_1 by iteratively adding alternating edges from μ_0 and μ' to $\{a_1, \mu'(a_1)\}$, first with increasing indices and then with decreasing indices.

Starting $e_1 \equiv \{a_1, \mu'(a_1)\}$ and $P_1 \equiv (e_1)$, do the following iteratively. Choose an edge e_{k+1} from μ_0 or μ' that is not already present in P_k such that $e_k \cap e_{k+1} \neq \emptyset$, then define P_{k+1} by appending e_{k+1} to P_k. Continue until no more edges may be added in this way. Finally, repeat the same process starting from e_1, but *prepending* edges $e_0, e_{-1}, \ldots$ to P_k.

[15] namely, μ_0

Observe that P is a path in $(A, \mu_0, \mu \cup I(\mu))$ because $\mu' \succsim_C \mu$. Next, observe that because every agent in P is contained in at most two edges ($\{a, \mu_0(a)\}$ and $\{a, \mu'(a)\}$); every agent in P is contained in at least two edges because edges are added until no more can be added without including repeats and therefore P is complete. Also, P is alternating because $e_{2k} \in \mu_0$ and $e_{2k-1} \in \mu'$. Finally, observe that $e_1 \in I(\mu)$. Therefore, P is a blocking path of μ. Therefore $(A, \mu_0, \mu \cup I(\mu))$ contains a blocking path, completing the proof. □

Introduction to the proof of lemma 1:

Before proving lemma 1, we first introduce some notation and a short result: We say that loop $e = \{a\}$ is a *proposal source* if either

1(a) : $a \in \mu_0$ and $a \in W$, or

1(b) : $a \notin \mu_0$ and $a \in F$.

We say that loop $e = \{a\}$ is a *proposal sink* if e in not a proposal souce; that is, if either

2(a) : $a \notin \mu_0$ and $a \in W$ or

2(b) : $a \in \mu_0$ and $a \in F$.

A straightforward parity argument shows that if $P = (e_1, \ldots, e_n)$ is a complete, alternating, and linear path in $(A, \mu_0, \mu \cup I(\mu))$, then one of $\{e_1, e_n\}$ is a proposal source and the other is a proposal sink and vice versa.

Lemma A.3. *Let $P = (e_1, \ldots, e_n)$ be a complete, alternating, and linear path in $(A, \mu_0, \mu \cup I(\mu))$ with $n \geq 3$. Then one of $\{e_1, e_n\}$ is a proposal source and the other is a proposal sink.*

Proof. Because P is linear and complete, e_1 and e_n are both loops. Hence e_1 and e_n are (individually) either proposal sinks or proposal sources (but not both). Let $e_1 = \{a_1\}$ and $e_n = \{a_{n-1}\}$. The remainder of the proof shows that if one is a proposal source then the other is a proposal sink.

Suppose e_1 is a proposal source. Note that because μ_0 and $\mu \cup I(\mu)$ are both bipartite and P is complete and alternating, except for the loops that open and close P all edges must alternate between worker to firm and firm to worker. That is, if write $e_k = \{a_{k-1}, a_k\}$ for $1 < k < n$, then the agent with the odd index is from the same set (workers or firms) as a_1 and the agent with the even index is from the other.

There are two cases:

1. *n is even:* There are two subcases:

   (a) *$e_1 \in \mu_0$ and $a_1 \in W$:* Then $e_n \notin \mu_0$. Then $a_{n-1} \in W$. Thus e_n is a proposal sink.

   (b) $e_1 \in \mu$ *and* $a_1 \in F$*:* Then $e_n \in \mu_0$. Then $a_{n-1} \in F$. Thus e_n is a proposal sink.

2. *n is odd:* There are two subcases:

   (a) $e_1 \in \mu_0(a_1)$ *and* $a_1 \in W$*:* Then $e_n \in \mu_0$. Then $a_{n-1} \in F$. Thus e_n is a proposal sink.

   (b) $e_1 \in \mu(a_1)$ *and* $a_1 \in F$*:* Then $e_n \notin \mu_0$. Then $a_{n-1} \in W$. Thus e_n is a proposal sink.

A symmetric argument shows that if a_n is a proposal source, then a_1 is a proposal sink. This proves the claim. □

Proof of lemma 1:

Suppose (toward a contradiction) that $P = (e_1, \ldots, e_n)$ is linear blocking path of μ_1. Because P is linear and complete, e_1 and e_n are both loops. By construction μ_1 is individually rational and hence $n \geq 3$ (lemma A.1). By lemma A.3 we relabel P such that e_1 is a proposal source and e_n is a proposal sink. Let

$$\begin{aligned} e_1 &= \{a_1\} \\ e_2 &= \{a_1, a_2\} \\ &\vdots \\ e_{n-1} &= \{a_{n-2}, a_{n-1}\} \\ e_n &= \{a_{n-1}\} \end{aligned}$$

By our relabeling of P, it follows that if $\{a_k, a_{k+1}\} \in \mu^P$, then $a_k \in W$.

