Stability, Strategy-Proofness, and Respect for Improvements*

Daisuke Hirata† Yusuke Kasuya‡ Yasunori Okumura§

This Version: February 20, 2023

Graduate School of Economics Discussion Paper No. 2023-01

Abstract

In priority-based two-sided matching, respect for improvements requires a mechanism that an agent should get weakly better off when she is higher prioritized. Not only is it a normative desideratum, but this property is also substantial for ex-ante investments and disclosure of non-preference information. In the general model of matching with contracts, we demonstrate that a stable mechanism respects improvements if and "almost" only if it is strategy-proof, although the precise statement varies across our assumptions. Our results suggest that strategy-proofness is desirable not only as a strategic property but *also for its normative implication*. We also examine (i) simultaneous manipulations of both reported preferences and priority structures, (ii) collective effects of affirmative action policies, and (iii) properties of priority structures sufficient for strategy-proofness and respect for improvements.

JEL Classification: C78, D47

*We are grateful to Akifumi Ishihara, Fuhito Kojima, Shintaro Miura, Takeshi Murooka, Shin Sato, Jan Christoph Schlegel, Kentaro Tomoeda, and Yuichi Yamamoto for their comments. We also thank seminar participants at SWET 2019, CTW Summer Camp 2019, AETW 2020, ES World Congress 2020, GAMES 2020, Doshisha University, Hitotsubashi University, Kyoto University, Nagoya University, Shinshu University, and Waseda University for helpful discussions. Much of this research was conducted while Hirata was visiting UNSW Sydney; he is grateful for their hospitality. This study is financially supported by JSPS KAKENHI 19K01542. The usual disclaimer applies.

†Hitotsubashi University; d.hirata@r.hit-u.ac.jp

‡Kobe University; kasuya@econ.kobe-u.ac.jp

§Tokyo University of Marine Science and Technology; okuyasu@gs.econ.keio.ac.jp

Contents

1 Introduction

Priority-based matching is a problem of matching agents (such as students, doctors, cadets, and lawyers) to institutions (such as schools/colleges, hospitals, military branches, and courts) based on agents' preferences and institutions' priorities. For instance, in many real-world school choice programs, a student is given a higher priority at a school, if she lives in its neighborhood and/or has a sibling attending the same school. In such priority-based matching, it would be unequivocal, simply by the definition of the word, that *being higher prioritized should be good for an agent*. However, we cannot judge if this principle actually holds from priorities alone because they are just an input (or a parameter) of a mechanism. That is, we need to check it *as a property of a matching mechanism*.

While it might sound self-evident, the above principle that being higher prioritized should be good has two interpretations: The first is an *interpersonal* comparison. If an agent is assigned to an institution while it gives another agent a higher priority, the latter agent should be better off than the former in the sense that she should be assigned to a (weakly) better institution. This requirement is tantamount to the *stability* of a matching. A matching would be against the spirit of the priority and thereby be unfair if it is unstable, i.e., if the high-priority agent wants to but cannot get into the institution while the low-priority agent can. Therefore, stability is a central desideratum as a fairness condition, even when a matching authority has enough power to enforce unstable matchings.

The second interpretation is *intrapersonal* comparative statics. If an agent becomes higher prioritized at an institution with everything else being equal, she should get (weakly) better off than before. Balinski and Sönmez (1999) name this requirement as *respect for improvements*. Practically speaking, a matching authority would increase an agent's priority when it aims to increase her welfare. Further, such a policy goal should reflect the social value that the target agent deserves better, not worse. Thus, it would be against the social value if the higher priority makes her worse off (due to the lack of respect for improvements). And perhaps surprisingly, mechanisms in the real world often fail to respect improvements (e.g., Aygün and Bó, 2021; Balinski and Sönmez, 1999; Sönmez, 2013; Sönmez and Switzer, 2013). Hence, respect for improvements is practically significant to ensure that policy changes have fair consequences in accordance with the social value behind them.

For a matching mechanism to truly meet the principle that being higher prioritized should be good, therefore, it should both be stable and respect improvements. However, it is non-trivial to design such a mechanism because, despite the similarity in their spirit, the two desiderata are known to be logically independent (Balinski and Sönmez, 1999). To explore the two desiderata as generally as possible, we adopt the model of *matching with contracts* à la Hatfield and Milgrom (2005). This model generalizes the classic two-sided matching, and the literature has identified a growing number of its applications.[1] As such, matching with contracts enables us to examine different assumptions in a unified way, thereby crystallizing their logical consequences. In this general framework, we investigate conditions for a stable mechanism to respect improvements.

Our main contribution is to demonstrate that *strategy-proofness* is the key condition for a stable mechanism to respect improvements. Namely, a stable mechanism respects improvements, if and "almost" only if it is strategy-proof in the sense that reporting true preferences is a dominant strategy for each agent. This equivalence holds, albeit in a relatively weak form, without any assumptions on institutions' choice functions (Theorem 1). In addition, we establish stronger forms of equivalence under certain assumptions on the domain of priority structures (Theorems 2–4). We will explain these results in detail, along with several additional results, in Section 1.1 below.

Our results shed new light on the importance of strategy-proofness in priority-based matching markets, possibly justifying the prevalent focus on strategy-proof mechanisms in the literature. Purely as an implementability condition, one could argue that strategy-proofness is unnecessarily strong and a weaker condition would suffice to implement a desirable mechanism.[2] Given our results, however, using a non-strategy-proof mechanism has a serious drawback that it (almost) needs to be unstable or to disrespect improvements. The key takeaway of this paper is that strategy-proofness is important not only as an incentive condition but *also through its normative implication*.[3]

Further, by reinterpreting respect for improvements, our equivalence results suggest that strategy-proofness is also substantial for two additional reasons. First, we can see

[1] For applications of matching with contracts, see, e.g., Aygün and Turhan (2019, 2020), Dimakopoulos and Heller (2019), Greenberg et al. (2021), Hafalir et al. (2022), Hassidim et al. (2017), Kominers and Sönmez (2016), Sönmez (2013), Sönmez and Switzer (2013), and Westkamp (2013).

[2] Indeed, several existing studies investigate other implementability conditions in matching problems; e.g., see Haeringer and Klijn (2009), Iwase et al. (2022), and Kumano (2017) for studies on Nash implementation.