We argue by induction that every worker $a_k \in P$ makes a proposal during the Propose algorithm. Because every agent contained in P weakly prefers μ^P to μ_1, it follows that every worker contained in P *who proposes* proposed to his μ^P-partner. In our base case we show that the worker with the lowest index contained in P proposes during the Propose algorithm. There are two possibilities:

1. *a_1 is a worker:* Because e_1 is a proposal source by definition $\mu_0(a_1) = a_1$. Hence a_1 begins the Propose algorithm matched to $\varnothing$. Therefore, a_1 proposes during the Propose algorithm.

2. *a_1 is a firm:* Because e_1 is a proposal source, by definition $\mu_0(a_1) \neq a_1$. Therefore $\mu_0(a_1) = a_2$. Because a_1 prefers μ^P to μ_0 and $\mu^P(a_1) = a_1$ because e_1 is loop, it follows that a_2 is matched to $\varnothing$ at the start of the Propose algorithm. Therefore, a_2 proposes during the Propose algorithm.

For the inductive step, suppose $a_{k-1} \in W$ makes a proposal; we will show that the worker with the next highest index makes a proposal. If $k - 1 \geq n - 2$, then a_{k-1} is the worker with the highest index and the claim is vacuous; therefore, suppose $k-1 < n-2$. Because $\mu^P(a_{k-1}) = a_k$, it follows that a_{k-1} proposes at some point to a_k. Because μ_1 is individually rational and $\mu_0(a_k) = a_{k+1}$, it follows that a_k weakly prefers a_{k-1} to a_{k+1}. Therefore a_{k+1} is matched to $\varnothing$ at some point and thus a_{k+1} makes at least one proposal during the Propose algorithm, concluding our inductive argument.

Next, we show that an agent contained in a proposal sink never rejects a proposal from their μ^P-partner. If a_{n-1} is a worker, then he never rejects a proposal from himself. If a_{n-1} is a firm, then $\mu_0(a_{n-1}) = a_{n-1}$ by definition. Because a_{n-1} prefers μ^P to both μ_0 and μ_1 and because a_{n-1} receives no proposals she prefers to $\mu_1(a_{n-1})$ (by construction of μ_1), it follows that a_{n-1} does not reject a proposal from $\mu^P(a_{n-1})$.

Finally, we show that no worker contained in P is rejected by his μ^P-partner. To see this, suppose (toward a contradiction) that $k-1$ is the largest index such that a_{k-1} is rejected by $\mu^P(a_{k-1})$. Because a proposal sink does not reject a proposal by his or her μ^P-partner, it follows that $k-1 < n-2$ (that is, a_{k-1} is not one of the last two agents in the path).

Because a_k prefers a_{k-1} to $\mu_1(a_k)$ and yet a_k rejects a_{k-1}, it must be that $\mu_0(a_k) = \mu_1(a_k)$ (by construction of μ_1). Therefore a_k is matched to a_{k+1} by both μ_0 *and* μ_1. Because matches are bijective, we have $\mu_1(a_{k+1}) = \mu_0(a_{k+1}) = a_k$. Consider that, because P is a complete and $n \geq 3$, it follows that $\mu^P(a_{k+1}) \neq \mu_1(a_{k+1})$. Therefore a_{k+1} must be rejected by $\mu^P(a_{k+1})$, a contradiction to our supposition that $k-1$ is the largest index for which a worker is rejected by his μ^P-match.

Therefore, because no worker in P is rejected by his μ^P-partner, it follows that μ^P agrees with μ_1 on P. Hence, every edge in P from $\mu_1 \cup I(\mu_1)$ is from μ_1. But because P is a blocking path, it must contain an edge from $I(\mu_1)$. Because $\mu \cap I(\mu_1) = \emptyset$, this is a contradiction. Therefore no blocking path of μ_1 is linear. $\square$

Proof of lemma 2

Suppose (toward a contradiction) $P = (e_1, \ldots e_n)$ is a cyclic blocking path in $(A, \mu_0, \mu_2 \cup I(\mu_2))$. Because there is a bijection[16] between the workers and firms contained in P, n is even. Define $m \equiv \frac{n}{2}$.

From P (after a possible relabeling) define a vector of agents $(a_1, a_2, \ldots, a_n \equiv a_0)$ such that $\{a_{k-1}, a_k\} = e_{k-1}$, $a_1 \in W$, and $e_1 \in I(\mu_2)$. Because P is alternating, every odd agent is a worker and every even agent is a firm.

We first show that every agent in P stands up in the Exchange phase. To see this, suppose (toward a contradiction) that some worker a_k in P does not stand up during the Exchange phase. Then a_k makes a proposal during the Propose phase to a_{k+1}. Therefore, a_{k+2} makes a proposal

[16] namely, μ_0

during the Propose phase. We can iterate this argument to show that every worker in P makes a proposal during the Propose phase. Because P is a blocking path, each firm in P prefers her respective proposal to her μ_1-partner. Because a_{k-2} is rejected by a_{k-1}, it necessarily follows that $\mu_1(a_k) = a_{k-1}$. Therefore, a_k stands up in the Exchange phase, a contradiction. Therefore, every agent in P stands up during the Propose phase.

Let t_k be the iteration of the **while** ... **do** loop of the Exchange algorithm that a_k sits down in.[17] During the Exchange algorithm every worker a_{2k-1} points to firm a_{2k}; hence, firm a_{2k} sits down weakly earlier than worker a_{2k-1}. In symbols, $t_{2k-1} \geq t_{2k}$ for all $1 \leq k \leq m$. Because $e_1 \in I(\mu_2)$, it follows that $t_1 > t_2$. Therefore,

$$\sum_{k=1}^{m} t_{2k-1} > \sum_{k=1}^{m} t_{2k}$$

However, every worker a_{2k+1} sits down at the same time firm a_{2k} sits down. In symbols, $t_{2k+1} = t_{2k}$ for all $1 \leq k \leq m$. Therefore,

$$\sum_{k=1}^{m} t_{2k+1} = \sum_{k=1}^{m} t_{2k}$$

Because $\sum_{k=1}^{m} t_{2k+1} = \sum_{k=1}^{m} t_{2k-1}$, we reach a contradiction. $\square$

Proof of theorem 2:

Consider the following counterexample. There are three workers denoted by the numbers 1, 2, and 9, and three firms denoted by the letters A, B, and Z. Workers 1 and 2 are reference matched to A and B, respectively, while worker 9 and firm Z are each reference matched to him or herself. Formally:

$$\mu_0(1) = A \qquad \mu_0(2) = B \qquad \mu_0(9) = 9$$
$$\mu_0(A) = 1 \qquad \mu_0(B) = 2 \qquad \mu_0(Z) = Z$$

A profile of preferences $\succ$ and an alternate profile of worker preferences are given in figure 3. We use the circles to indicate match μ°, the squares to indicate match $\mu^{\blacktriangle}$, the $\hat{}$ to indicate $\hat{\mu}$, and

[17] That is, if a_k sits down on the fourth iteration of the while loop, then $t_k = 4$.

P_1

$\succ_1$	$\succ_2$	$\succ_9$
B°	Z	B
Z	A°	Z°
A	B	$\vdots$
$\vdots$	$\vdots$	

P_2

$\succ_1'$	$\succ_2$	$\succ_9$
B°	$Z^\blacktriangle$	$B^\blacktriangle$
Z	A°	Z°
$A^\blacktriangle$	B	$\vdots$
$\vdots$	$\vdots$	

P_3

$\succ_1'$	$\succ_2'$	$\succ_9$
B	$Z^\blacktriangle$	$B^\blacktriangle$
Z	A	$\tilde{Z}$
$\tilde{A}^\blacktriangle$	$\tilde{B}$	$\vdots$
$\vdots$	$\vdots$	

P_4

$\succ_1'$	$\succ_2'$	$\succ_9'$
B	$Z^\blacktriangle$	$B^\blacktriangle$
Z	A	Z
$A^\blacktriangle$	B	$\vdots$
$\vdots$	$\vdots$	

$\succ_A$	$\succ_B$	$\succ_Z$
A	$9^\blacktriangle$	$\tilde{9}^\circ$
2°	1°	1
$\tilde{1}^\blacktriangle$	$\tilde{2}$	$2^\blacktriangle$
$\vdots$	$\vdots$	$\vdots$

Figure 3: Tables provide preferences $\succ$ and alternate worker preferences $\succ'$. A grayed-out firm in $\succ'$ indicates that the worker matching to himself more than to that firm.

$\tilde{}$ to indicate $\tilde{\mu}$.

$$\mu^\circ(1) = B \qquad \mu^\circ(2) = A \qquad \mu^\circ(9) = Z$$
$$\mu^\circ(A) = 2 \qquad \mu^\circ(B) = 1 \qquad \mu^\circ(Z) = 9$$

$$\mu^\blacktriangle(1) = A \qquad \mu^\blacktriangle(2) = Z \qquad \mu^\blacktriangle(9) = B$$
$$\mu^\blacktriangle(A) = 1 \qquad \mu^\blacktriangle(B) = 9 \qquad \mu^\blacktriangle(Z) = 2$$

$$\tilde{\mu}(1) = A \qquad \tilde{\mu}(2) = B \qquad \tilde{\mu}(9) = Z$$
$$\tilde{\mu}(A) = 1 \qquad \tilde{\mu}(B) = 2 \qquad \tilde{\mu}(Z) = 9$$

We keep the firm preference profile fixed at $\succ_A$, $\succ_B$ and $\succ_Z$ for the firms and only specify preferences for the workers.

To prove the result, suppose that ψ is not manipulable. We consider the sequence of preference profiles P_1, P_2, P_3, and P_4 formed by swapping $\succ_1'$ for $\succ_1$, then $\succ_2'$ for $\succ_2$, and then $\succ_9'$ for $\succ_9$. We use the non-manipulability of ψ to restrict ψ to a unique match in each case. We then show that at P_3 worker 9 can profitably deviate to $\succ_9'$, a contradiction to the non-manipulability of ψ.

First, we limit the scope of matches we consider. Consider any μ and any P_j.