[3] The existing literature has argued that strategy-proofness is normatively important in "leveling the playing field" between strategic and naive agents (e.g., Pathak and Sönmez, 2008). Nevertheless, our "normative implication" differs from this existing view, as ours is relevant without the two types of agents.

respect for improvements as a proper incentive for ex-ante efforts and investments. For instance, suppose that colleges higher prioritize students with higher GPAs. Disrespect for improvements in this scenario means that students can be better off by getting a lower GPA, undermining their incentive to study hard at high schools. This possibility is not merely a theoretical concern: Sönmez (2013) reports anecdotal evidence on an Internet forum that cadets at the US Military Academy were aware of such disincentives and discussed the possibility of deliberately lowering their grades. In order to eliminate such disincentives, a mechanism should respect improvements, and thus, it should be strategy-proof.[4]

Second, respect for improvements also incentivizes agents to provide *non-preference* information.[5] Typically as an affirmative action measure, priorities are often based on the agents' personal information such as ethnic backgrounds. If disclosing their background leads to a higher priority that in turn results in a worse outcome, minority agents would refuse to do so, thereby making the affirmative action policy ineffective. Again, this is a practical concern: See Aygün and Turhan (2020) and Sönmez and Yenmez (2022) for the case in India. To induce voluntary disclosure of non-preference information, respect for improvements is essential, and thus, so is strategy-proofness in terms of preference reports.

Technically speaking, our equivalence results generalize several existing results in the literature. In the classic two-sided matching, the deferred acceptance mechanism (Gale and Shapley, 1962) is both (i) the unique stable mechanism that is strategy-proof (Alcalde and Barberà, 1994; Dubins and Freedman, 1981; Roth, 1982) and (ii) the unique stable mechanism that respects improvements (Balinski and Sönmez, 1999). In matching with contracts, the cumulative offer mechanism, which is a variant of the deferred acceptance, is known to satisfy all of stability, strategy-proofness, and respect for improvements when institutions' priorities meet certain conditions (e.g., Afacan, 2017; Avataneo and Turhan, 2021; Aygün and Turhan, 2019, 2020; Kominers and Sönmez, 2016; Sönmez, 2013; Sönmez and Switzer, 2013). These existing results are not a coincidence: From our main results, we can mechanically conclude a stable mechanism respects improvement if it is strategy-proof. It should be noted that our Theorem 1 is so general that it covers the cases where a (unique) stable and strategy-proof mechanism is *not* the cumulative offer mechanism.

The rest of this paper is organized as follows: Section 1.1 provides a brief overview of

[4] For strategy-proofness and investment incentives, see also Hatfield et al. (2021a) and Tomoeda (2019).

[5] For this reason, some recent studies indeed refer to respect for improvements as "incentive compatibility" (e.g., Aygün and Bó, 2021; Aziz and Brandl, 2021; Sönmez and Yenmez, 2022; Pathak et al., 2020).

our results. Sections 2 and 3 introduce the model and several new concepts. Sections 4 and 5 present the main results. Section 6 provides several additional results. The proofs of all Facts, Propositions, and Theorems are gathered in Section 7.

1.1 Preview of the Results

For the purpose of this study, our first task is to define improvements based on *choice functions*. This is because matching with contracts formulates the priority structure at an institution as a choice function, which specifies the subset of contracts that the institution would choose from each possible menu (i.e., each possible set of applications). This is a key feature of the model, as it encompasses priority structures that cannot be reduced to simple linear orders, thereby broadening the scope of the model.[6] At the same time, this formulation makes the definition of improvements (i.e., what changes in priority structures should be favorable for a particular agent) less obvious. In this study, we propose and examine two definitions, strong and weak, of improvements. As the names suggest, strong improvements are a subset of weak improvements; i.e., if a choice function is a strong improvement over another, it is also a weak improvement. Accordingly, *respect for strong improvements* is a weaker requirement for a mechanism than *respect for weak improvements*.

With these definitions, we establish four equivalence results between strategy-proofness and respect for improvements, making different assumptions about the domain of admissible choice functions. First of all, we establish a general result concerning respect for strong improvements, making no assumptions on the domain at all. Even with such unstructured domains, Theorem 1 ensures that a stable mechanism is strategy-proof if and only if it respects improvements and satisfies an auxiliary condition, which we name the irrelevance of unchosen contracts (for short, IUC). It is particularly noteworthy that for this theorem, we do not assume any substitutability conditions, which play a central role in theoretical studies of matching with contracts.[7] Consequently, Theorem 1 covers the cases where the desirable mechanism it characterizes is not the cumulative offer mechanism (for short, COM; Hatfield

[6] A typical example is school choice programs with affirmative action policies (e.g., Ehlers et al., 2014; Hafalir et al., 2013; Kojima, 2012). A school may prioritize a minority student over a non-minority if it admits too few other minority students, while it may reverse the ranking otherwise. See also Echenique and Yenmez (2015) and Imamura (2020) for axiomatic approaches to priority structures as choice functions.

[7] For those conditions, see Afacan and Turhan (2015), Hatfield and Kojima (2010), Hatfield and Kominers (2019), Hatfield et al. (2021b), and Hatfield and Milgrom (2005) among others.

and Milgrom, 2005). The COM is a variant of the deferred acceptance and has been the leading candidate for a "desirable" mechanism in the literature. Nevertheless, Hirata and Kasuya (2017, Example 1) demonstrate that in the general case, a stable and strategy-proof mechanism need not be the COM. Thus, the connection between strategy-proofness and respect for improvements we identify in Theorem 1 arises purely from the definitions of the desiderata, independently of the nature of any particular mechanisms.

Next, we characterize respect for weak improvements assuming admissible choice functions are *observably substitutable* (Hatfield et al., 2021b). Without this assumption, a stable and strategy-proof mechanism can fail to respect weak improvements, although it should respect strong improvements by Theorem 1. In the case of observable substitutability, however, a stable mechanism is strategy-proof if and only if it respects weak improvements and satisfies the IUC (Theorem 2). It should be noted that we cannot drop the IUC in this statement. Under observable substitutability, a stable mechanism may respect weak improvements without being strategy-proof. Then, it would be natural to ask under what condition, if any, respect for improvements by itself is fully equivalent to strategy-proofness.

Our answer to this question is two-fold: First, we identify such a condition, restricting our attention to the COM rather than any stable mechanisms. To this end, we define a new property strengthening observable substitutability. When choice functions satisfy this *strong observable substitutability*, the COM is strategy-proof if and only if it respects (either weak or strong) improvements (Theorem 3). Second, we establish the full equivalence for an arbitrary stable mechanism, with further strengthening the substitutability condition: When choice functions satisfy *unilateral substitutability* (Hatfield and Kojima, 2010), a stable mechanism is strategy-proof if and only if it respects (either weak or strong) improvements (Theorem 4). This last result defines a sharp limit for non-strategy-proof mechanisms: Without assumptions, such a mechanism can be stable and respect improvements. However, according to Theorem 4, this is possible only if we restrict our attention to choice functions that are *not* unilaterally substitutable.