- If $A \succ_1 \mu(1)$ then 1 strictly prefers $\mu_0(1)$ to $\mu(1)$, hence μ is not in the agreeable core; the same holds for $B \succ_2 \mu(2)$, $2 \succ_B \mu(B)$, and $1 \succ_A \mu(A)$.
- If $j \neq 4$ and $Z \succ_9 \mu(9)$, then $\{9, Z\}$ is an agreeable coalition that blocks μ.
- If $j = 4$ and $Z \succ_9 \mu(9)$, then μ in the agreeable core implies that $\mu(1) \neq Z$ and hence $B \succ_9 \mu(9)$ implies that $\{2, 9, B, Z\}$ is an agreeable coalition that blocks μ; hence, if μ is in the agreeable core then $\mu(9) = B$.
- If $\mu(1) = Z$ and $\mu(2) = A$, then for P_1 $\{1, A, Z\}$ is an agreeable blocking coalition and for P_2, P_3, and P_4 $A \succ_1' Z$. Hence for all P_j $\mu(1) = Z$ and $\mu(2) = A$ imply that μ is not in the agreeable core.

It follows that every worker is matched to a firm, and thus every firm is matched to a worker. Therefore, any match in the agreeable core only occurs between agents who are listed on each other's preferences in figure 3. An exhaustive search reveals that μ°, $\mu^\blacktriangle$, and $\tilde{\mu}$ are the only matches that meet these criteria.

For P_1, the agreeable core is $\{\mu^\circ\}$ because:

✓ μ° is the output of the PE algorithm and hence is in the agreeable core.

✗ $\mu^\blacktriangle$ is blocked by the agreeable coalition $\{1, A, Z\}$ with any deviation μ' such that $\mu'(1) = Z$ and $\mu'(A) = A$.

✗ $\tilde{\mu}$ is blocked by the agreeable coalition $\{1, 2, A, B\}$ with any deviation μ' such that $\mu'(1) = B$ and $\mu'(2) = A$.

Hence, $\psi(P_1) = \mu^\circ$.

For preferences P_2, the agreeable core is $\{\mu^\circ, \mu^\blacktriangle\}$ because:

✓ μ° does not match any worker to a firm he dropped from his preference, so every blocking coalition under these preferences forms under the prior preferences.

✓ $\mu^\blacktriangle$ is the output of the PE algorithm and hence is in the agreeable core.

✗ $\tilde{\mu}$ is blocked by the agreeable coalition $\{1, 2, A, B\}$ with any deviation μ' such that $\mu'(1) = B$ and $\mu'(2) = A$.

If $\psi(P_2) = \mu^\blacktriangle$, then consider the deviation by worker 1 of reporting $\succ_1$ at P_2. Because $\mu^\circ(1) \succ_1' \mu^\blacktriangle(1)$, this is a profitable deviation. Therefore, because ψ is not manipulable, $\psi(P_2) = \mu^\circ$.

For preferences P_3, the agreeable core is $\{\mu^\blacktriangle, \tilde{\mu}\}$ because:

- ✗ μ° matches worker 2 to firm A, which violates the requirement that $\mu(2) \succsim_2 B$.
- ✓ $\mu^\blacktriangle$ is the output of the PE algorithm and hence is in the agreeable core.
- ✓ $\tilde{\mu}$: Observe that Z cannot be strictly better off in any blocking coalition, and thus 2 cannot be strictly better any blocking coalition. Furthermore, any agreeable coalition that makes 1 strictly better off must include B and hence, because the coalition is agreeable, 2. Therefore, any agreeable blocking coalition cannot make any worker strictly better off. Hence, $\tilde{\mu}$ is also in the agreeable core.

If $\psi(P_3) = \mu^\blacktriangle$, then consider the deviation by worker 2 of reporting $\succ_2'$ at P_2. Because $\mu^\blacktriangle(2) \succ_2 \mu^\circ(2)$, this is a profitable deviation. Therefore, because ψ is not manipulable, $\psi(P_3) = \tilde{\mu}$.

In this final step, we note that the core under P_4 is the singleton $\mu^\blacktriangle$. To see this, observe that μ° and $\tilde{\mu}$ each match a worker to a firm he lists below his reference match, and therefore none of these three matches is in the agreeable core. $\mu^\blacktriangle$ is the output of the PE algorithm and hence is in the agreeable core. However, consider the deviation by worker 9 of reporting $\succ_9'$ at P_3. Because $\mu^\blacktriangle(9) \succ_9 \tilde{\mu}(9)$, this is a profitable deviation. Therefore, ψ is manipulable, a contradiction. □

Proof of lemma 3: For the first claim, suppose (toward a contradiction) that w and f are both free agents in μ who also both prefer each other to $\mu(w)$ and $\mu(f)$, respectively. We construct a blocking path in $(A, \mu_0, \mu \cup I(\mu))$, a contradiction to the supposition that μ is in the agreeable core.

Because w is a free agent in μ, w lies on a linear, complete, and alternating path P_w of (A, μ_0, μ). Rewrite P_w such that

$$P_w = (e_1, \dots, e_{k-1}, \{\mu_0(w), w\}, \{w, \mu(w)\}, \dots)$$

Similarly, there is a complete and alternating P_f such that

$$P_f = (\dots, \{\mu(f), f\}, \{f, \mu_0(f)\}, e_{k+1}, \dots, e_n)$$

There are two cases:

1. *P_w and P_f do not intersect:* Then

$$(e_1, \dots, e_{k-1}, \{w, f\}, e_{k+1}, e_n)$$

is a blocking path of μ.