Additional Results

In addition to the main equivalence results, we also examine three related issues. First, we consider two-dimensional manipulations of a mechanism. As we discussed above, in certain markets, agents could intentionally disimprove their own positions in priority structures. If so, they might be able to benefit by manipulating both their preference reports and the choice functions, even though they cannot do so with preference manipulations only. However, our main results entail that this is *not* the case for stable mechanisms: A stable mechanism is immune to two-dimensional manipulations if (and only if) it is strategy-proof in the standard sense (Theorem 5).

Second, we investigate the *collective* effects of priority changes. To this end, we consider the COM under observable substitutability and assume it is strategy-proof. Then, when the priority structures change in favor of a group of agents, the target group never strictly Pareto deteriorates; i.e., at least one of them should be weakly better off (Theorem 6). While this conclusion might appear not appealing enough, in fact, it is almost impossible to satisfy stronger, more appealing requirements. See Section 6.2 for details and the discussion of related results in the literature.

Lastly, we explore under what conditions, in terms of primitives of the model, there is a stable mechanism that respects improvements. Notice that our main results indirectly answer this question: The existing literature has identified several sufficient conditions for the COM to be stable and strategy-proof. Our main results imply that those conditions also suffice for respect for improvements. In addition to such an indirect answer, we also provide a direct one; i.e., we establish a novel sufficient condition on choice functions so that the COM is stable, strategy-proof, and respects improvements. Our condition is stronger than the one by Hatfield et al. (2021b), which is the weakest to date, but ours has its own merits relative to theirs. See Section 6.3 for details, including the discussion of the literature.

2 Environment

Let D and H be finite sets of *agents* and *institutions*, respectively. The finite set X^G of possible *contracts* is given by a subset of $D \times H \times \Theta$ for some finite Θ. The elements of Θ, called *contractual terms*, represent different ways for a pair $(d, h) \in D \times H$ to be matched.[8] For each contract $x \in X^G$, let $\mathsf{d}(x)$ and $\mathsf{h}(x)$ be its projections onto D and H, i.e., $x = (\mathsf{d}(x), \mathsf{h}(x), \theta)$ for some $\theta \in \Theta$. In other words, each x is a bilateral contract

[8] Examples of different contractual terms include salary levels and jobs at an employer (Kelso and Crawford, 1982; Roth, 1984), tuition levels at a university (Artemov et al., 2021; Biró et al., 2021), lengths of service at a military branch (Greenberg et al., 2021; Sönmez, 2013; Sönmez and Switzer, 2013), and waiting times for legal traineeships at a regional court (Dimakopoulos and Heller, 2019).

between agent $\mathsf{d}(x) \in D$ and institution $\mathsf{h}(x) \in H$. For a subset X of contracts, we write $\mathsf{d}(X)$ and $\mathsf{h}(X)$ to denote $\{\mathsf{d}(x) : x \in X\}$ and $\{\mathsf{h}(x) : x \in X\}$. The power set of X^G is denoted by 2^{X^G}.

A subset $X \subseteq X^G$ of contracts is said to be an *allocation* if it includes at most one contract for each agent, i.e., if $x, x' \in X$ and $x \neq x'$ imply $\mathsf{d}(x) \neq \mathsf{d}(x')$. The set of all possible allocations is denoted by $\mathscr{A} \subseteq 2^{X^G}$. For each allocation $X \in \mathscr{A}$ and agent $d \in D$, let $\mathsf{x}(d, X)$ denote the contract that X assigns to d; i.e., $\mathsf{x}(d, X) = x$ if $x \in X$ and $\mathsf{d}(x) = d$. If there is no such contract in X for agent d, then d is said to be assigned the *null contract* and we let $\mathsf{x}(d, X) = \varnothing$. In what follows, we use the symbols $\varnothing$ and $\emptyset$ to denote the null contract and the empty set, respectively.

Each agent $d \in D$ has a strict preference relation represented by a linear order $\succ_d$ over some $\mathrm{Ac}(\succ_d) \subseteq \{x \in X^G : \mathsf{d}(x) = d\}$, where $\mathrm{Ac}(\succ_d)$ denotes the set of all *acceptable* contracts (i.e., the set of those which are preferred to the null contract). That is, we identify d's preference with his ranking over the acceptable contracts and ignore the ranking among unacceptable contracts. This is without loss of generality as long as the mechanisms we consider also ignore such information (i.e., as long as their outcomes are invariant in regard to the preferences among unacceptable contracts). Let $\mathscr{P}_d$ be the set of all such preferences for agent d. A profile of the agents' preferences and the domain of all possible profiles are denoted by $\succ_D := (\succ_d)_{d \in D}$ and $\mathscr{P}_D := \prod_{d \in D} \mathscr{P}_d$, respectively.

With our formulation of preferences, the following concept becomes well-defined: Taking a subset Y of contracts and a preference $\succ_d$ as given, the *dropping* of Y from $\succ_d$, denoted by $\succ_d^{-Y}$, is the unique preference order such that (i) $\mathrm{Ac}\left(\succ_d^{-Y}\right) = \mathrm{Ac}(\succ_d) - Y$ and (ii) $w \succ_d^{-Y} w' \Leftrightarrow w \succ_d w'$ for all $w, w' \in \mathrm{Ac}\left(\succ_d^{-Y}\right)$. Note that $\succ_d^{-Y}$ is well-defined even if $\mathsf{d}(y) \neq d$ for some $y \in Y$. In particular, $\succ_d^{-Y} = \succ_d$ if $d \notin \mathsf{d}(Y)$. Given a profile $\succ_D = (\succ_d)_{d \in D}$, we call $\succ_D^{-Y} = \left(\succ_d^{-Y}\right)_{d \in D}$ the dropping of Y from $\succ_D$. When the dropped set of contracts is a singleton, for brevity, we write $\succ_d^{-x}$ and $\succ_D^{-x}$ instead of $\succ_d^{-\{x\}}$ and $\succ_D^{-\{x\}}$, respectively.

For the sake of notational simplicity, we extend the agents' preferences in three natural steps: First, we slightly abuse notation and write $x \succ_d \varnothing$ and $x \succ_d y$ when $x \in \mathrm{Ac}(\succ_d)$ and $y \notin \mathrm{Ac}(\succ_d)$. These are in line with the definition that $\mathrm{Ac}(\succ_d)$ is the set of "acceptable" contracts, which should be preferred to the null contract and to any unacceptable contract. Second, we further abuse notation and use the same symbol to compare allocations. For two allocations $X, X' \in \mathscr{A}$, we write $X \succ_d X'$ to denote $\mathsf{x}(d, X) \succ_d \mathsf{x}(d, X')$. Likewise, $X \succeq_d X'$ and $X =_d X'$ denote $\mathsf{x}(d, X) \succeq_d \mathsf{x}(d, X')$ and $\mathsf{x}(d, X) = \mathsf{x}(d, X')$, respectively. Third, we also compare an allocation and a contract in an analogous way; e.g., $X \succ_d x$ denotes $\mathsf{x}(d, X) \succ_d x$.