2. *P_w and P_f do intersect:* Then let i be the greatest index less than k such that e_i is in P_f. Let e_j be the edge in P_f such that $e_i = e_j$. Therefore the path

$$(e_j, \ldots, e_{k-1}, \{w, f\}, e_{k+1}, \ldots e_{j-1})$$

is a blocking path of μ.

In either case there is a blocking path of μ. But then μ is not in the agreeable core, a contradiction.

For the second claim we can repeat the argument from the first claim, substituting the edge $\{a\}$ for $\{a, \mu(a)\}$ in path P_a. □

Lemma A.4. *Let μ and ν be structurally similar matches in the agreeable core. Then $(\mu \vee \nu)(w) \in F$ if and only if $\mu(w) \in F$ <u>or</u> $\nu(w) \in F$. Similarly, $(\mu \vee \nu)(f) \in W$ if and only if $\mu(f) \in W$ <u>and</u> $\nu(f) \in W$. A symmetric result holds for $\wedge$.*

Proof. Both statements clearly hold for every agent that is not free in μ (and ν because μ and ν are structurally similar). Hence, we show that the statements hold for the free agents in μ.

For the first statement:

- *For the ($\Rightarrow$) direction:* We show that if $\mu(w) \notin F$ and $\nu(w) \notin F$, then $(\mu \vee \nu)(w) \notin F$. Then $\mu(w) = \nu(w) = w$, which implies $(\mu \vee \nu)(w) = w$. Thus $(\mu \vee \nu)(w) \notin F$.

- *For the ($\Leftarrow$) direction:* We show that if $\mu(w) \in F$ or $\nu(w) \in F$, then $(\mu \vee \nu)(w) \in F$. To see this, note that if $\mu(w) = f$ or $\nu(w) = f$, then w strictly prefers f to being unmatched (w) by lemma 3. Therefore, $\mu \vee \nu$ cannot leave w unmatched and therefore $(\mu \vee \nu)(w) \in F$.

For the second statement:

- *For the ($\Rightarrow$) direction:* We show that if either $\mu(f) \notin W$ or $\nu(f) \notin W$, then $(\mu \vee \nu)(f) \notin W$. Then $\mu(f) = f$ or $\nu(f) = f$. By lemma 3, f weakly prefers both $\mu(f)$ and $\nu(f)$ being unmatched. By the definition of $\vee$, $(\mu \vee \nu)(f) = f$. Therefore, $(\mu \vee \nu)(f) \notin W$.

- *For the ($\Leftarrow$) direction:* We show that if $\mu(f) \in W$ and $\nu(f) \in W$, then $(\mu \vee \nu)(f) \in W$. Then $\{\mu(f), \nu(f)\} \subseteq W$. Therefore $(\mu \vee \nu)(f) \in W$.

This completes the proof. □

Proof of lemma 4: We draw our proof from the proof of Theorem 2.16 in Roth and Sotomayor (1990). We show that $\mu \vee \nu$ is a match; the argument for $\mu \wedge \nu$ is symmetric.

Because the free agents are the same in μ and ν, we need only to show that $\mu \vee \nu$ is a match on the free agents of μ and ν; all other matches are left unchanged because μ and ν are structurally similar. It is immediate from the definition of $\vee$ that items 1 and 2 from the definition of a match hold. That is, we only need that $(\mu \vee \nu)(a) = b \iff (\mu \vee \nu)(b) = a$. Of course, if $a = b$ then the statement is tautological; hence, we prove for $w \in W$ and $f \in F$:

$$(\mu \vee \nu)(w) = f \iff (\mu \vee \nu)(f) = w.$$

For the ($\Rightarrow$) *direction:* We show that $(\mu \vee \nu)(w) = f$ implies $(\mu \vee \nu)(f) = w$. We consider the case when $\mu(w) = f$; the other case is symmetric. Suppose (toward a contradiction) that $(\mu \vee \nu)(f) \neq w$. Then $(\mu \vee \nu)(f) = \nu(f)$. Then f strictly prefers w to $\nu(f)$ and w strictly prefers f to $\nu(w)$, so w and f is a blocking pair of ν, a contradiction by lemma 3. This completes this direction.

For the ($\Leftarrow$) *direction:* We show that $(\mu \vee \nu)(f) = w$ implies $(\mu \vee \nu)(w) = f$. We define a sequence of sets, then study their cardinality. Let

$$\begin{aligned} W' &\equiv \{w \in W : (\mu \vee \nu)(w) \in F\} \\ &= \{w \in W : \mu(w) \in F \text{ or } \nu(w) \in F\} && \because \textit{lemma A.4}. \end{aligned}$$

and

$$\begin{aligned} F' &\equiv \{f \in F : (\mu \vee \nu)(f) \in W\} \\ &= \{f \in F : \mu(f) \in W \text{ and } \nu(f) \in W\} && \because \textit{lemma A.4}. \end{aligned}$$

Observe the following relations:

$$\begin{aligned} |F'| &= |\mu(F')| && \because \mu \text{ is a match} \\ \mu(F') &\subseteq W' && \because \text{Definition of } F' \text{ and } W' \end{aligned}$$

Therefore $|F'| \leq |W'|$. Similarly,

$$\begin{aligned} |W'| &= |(\mu \vee \nu)(W')| && \because (\Rightarrow) \text{ implication} \\ (\mu \vee \nu)(W') &\subseteq F' && \because (\Rightarrow) \text{ implication} \end{aligned}$$

Therefore $|W'| \leq |F'|$ and thus $|W'| = |F'|$. Therefore $|(\mu \vee \nu)(W')| = |F'|$ and thus $(\mu \vee \nu)(W') = F'$.

The final string of implications is as follows: If $(\mu \vee \nu)(f) \in W$, then $f \in F'$. If $f \in F'$, then there exists w in $w \in W'$ such that $(\mu \vee \nu)(w) = f$. This completes this direction.

Therefore, $\mu \vee \nu$ satisfies item 3 from the definition of a match and thus $\mu \vee \nu$ is a match. $\square$

Lemma A.5. *Let μ and ν be structurally similar matches in the agreeable core. Then $\mu \vee \nu \subseteq \mu \cup \nu$ and $I(\mu \vee \nu) \subseteq I(\mu) \cup I(\nu)$. The same holds for $\mu \wedge \nu$.*

Proof. By construction, $\mu \vee \nu$ only contains matches from μ and ν and thus $\mu \vee \nu \subseteq \mu \cup \nu$.

Let $\{w, f\} \in I(\mu \vee \nu)$ and let A^F be the free agents in μ (and ν because μ and ν are structurally similar). There are three cases:

1. $|\{w, f\} \cap A^F| = 0$: Then $(\mu \vee \nu)(w) = \mu(w)$ and $(\mu \vee \nu)(f) = \mu(f)$ by construction, so $\{w, f\} \in I(\mu)$.

2. $|\{w, f\} \cap A^F| = 1$: Suppose that $w \in A^F$; the other case is symmetric. Then either $(\mu \vee \nu)(w) = \mu(w)$ or $(\mu \vee \nu)(w) = \nu(f)$; again, let $(\mu \vee \nu)(w) = \mu(w)$ and the other case is symmetric. Then $(\mu \vee \nu)(f) = \mu(f)$ by construction, so $\{w, f\} \in I(\mu)$.

3. $|\{w, f\} \cap A^F| = 2$: This contradicts lemma 3 and thus cannot happen.

In the cases that do not lead to a contradiction we see that $\{w, f\} \in I(\mu) \cup I(\nu)$, which completes the proof. □

Lemma A.6. *Let μ and ν be structurally similar matches in the agreeable core. Then any blocking path of $\mu \vee \nu$ must contain both a free agent in μ and an agent who is not free in μ. A symmetric result holds for $\mu \wedge \nu$.*

Proof. Let A^F denote the free agents in μ (and ν because μ and ν are structurally similar), and let P be a blocking path of $\mu \vee \nu$. We show that $P \not\subseteq A^F$ and $P \not\subseteq A \backslash A^F$. To see this, consider both cases (toward a contradiction in each case):

1. *Suppose* $P \subseteq A^F$: Then exists an edge e in P such that $e \in I(\mu \vee \nu)$. By lemma A.5, $e \in I(\mu)$ (the other case is symmetric). If $e = \{w, f\}$, then e constitutes a blocking pair and contradicts lemma 3. If $e = \{a\}$, then a strictly prefers being unmatched to μ and contradicts lemma 3. Therefore, $P \not\subseteq A^F$.

2. *Suppose* $P \subseteq A \backslash A^F$: Note that$\mu \vee \nu$ agrees with μ on $A \backslash A^F$. If P blocks $\mu \vee \nu$ then P blocks μ, a contradiction to the supposition that μ is in the agreeable core. Therefore, $P \not\subseteq A \backslash A^F$.

Therefore, P intersects both A and $A \backslash A^F$. By the definition of a path, there exists some e in P that contains an agent in A^F and an agent in $A \backslash A^F$. □

Proof of theorem 3: We show that $\mu \vee \nu$ is in the agreeable core; the argument for $\mu \wedge \nu$ is symmetric. By lemma 4, $\mu \vee \nu$ is a match. Because μ and ν are both individually rational, $\mu \vee \nu$ is individually rational. The remaining step is to show that there are no blocking paths of $\mu \vee \nu$.

Suppose (toward a contradiction) that $\mu \vee \nu$ is blocked by an agreeable coalition. By proposition 1, there is a blocking path P of $\mu \vee \nu$. Let A^F denote the free agents in μ (and ν because μ and ν are structurally similar).