Each institution $h \in H$ has a *choice function* $C_h : 2^{X^G} \to \mathscr{A}$ such that for every menu $X \subseteq X^G$ of contracts, (i) $C_h(X) \subseteq X$ and (ii) $\mathsf{h}(x) = h$ for all $x \in C_h(X)$. Throughout the paper, we assume that the choice functions satisfy the following mild requirement: Institution h's choice function $C_h(\cdot)$ is said to satisfy the *irrelevance of rejected contracts* (for short, IRC; Aygün and Sönmez, 2013) if $x \notin C_h(X \cup \{x\})$ implies $C_h(X' \cup \{x\}) = C_h(X)$ for all $x \in X^G$ and $X \subseteq X^G$. Note that this condition is satisfied whenever a choice function is induced by a strict preference over subsets of contracts. Taking a choice function C_h as given, the *rejection function R_h associated with C_h* is defined by $R_h(X) := X - C_h(X)$ for each $X \subseteq X^G$.

A profile of the institutions' choice functions is denoted by $C_H = (C_h)_{h \in H}$. With slight abuse of notation, we will often identify C_H with the aggregate choice function, letting $C_H(X)$ denote $\cup_h C_h(X)$ for each $X \subseteq X^G$. Note that the aggregate $C_H(\cdot)$ should satisfy the IRC, given that each component C_h does. The aggregate rejection function associated with C_H is defined by $R_H(X) := X - C_H(X) = \cap_h R_h(X)$ for each $X \subseteq X^G$. The domain of profiles of choice functions under consideration is denoted by $\mathscr{C}_H$. We will impose some restrictions on $\mathscr{C}_H$ later, but for the moment, suppose that $\mathscr{C}_H$ is arbitrary except that each $C_H \in \mathscr{C}_H$ satisfies the IRC.

Given $\succ_D$ and C_H, we define the following concepts on the set $\mathscr{A}$ of all allocations: An allocation $X \in \mathscr{A}$ is said to be *individually rational* at $(\succ_D, C_H)$, if $\mathsf{x}(d, X) \succeq_d \varnothing$ for all $d \in D$ and $C_h(X) = \{x \in X : \mathsf{d}(x) = h\}$ for all $h \in H$. A pair of an institution $h \in H$ and a subset $X' \subseteq X^G$ of contracts is said to *block* an allocation $X \in \mathscr{A}$ at $(\succ_D, C_H)$ if $C_h(X \cup X') \neq C_h(X)$ and $C_h(X \cup X') \succeq_d X$ for all $d \in \mathsf{d}(C_h(X \cup X'))$.[9] An allocation X is said to be *stable* at $(\succ_D, C_H)$ if it is individually rational and not blocked by any $(h, X') \in H \times 2^{X^G}$.

Given (D, H, X^G) as well as the domain $\mathscr{C}_H$ of admissible profiles of choice functions, a *mechanism* is a mapping $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$. A mechanism $F(\cdot, \cdot)$ is said to be *stable*

[9] Requiring $C_h(X \cup X') = X'$ is redundant here, although it is often a part of the definition in the literature. See Hirata and Kasuya (2017, Lemma 1) for details.

(resp. *individually rational*) if for each $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$, its output $F(\succ_D, C_H)$ is stable (resp. individually rational) at $(\succ_D, C_H)$. Lastly, a mechanism $F(\cdot,\cdot)$ is said to be *strategy-proof* if $F(\succ_D, C_H) \succeq_d F((\rhd_d, \succ_{-d}), C_H)$ holds for all $\succ_D \in \mathscr{P}_D$, $C_H \in \mathscr{C}_H$, $d \in D$, and $\rhd_d \in \mathscr{P}_d$, where $\succ_{-d} = (\succ_{d'})_{d' \in D - \{d\}}$.

3 Strong Improvements and Related Concepts

In this section, we formally introduce the concept of strong (dis)improvements and then define two properties for a matching mechanism. To start with, taking a choice function C_h and a subset Y of contracts as given, define another choice function C_h^{-Y} by

$$C_h^{-Y}(X) := C_h(X - Y) \text{ for all } X \subseteq X^G.$$

That is, C_h^{-Y} differs from the original C_h in ignoring the contracts in Y as if they are not in the menu, *even when they actually are*. It is easy to check $C_h^{-Y}(\cdot)$ meets all the requirements to be a choice function for h. In particular, it meets the IRC given C_h does; see Appendix E for a proof. Note also that C_h^{-Y} is well-defined even if Y contains a contract x such that $\mathsf{h}(x) \neq h$; in particular, by the IRC, $C_h^{-Y} = C_h$ when $h \notin \mathsf{h}(Y)$. Given a profile $C_H = (C_h)_{h \in H}$ of choice functions, we write C_H^{-Y} to denote the profile $\left(C_h^{-Y}\right)_{h \in H}$. When $Y = \{x\}$ is a singleton, for simplicity, we write C_h^{-x} and C_H^{-x} instead of $C_h^{-\{x\}}$ and $C_H^{-\{x\}}$.

In what follows, we call C_h a *strong Y-improvement* over C_h^{-Y}; conversely, we also refer to the latter as the *strong Y-disimprovement* of the former. Note that a strong Y-improvement over a given choice function is not unique, because $C_h^{-Y} = \widetilde{C}_h^{-Y}$ can hold even if $C_h \neq \widetilde{C}_h$. In contrast, for any C_h and Y, the strong Y-disimprovement is unique. Comparing two profiles of choice functions, $C_H = (C_h)_{h \in H}$ and $C_H^{-Y} = \left(C_h^{-Y}\right)_{h \in H}$, we call the former (resp. the latter) a strong Y-improvement over the latter (resp. the strong Y-disimprovement of the former). When Y is a singleton, we refer to $\{x\}$-(dis)improvements simply as x-(dis)improvements.

We can view strong improvements as an introduction of new matching opportunities to the market. Suppose that a new contract, say x, is newly introduced and it is the only change in the market. First, then, the choice functions before the change should not have chosen x from any menu. Second, the choice functions before and after the change should agree with each other unless x is not in the menu, reflecting the fact that the introduction of x is the only change. Actually, it then follows that the choice function after the change should be an x-improvement over the original one. More generally, under the assumption of the IRC, two choice functions C_h and C'_h satisfy (i) $Y \cap C'_h(X) = \emptyset$ and (ii) $Y \cap X = \emptyset \Rightarrow C_h(X) = C'_h(X)$ for all $X \subseteq X^G$, if and only if $C'_h = C_h^{-Y}$.

Now we define two properties for a mechanism concerning changes in the priority structures. The first is respect for improvements, where "improvements" are taken to be the strong improvements defined above. When $\mathsf{d}(Y) = \{d\}$ for some $d \in D$, a strong Y-improvement opens up new opportunities only for the single agent d, with everything else being kept constant. As such, it would be natural to argue that it should be a favorable change for d and make her better off.

Definition 1. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to *respect strong improvements* if $F(\succ_D, C_H) \succeq_d F(\succ_D, C_H^{-Y})$ holds for all $d \in D$, $\succ_D \in \mathscr{P}_D$, $C_H \in \mathscr{C}_H$, and $Y \subseteq X^G$ such that $\mathsf{d}(Y) = \{d\}$ and $C_H^{-Y} \in \mathscr{C}_H$. □

The second is an invariance property, which we name the *irrelevance of unchosen contracts* (for short, IUC). It requires that introducing new opportunities should affect the matching outcome only if some of the new contracts are chosen. Conversely, it necessitates that an "abolishment" of a contract x (i.e., a change from C_h to C_h^{-x}) should be irrelevant unless, with the original choice functions, the mechanism would have chosen the "abolished" x. Note that our IUC is a property of a mechanism while the IRC is of a choice function. These two are thus logically incomparable, despite the similarity in their spirit. Notice also that the following definition is relatively weak in that it focuses on the case of $\mathsf{d}(Y) = \{d\}$ and it requires the equality only for d. (Recall that for two allocations X and X', $X =_d X'$ means agent d signs the same contract at both.)

Definition 2. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to satisfy *the irrelevance of unchosen contracts* (for short, IUC) if

$$[Y \cap F(\succ_D, C_H) = \emptyset] \Longrightarrow F(\succ_D, C_H) =_d F\left(\succ_D, C_H^{-Y}\right),$$

for all $d \in D$, $\succ_D \in \mathscr{P}_D$, $C_H \in \mathscr{C}_H$, and $Y \subseteq X^G$ such that $\mathsf{d}(Y) = \{d\}$ and $C_H^{-Y} \in \mathscr{C}_H$. □

While it has a natural interpretation we discussed above, it should be also noted that the concept of strong improvements is rather extreme. In certain special cases, what should be regarded as an improvement becomes less ambiguous. In those cases, our strong improvements reduce to a proper subset of the naturally-defined improvements. The following example highlights this fact. However, note that the narrowness of strong improvements makes respect for them weak. Therefore, it should not undermine the *necessity* of respect for strong improvements, even though one might reasonably argue that it is not sufficient for a desirable mechanism. We will consider a broader concept of improvements (and hence, a stronger definition of respect for improvements) in Section 5.

Example 1 (the classic model)**.** Suppose $X^G = D \times H$, i.e., there is no contractual terms. Let P be a linear priority order over $D \cup \{\emptyset\}$, where $\emptyset$ represents vacancy. Also let $q \in \mathbb{N}$ refer to the quota of an institution. The *classic choice function* $C_h^{P,q}(\cdot)$ for institution h is induced by (P, q) as follows: It is defined to be the choice function that chooses $x = (d, h)$ if and only if d is among the best q acceptable applicants according to P.[10] Comparing two priority orders, we call P a *classic improvement* over Q for d, if

$$[d\ Q\ e \Rightarrow d\ P\ e] \text{ and } [e\ Q\ e' \Leftrightarrow e\ P\ e'],$$

for any $e, e' \in (D - \{d\}) \cup \{\emptyset\}$. One can confirm that $C_h^{P,q}$ is a strong (d, h)-improvement over $C_h^{Q,q}$ if and only if (i) P is a classic improvement over Q and (ii) $\emptyset\ Q\ d$. Notice that without the second requirement, $C_h^{P,q}$ is not necessarily a strong (d, h)-improvement over $C_h^{Q,q}$ even if P is a classic improvement over Q. □

3.1 Richness of the Domain

To meaningfully study the two properties we introduced above, we need some appropriate assumption concerning the richness of the domain $\mathscr{C}_H$ of choice function profiles. On the one hand, the domain must be sufficiently rich for the two properties to take effect: As an extreme example, they become vacuous if $\mathscr{C}_H$ is a singleton or if no profile in $\mathscr{C}_H$ is a strong improvement over another. On the other hand, $\mathscr{C}_H$ cannot be too inclusive because otherwise, it would become inconsistent with the existence of a desirable mechanism we will consider: For instance, if we let $\mathscr{C}_H$ be the set of all possible profiles of choice functions, no mechanism on $\mathscr{P}_D \times \mathscr{C}_H$ is stable. In balancing these opposing needs, we adopt the

[10] More formally, $(d, h) \in C_h^{P,q}(X)$ if and only if $d\ P\ \emptyset$ and $q > \#\{(d', h) \in X : d'\ P\ d\}$.

following definition of a *rich domain* of choice function profiles.[11]

Definition 3. A domain $\mathscr{C}_H$ of profiles of choice functions is said to be *rich* if for any $C_H \in \mathscr{C}_H$ and $x \in X^G$, we have $C_H^{-x} \in \mathscr{C}_H$. □

While this definition is enough for our purpose, it is actually quite weak from the following two perspectives: First, our richness is consistent with a variety of conditions for choice functions, including the IRC and those which we will assume in later sections. Specifically, given any domain satisfying any subset of those conditions, we can trivially expand it to a rich one without violating them. This is because strong disimprovements preserve all of those conditions as we demonstrate in Appendix E. In this sense, assuming richness is without loss of generality, at least in direct relation to those additional assumptions.

Second, richness defined as above is orthogonal to the existence of a "desirable" mechanism. Recall that our main results relate strategy-proofness and respect for improvements of stable mechanisms. Hence, they are meaningful only for domains where there is a stable mechanism that is strategy-proof or respects improvements. A priori, richness and the existence of such a "desirable" mechanism could *jointly* require some additional properties of $\mathscr{C}_H$. It turns out, however, that this is not the case and the existence on a rich domain is no stronger than that for a single profile of choice functions, as we can take $\mathscr{C}_H$ to be a singleton in the following fact.