By lemma A.6, there is an edge in P that contains both an agent in A^F and an agent in $A \backslash A^F$. We choose an edge $e = \{a, b\}$ such that $a \in A^F$ and $b \in A \backslash A^F$ in the following manner. By the definition of a blocking path, there must exist some edge e^* in P that is in $I(\mu \vee \nu)$. By lemma 3, e^* must intersect $A \backslash A^F$. Take e to be the closest such edge in P (if two are equally close, choose an arbitrary one). Note that $\mu_0(a) \in A^F$ by definition, so we can write P in the following manner:

$$P = (\overbrace{e_1, \ldots, e_{n-1}, \{\mu_0(a), a\}}^{\text{contained in } A^F}, \overbrace{\underbrace{\{a, b\}}_{e}, e_n, \ldots, e_N}^{\text{contains } e^*}).$$

We show that there must exist an edge e' (distinct from e) in $e_n, \ldots, e_N$ that also contains an agent in A^F and an agent in $A \backslash A^F$. Suppose (toward a contradiction) that no such e' exists. Then the edges $e_n, \ldots, e_N$ are all subsets of $A \backslash A^F$. Note that $a \in A^F$ implies that there is a linear, complete, and alternating path in (A, μ_0, μ)

$$P^a = (e_1^a, \ldots, e_{k-1}^a, \{\mu_0(a), a\}, \{a, \mu(a)\}, e_k^a \ldots, e_K^a).$$

Then

$$(\underbrace{e_1^a, \ldots, e_{k-1}^a, \{\mu_0(a), a\}}_{P^a}, \{a, b\}, \overbrace{e_n, \ldots, e_N}^{P}).$$

is a path in $(A, \mu_0, \mu \cup I(\mu))$ because μ and $\mu \vee \nu$ agree on $A \backslash A^F$. Furthermore, it is complete and alternating by construction. Thus P is a blocking path of μ, a contradiction to the supposition that μ is in the agreeable core. Therefore such an e' exists. Let $e' = \{a', b'\}$ in P be the first such edge in $e_n, \ldots, e_N$, with $a' \in A \backslash A^F$ and $b' \in A^F$. Because e is chosen to be the closest such edge to e^*, it follows that e^* is between e and e' in P.

Before completing the proof, we assess our current construction. We have two edges in P, $e = \{a, b\}$ and $e' = \{a', b'\}$, such that $a, b' \in A^F$ and $b, a' \notin A \backslash A^F$. Furthermore, an edge e^* in $I(\mu \vee \nu)$ is between e and e' in P. By the definition of free agents, we have that $e, e' \notin \mu_0$. Every edge between e and e' is contained entirely in $A \backslash A^F$. A straightforward counting argument[18] shows that one of $\{a, b'\}$ is a worker and the other is a firm, so we rewrite[19]

$$e = \{w, f\}$$
$$e' = \{w', f'\}$$

with $w, f' \in A^F$.

As we observed in the previous paragraph, $e, e' \in (\mu \vee \nu) \cup I(\mu \vee \nu)$. Because w weakly prefers $(\mu \vee \nu)(w)$ to both $\mu(w)$ and $\nu(w)$, and because $(\mu \vee \nu)(f) = \mu(f) = \nu(f)$, it follows that $e \in \mu \cup I(\mu)$

[18]We can repurpose the proof of lemma A.3.

[19]The problem is symmetric here as well, so this is without loss of generality.

and $e \in \nu \cup I(\nu)$. By lemma A.5, we have that $e' \in \mu \cup I(\mu)$ or $e' \in \nu \cup I(\nu)$. Suppose the former case (the latter case is symmetric).

Because $w, f' \in A^F$ implies that there are a linear, complete, and alternating paths (the first is a relabeling of P^a)

$$P^w = (e_1^w, \ldots, e_{k-1}^w, \{\mu_0(w), w\}, \{w, \mu(w)\}, e_k^w, \ldots, e_K^w).$$
$$P^{f'} = (e_1^{f'}, \ldots, e_{l-1}^{f'}, \{\mu(f'), f'\}, \{f', \mu_0(f')\}, e_l^{f'}, \ldots, e_L^{f'}).$$

Then observe that the following path P^* is a complete and alternating path of μ by construction:

$$P^* = (\underbrace{e_1^w, \ldots, e_{k-1}^w, \{\mu_0(w), w\}}_{P^w}, \overbrace{\{w, f\}, \ldots, \{w', f'\}}^{P}, \underbrace{\{f', \mu_0(f')\}, e_l^{f'}, \ldots, e_L^{f'}}_{P^{f'}}).$$

Additionally, the P^* contains $e^* \in I(\mu \vee \nu)$ because e^* is between e and e' in P. By lemma A.5, $e^* \in I(\mu)$. Therefore, P^* is a blocking path of μ, a contradiction to the supposition that μ is in the agreeable. Therefore, there are no blocking paths of $\mu \vee \nu$, which implies that $\mu \vee \nu$ is in the agreeable core. $\square$

References

A. Abdulkadirog. Generalized Matching for School Choice. Jan. 2011.

A. Abdulkadiroğlu and T. Sönmez. House Allocation with Existing Tenants. *Journal of Economic Theory*, 88(2):233–260, Oct. 1999. ISSN 00220531. doi: 10.1006/jeth.1999.2553. URL https://linkinghub.elsevier.com/retrieve/pii/S002205319992553X.

A. Abdulkadiroğlu and T. Sönmez. School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3):729–747, May 2003. ISSN 0002-8282. doi: 10.1257/000282803322157061. URL https://pubs.aeaweb.org/doi/10.1257/000282803322157061.

I. Balbuzanov and M. H. Kotowski. Endowments, Exclusion, and Exchange. *Econometrica*, 87(5):1663–1692, 2019. ISSN 0012-9682. doi: 10.3982/ECTA15676. URL https://www.econometricsociety.org/doi/10.3982/ECTA15676.