Fact 1. *Suppose that there exists a stable and strategy-proof mechanism F on $\mathscr{P}_D \times \mathscr{C}_H$. Then, there exists an extension of F that is stable and strategy-proof on $\mathscr{P}_D \times \mathscr{C}_H^*$, where $\mathscr{C}_H^* := \{C_H^{-Y} : C_H \in \mathscr{C}_H$ and $Y \subseteq X^G\}$ is the smallest rich domain containing $\mathscr{C}_H$.*

4 Respect for Strong Improvements on General Domains

In this section, we provide the most general form of our main results, relating strategy-proofness and respect for improvements *without any assumptions* on choice functions. Notice that the two properties we introduced in Section 3 restrict a matching mechanism in a different dimension from the one strategy-proofness does: On the one hand, respect for improvements and the IUC are about changes in institutions' choice functions, and

[11] Note that the following richness condition is logically independent of *unitality* of Hatfield et al. (2021b), which requires $\mathscr{C}_H$ to contain all possible combinations of "unit-demand" choice functions.

their definitions take agents' preferences as fixed. On the other hand, strategy-proofness concerns misrepresentations of agents' preferences, taking institutions' choice functions as given. Nevertheless, some fundamental links exist between the two dimensions, even on the general "unstructured" domains, where we can rely only on the definitions of the desiderata.

Theorem 1. *Let $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ be a stable mechanism. Then, F respects strong improvements and satisfies the IUC if it is strategy-proof. When $\mathscr{C}_H$ is rich, the converse is also true: F respects strong improvements and satisfies the IUC (if and) only if it is strategy-proof.*

We can summarize the key components of the proof of Theorem 1 into two propositions, which we present below. First, for possible mechanisms we consider in the above theorem, an outcome equivalence holds between certain changes in preferences and those in priority structures:

Proposition 1. *Let $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ be a stable mechanism and suppose either (i) F is strategy-proof or (ii) $\mathscr{C}_H$ is rich, F respects strong improvements, and it satisfies the IUC. Then, $F\left(\succ_D, C_H^{-Y}\right) = F\left(\succ_D^{-Y}, C_H\right)$ holds for all $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$ and $Y \subseteq X^G$ such that $\mathsf{d}(Y) = \{d\}$ for some $d \in D$ and $C_H^{-Y} \in \mathscr{C}_H$.*

Roughly speaking, the conclusion of this proposition states that *not applying* for a contract x (i.e., submitting $\succ_d^{-x}$ instead of $\succ_d$) should lead to the same consequence as of the application for x being *nullified and ignored* (i.e., the priority structures changing from C_H to C_H^{-x}). As we will present as Fact 2 in Section 5.2, under a relatively mild condition, the COM (Definition 5 below) satisfies this property solely by its algorithmic definition. The above proposition might thus appear not surprising to those who are familiar with the literature.

However, we should emphasize two points here: First, the mechanism F in Proposition 1 need not be the COM. With the general domains we are analyzing, Hirata and Kasuya (2017) provide an example where the unique stable strategy-proof mechanism is not the COM, but even in such a case, the above proposition is applicable.[12] Proposition 1 is non-trivial in identifying a universal property of a "desirable" mechanism, whether or not it is the COM. Second, neither the Boston mechanism nor the top trading cycles mechanism

[12]The example by Hirata and Kasuya (2017) is for a single profile of choice functions, but combined with Fact 1, it can be generalized to a rich domain of profiles of choice functions.

(Abdulkadiroğlu and Sönmez, 2003) meets the above property. Note that these mechanisms naturally embody the concept of priorities, albeit in a different way from the COM; moreover, the latter is strategy-proof. Therefore, the equivalence property we identify in Proposition 1 applies only to stable mechanisms, not to any "reasonably priority-based" mechanisms.

The second key driver behind Theorem 1 is a reduction and decomposition of strategy-proofness for individually rational mechanisms:

Proposition 2. *Let C_H be an arbitrary profile of choice functions and $F : \mathscr{P}_D \times \{C_H\} \to \mathscr{A}$ an individually rational mechanism. Then, F is strategy-proof if and only if there are no $\succ \in \mathscr{P}_D$ and $x \in X^G$ such that either*

$$F(\succ_D^{-x}, C_H) \succ_{\mathsf{d}(x)} F(\succ_D, C_H), \text{ or} \tag{1}$$

$$F(\succ_D, C_H) \succ_{\mathsf{d}(x)} F(\succ_D^{-x}, C_H) \text{ and } x \notin F(\succ_D, C_H). \tag{2}$$

The above proposition states that for an individually rational mechanism, strategy-proofness reduces to non-manipulability via two special classes of preference misreports. First, equation (1) represents the situation where $\mathsf{d}(x)$ can get strictly better off by *dropping* a contract x from her list of acceptable contracts, when her true preference is $\succ_d$. Second, if equation (2) holds, $\mathsf{d}(x)$ can profitably manipulate F by *adding* x to (an appropriate position in) her list of acceptable contracts, when her true preference is $\succ_{\mathsf{d}(x)}^{-x}$ and hence x is actually unacceptable. Proposition 2 guarantees that an individually rational mechanism is strategy-proof if it is immune to these two simple classes, even though there are various other possible manipulations.

To the best of our knowledge, a reduction of strategy-proofness to the above two classes is novel, while several similar results are known in various environments: Compared to the true preference, neither dropping nor adding x needs (i) to be a truncation (Ehlers, 2004; Roth and Rothblum, 1999; Roth and Vande Vate, 1991), (ii) to be an adjacent preference (Carroll, 2012; Sato, 2013a,b), or (iii) to maintain the upper-contour set (Chun and Yun, 2020; Roy and Sadhukhan, 2022). The most closely related to Proposition 2 above might be Kojima and Pathak (2009, Lemma 1), who show in the classic matching model that any manipulation of a stable mechanism by an *institution* can be mimicked by a dropping strategy. It should be noted, however, that the immunity to adding strategies is not redundant on the *agent* side in our generalized framework, as we will elaborate in Section 5.

Once we establish the above two propositions, the rest of the proof of Theorem 1 is rather straightforward. That is, Proposition 1 translates the restrictions imposed by respect for strong improvements and the IUC into the negations of (1) and (2), which in turn are necessary and sufficient for strategy-proofness by Proposition 2. For the "only if" part, the richness of $\mathscr{C}_H$ is necessary to negate (1) and (2) for all $x \in X^G$. Without richness, F can fail to be strategy-proof while satisfying the other two properties. For an extreme example, recall that respect for strong improvements and the IUC are trivially met if $\mathscr{C}_H$ is a singleton.

5 Respect for Weak Improvements on Various Domains

We have thus far focused on strong (dis)improvements of choice functions, and it allows our analysis on the least structured domains as possible. However, one might argue that respect for strong improvements are insufficient as a desideratum, since strong improvements are rather extreme as we have seen in Example 1. To address this concern, we now consider a weaker concept of improvements.

Taking two choice functions and a subset Y of contracts as given, we say that C_h is a *weak Y-improvement* over C'_h if

$$C_h(X) \neq C'_h(X) \Leftrightarrow \text{ there exists } y \in Y \text{ such that } y \in \left[C_h(X) - C'_h(X)\right].$$

Note that under the IRC, any strong Y-improvement over C'_h is also a weak Y-improvement over C'_h, but not vice versa. A profile $C_H = (C_h)_{h \in H}$ is a weak Y-improvement of $C'_H = (C'_h)_{h \in H}$ if every C_h is a weak Y-improvement over C'_h. When Y is a singleton, we refer to a weak $\{x\}$-improvement simply as a weak x-improvement.

A weak Y-improvement C_h differs from its baseline C'_h only when the former chooses a contract from Y whereas the latter does not. When $\mathsf{d}(Y) = \{d\}$ for some agent d, in particular, the former chooses her contracts from a wider variety of the menus than the latter does. It would thus be natural to argue such a change should be (intended to be) favorable for the agent d.[13] This leads us to defining respect for weak improvements as follows.

[13] In the same spirit, Afacan (2017) proposes an alternative definition of improvements based on choice functions. His definition is similar to but slightly different from our weak improvements, and as a consequence, our results do not hold with his definition. See Appendix H for details.

Definition 4. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to *respect weak improvements* if $F(\succ_D, C_H) \succeq_d F(\succ_D, C'_H)$ holds for all $d \in D$, $\succ_D \in \mathscr{P}_D$, and $C_H, C'_H \in \mathscr{C}_H$ such that C_H is a weak Y-improvement over C'_H for some $Y \subseteq X^G$ with $\mathsf{d}(Y) = \{d\}$. □

In supporting our definition, it should also be noted that weak improvements defined above reduce to the standard concepts in certain special cases: First, in the classic environment that we specified in Example 1, our weak improvements coincide with the classic improvements. Specifically, a classic choice function $C_h^{P,q}$ is a weak (d, h)-improvement over another $C_h^{Q,q}$ if and only if the priority order P is a classic improvement over Q for d. Second, it also boils down to a natural, order-based concept of improvements when the choice functions are induced by *slot-specific priorities* (Kominers and Sönmez, 2016).

Example 2 (slot-specific priorities)**.** Taking an institution h as given, let $X_h^G := \{x \in X^G : \mathsf{h}(x) = h\}$ be the contracts relevant to it and $q \in \mathbb{N}$ its quota. Also let $\mathbf{P} = (P_s)_{s=1,\ldots,q}$ be an ordered list of q priority orders for h, where each P_s is a linear order over $X_h^G \cup \{\varnothing\}$ and represents the priority for the s-th slot of h. The *slot-specific priorities* $\mathbf{P}$ induce a choice function $C_h^{\mathbf{P}}$ for h as follows:

- Given a menu X of contracts, let $X^1 := X \cap X_h^G$.
- For each $s = 1, \ldots, q$, recursively, let x_s^* be the best (possibly null) contract among $X^s \cup \{\varnothing\}$ according to P_s, and define $X^{s+1} := \{x \in X^s : \mathsf{d}(x) \neq \mathsf{d}(x_s^*)\}$ if $x_s^* \neq \varnothing$ whereas $X^{s+1} := X^s$ otherwise.
- The overall chosen set is defined to be $C_h^{\mathbf{P}}(X) := \{x_1^*, \ldots, x_{q_h}^*\} - \{\varnothing\}$.

When the setup is classic and $P_1 = \cdots = P_q = P$ for some P, this $C_h^{\mathbf{P}}$ is identical to the the classic choice function $C_h^{P,q}$ defined in Example 1.[14]

Comparing two lists of slot specific priorities for h, we say $\mathbf{P} = (P_s)_{s=1,\ldots,q}$ is an *unambiguous improvement over* $\mathbf{Q} = (Q_s)_{s=1,\ldots,q}$ for agent d if for all $s \in \{1, \ldots, q\}$,

- $x \; Q_s \; y \Rightarrow x \; P_s \; y$ and $x \; Q_s \; \varnothing \Rightarrow x \; P_s \; \varnothing$ for any $x \in X_d^G$ and $y \in X_{-d}^G$, and
- $z \; Q_s \; w \Leftrightarrow z \; P_s \; w$ and $z \; Q_s \; \varnothing \Leftrightarrow z \; P_s \; \varnothing$ for any $z, w \in X_{-d}^G$,

where $X_d^G := \{x \in X^G : \mathsf{d}(x) = d\}$ and $X_{-d}^G := X^G - X_d^G$.[15] Note that this is a generalization

[14] Strictly speaking, P is an order over $X_h^G \cup \{\varnothing\}$ here, while it is over $D \cup \{\varnothing\}$ in the classic setup. However, we can naturally identify $X_h = \{(d, h) : d \in D\}$ with D when the setup is classic.

[15] This definition slightly differs from the original by Kominers and Sönmez (2016) in that they do not require $z \; Q_s \; \varnothing \Leftrightarrow z \; P_s \; \varnothing$. If we take their original definition as it is, however, the COM actually does not respect unambiguous improvements, contradicting their Theorem 4. (See Appendix H for details.) Thus, our definition here would be the "right" one they intended to mean.

of the classic improvements we defined in Example 1: Suppose that the setup is classic and that $P_1 = \cdots = P_q = P$ and $Q_1 = \cdots = Q_q = Q$. Then, $\mathbf{P}$ is an unambiguous improvement over $\mathbf{Q}$ for d, if and only if P is a classic improvement over Q for d. Even outside the classic environment, our definition of weak improvements is equivalent to unambiguous improvements: Let $\mathbf{P}$ and $\mathbf{Q}$ be two lists of slot-specific priorities with the same quota q. Then, $C_h^{\mathbf{P}}$ is a weak $Y_{d,h}$-improvement over $C_h^{\mathbf{Q}}$, where $Y_{d,h} := \{y \in X^G : \mathsf{d}(y) = d \text{ and } \mathsf{h}(y) = h\}$, if and only if $\mathbf{P}$ is an unambiguous improvement over $\mathbf{Q}$ for d. □

Throughout the rest of this section, we investigate the relation between strategy-proofness and respect for weak improvements defined as above, under a number of different assumptions on $\mathscr{C}_H$. Specifically, we start with the unstructured domains and then add more structures step by step, thereby crystallizing the implication of each additional assumption on the relation between the two desiderata.

5.1 General Domains

In this subsection, we see that on the general unstructured domains, a stable mechanism does not necessarily respect weak improvements even if it is strategy-proof. To concisely present such an example, we now introduce the *cumulative offer mechanisms with precedence orders*. A *precedence order* is a bijection $\pi : D \to \{1, \ldots, |D|\}$. Roughly speaking, it specifies which agent should make an offer at each step in the following algorithm.

Definition 5. Given $(\succ_D, C_H)$ and a precedence order π, the *cumulative offer process with precedence order* π (for short, COP with π) computes a subset of contracts as follows.