J. Combe. Reallocation with priorities and minimal envy mechanisms. *Economic Theory*, 76 (2):551–584, Aug. 2023. ISSN 0938-2259, 1432-0479. doi: 10.1007/s00199-022-01465-x. URL https://link.springer.com/10.1007/s00199-022-01465-x.

J. Combe and J. C. Schlegel. Reallocation with priorities. *Games and Economic Behavior*, 143: 287–299, Jan. 2024. ISSN 08998256. doi: 10.1016/j.geb.2023.11.015. URL https://linkinghub.elsevier.com/retrieve/pii/S0899825623001872.

J. Combe, O. Tercieux, and C. Terrier. The Design of Teacher Assignment: Theory and Evidence. *The Review of Economic Studies*, 89(6):3154–3222, Nov. 2022. ISSN 0034-6527, 1467-937X. doi: 10.1093/restud/rdac002. URL https://academic.oup.com/restud/article/89/6/3154/6520472.

U. Dur, A. A. Gitmez, and Ö. Yılmaz. School choice under partial fairness. *Theoretical Economics*, 14(4):1309–1346, 2019. ISSN 1555-7561. doi: 10.3982/TE2482. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/TE2482. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE2482.

U. M. Dur and T. Morrill. The Impossibility of Restricting Tradeable Priorities in School Assignment. *Working Paper*, May 2017.

U. M. Dur and M. U. Ünver. Two-Sided Matching via Balanced Exchange. *Journal of Political Economy*, 127(3):1156–1177, June 2019. ISSN 0022-3808, 1537-534X. doi: 10.1086/701358. URL https://www.journals.uchicago.edu/doi/10.1086/701358.

A. Erdil and H. Ergin. What's the Matter with Tie-Breaking? Improving Efficiency in School Choice. *The American Economic Review*, 98(3):669–689, 2008. ISSN 0002-8282. URL https://www.jstor.org/stable/29730091. Publisher: American Economic Association.

D. Fragiadakis and P. Troyan. Improving matching under hard distributional constraints: Improving matching under constraints. *Theoretical Economics*, 12(2):863–908, May 2017. ISSN 19336837. doi: 10.3982/TE2195. URL http://doi.wiley.com/10.3982/TE2195.

D. Fragiadakis, A. Iwasaki, P. Troyan, S. Ueda, and M. Yokoo. Strategyproof Matching with Minimum Quotas. *ACM Transactions on Economics and Computation*, 4(1):1–40, Jan. 2016. ISSN 2167-8375, 2167-8383. doi: 10.1145/2841226. URL https://dl.acm.org/doi/10.1145/2841226.

D. Gale and L. S. Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962. ISSN 0002-9890. doi: 10.2307/2312726. URL https://www.jstor.org/stable/2312726. Publisher: Mathematical Association of America.

P. Guillen and O. Kesten. Matching Markets with Mixed Ownership: The Case for a Real-Life Assignment Mechanism. *International Economic Review*, 53(3):1027–1046, 2012. ISSN 0020-6598. URL https://www.jstor.org/stable/23251506. Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].

N. Hamada, C.-L. Hsu, R. Kurata, T. Suzuki, S. Ueda, and M. Yokoo. Strategy-proof school choice mechanisms with minimum quotas and initial endowments. *Artificial Intelligence*, 249: 47–71, Aug. 2017. ISSN 00043702. doi: 10.1016/j.artint.2017.04.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0004370217300528.

O. Kesten. On two competing mechanisms for priority-based allocation problems. *Journal of Economic Theory*, 127(1):155–171, Mar. 2006. ISSN 00220531. doi: 10.1016/j.jet.2004.11.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0022053104002194.

H. Kwon and R. I. Shorrer. Justified-Envy-Minimal Efficient Mechanisms for Priority-Based Matching. Jan. 2023.

T. Morrill. Making Efficient School Assignment Fairer. *Working Paper*, Nov. 2013.

J. S. Pereyra. A dynamic school choice model. *Games and Economic Behavior*, 80:100–114, July 2013. ISSN 08998256. doi: 10.1016/j.geb.2013.02.011. URL https://linkinghub.elsevier.com/retrieve/pii/S0899825613000341.

P. J. Reny. Efficient Matching in the School Choice Problem. *American Economic Review*, 112 (6):2025–2043, June 2022. ISSN 0002-8282. doi: 10.1257/aer.20210240. URL https://pubs.aeaweb.org/doi/10.1257/aer.20210240.

A. E. Roth and M. A. O. Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monographs. Cambridge University Press, 1990.

L. Shapley and H. Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37, Mar. 1974. ISSN 03044068. doi: 10.1016/0304-4068(74)90033-0. URL https://linkinghub.elsevier.com/retrieve/pii/0304406874900330.

P. Troyan, D. Delacrétaz, and A. Kloosterman. Essentially stable matchings. *Games and Economic Behavior*, 120:370–390, Mar. 2020. ISSN 08998256. doi: 10.1016/j.geb.2020.01.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0899825620300105.

S. Ueda, D. Fragiadakis, A. Iwasaki, P. Troyan, and M. Yokoo. Strategy-proof mechanisms for two-sided matching with minimum and maximum quotas (Extended Abstract). *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, June 2012.