- Initial condition: Let $D_0 := D$ and $O_0 := \emptyset$.
- Step $t \geq 1$: Let $d_t \in D_{t-1}$ be the agent with the smallest value of π among D_{t-1}; i.e., $\pi(d_t) = \min\{\pi(d) : d \in D_{t-1}\}$. Agent d_t offers her best contract, say x_t, among those remaining; i.e., x_t is the best among $X^G - O_{t-1}$. Let $O_t := O_{t-1} \cup \{x_t\}$ be the pool of contracts that have been offered up to this step. Among O_t, each institution h holds the best combination of contracts, $C_h(O_t)$. Lastly, let D_t be the set of agents for whom (i) no contract is currently held by any institution and (ii) not all acceptable contracts have been offered yet; i.e.,

$$D_t := \{d \in D : d \notin \mathsf{d}(C_H(O_t)) \text{ and } \mathrm{Ac}(\succ_d) - O_t \neq \emptyset\}.$$

(Recall that with our notation, $\mathsf{d}(C_H(O_t))$ denotes $\{\mathsf{d}(x) : x \in \cup_h C_h(O_t)\}$.) Proceed to step $t+1$ if D_t is non-empty, and terminate otherwise.

- Outcome: When the process terminates after step T, its outcome is $C_H(O_T)$.

The *cumulative offer mechanism with precedence order* π (for short, COM with π) assigns the outcome of the above process, assuming it is an allocation, to each $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$. In the sequel, $F^\star_\pi$ denotes the COM with precedence order π. □

Note that by definition, the COM with any precedence order is stable when it is well-defined as a mechanism (i.e., when the corresponding COP always outputs an allocation). Without further assumptions on C_H, however, the COM with a precedence order can fail to respect weak improvements even if it is well-defined, stable, and strategy-proof. The following is such an example.

Example 3. Let $D = \{d_1, d_2, d_3\}$, $H = \{h\}$, and $X^G = \{x_1, x_2, x_3, y_1, y_2\}$, where x_i and y_i are two contracts between d_i and h for each $i \in \{1, 2\}$ and x_3 is the only contract between d_3 and h. Define two preference relations, $\succ_h$ and $\succ'_h$, over the set $\mathscr{A}$ of all allocations by

$$\{x_1, x_2, x_3\} \succ_h \{y_1, y_2\} \succ_h \{y_1, x_3\} \succ_h \boldsymbol{\{y_2, x_3\}} \succ_h \boldsymbol{\{x_1\}} \succ_h \emptyset, \text{ and}$$
$$\{x_1, x_2, x_3\} \succ'_h \{y_1, y_2\} \succ'_h \{y_1, x_3\} \succ'_h \boldsymbol{\{x_1\}} \succ'_h \boldsymbol{\{y_2, x_3\}} \succ'_h \emptyset,$$

where all the subsets of X^G unspecified above are unacceptable. Let $\mathscr{C}_H$ be the minimal rich domain containing $\{C_h, C'_h\}$, where C_h and C'_h are the choice functions induced by $\succ_h$ and $\succ'_h$, respectively. The two choice functions disagree only at $X' = \{x_1, y_2, x_3\}$, where $C_h(X') = \{y_2, x_3\}$ and $C'_h(X') = \{x_1\}$. Hence, C_h is a weak x_3-improvement over C'_h.

Let $F^\star_\pi$ be the COM with π defined over $\mathscr{P}_D \times \mathscr{C}_H$, where π is the precedence order such that $(\pi(d_1), \pi(d_2), \pi(d_3)) = (3, 2, 1)$. As we do in Appendix B, one can confirm that this $F^\star_\pi$ is stable and strategy-proof. Nevertheless, it does *not* respect weak improvements: Let $\succ_D = (\succ_{d_1}, \succ_{d_2}, \succ_{d_3})$ be such that $x_1 \succ_{d_1} y_1 \succ_{d_1} \varnothing$, $y_2 \succ_{d_2} x_2 \succ_{d_2} \varnothing$, and $x_3 \succ_{d_3} \varnothing$. During the COP's with $(\succ_D, C_h)$ and $(\succ_D, C'_h)$, the agents offer (x_3, y_2, x_1, y_1) and (x_3, y_2, x_1, x_2), respectively, exactly in these orders. For agent d_3, $F^\star_\pi(\succ_D, C_h) = \{y_1, y_2\}$ is strictly less preferred to $F^\star_\pi(\succ_D, C'_h) = \{x_1, x_2, x_3\}$, although C_h is a weak x_3-improvement over C'_h as seen above. □

5.2 Observable Substitutability

Given our observation in the previous subsection, it would be natural to ask under what conditions strategy-proofness (in conjunction with stability) becomes sufficient for respect for weak improvements. In this subsection, we will see that *observable substitutability* of choice functions (Hatfield et al., 2021b) constitutes such a condition. In order to formally define this condition, we first introduce a few preliminary concepts.

Definition 6. An *offer process* is a finite sequence $(x_1, \ldots, x_n)$ of distinct contracts, and its range as a set (rather than a sequence) is denoted by $\mathbb{X}((x_1, \ldots, x_n)) := \{x_1, \ldots, x_n\}$. It is *observable at a profile C_H of choice functions* if for each $t \in \{1, \ldots, n-1\}$, agent $\mathsf{d}(x_{t+1})$ signs no non-null contract at $C_H(\{x_1, \ldots, x_t\})$; i.e.,

$$\mathsf{d}(x_{t+1}) \notin \mathsf{d}\left(C_H(\{x_1, \ldots, x_t\})\right) \text{ for each } t \in \{1, \ldots, n-1\}.$$

An offer process $(x_1, \ldots, x_n)$ is said to be *for institution h* if $\mathsf{h}(x_i) = h$ for all $i \in \{1, \ldots n\}$, and it is *observable at C_h* if $\mathsf{d}(x_{t+1}) \notin \mathsf{d}(C_h(\{x_1, \ldots, x_t\}))$ for each $t \in \{1, \ldots, n-1\}$. □

Roughly speaking, an offer process is observable if it arises during a COP with some (generalized) precedence order and some preference profile. As we define below, observable substitutability is a substitutability condition (i.e., the monotonicity of rejected sets) restricted on those possible paths of COPs. Recall that given C_h and C_H, the associated rejection functions are defined by $R_h(X) := X - C_h(X)$ and $R_H(X) := X - C_H(X) = \cap_h R_h(X)$, respectively.

Definition 7. A profile C_H of choice functions is said to be *observably substitutable* (for short, OS), if the associated rejection function R_H satisfies $R_H(\{x_1, \ldots, x_{n-1}\}) \subseteq R_H(\{x_1, \ldots, x_n\})$ for any offer process $(x_1, \ldots, x_n)$ that is observable at C_H.[16] A domain $\mathscr{C}_H$ of profiles of choice functions is said to be OS if every $C_H \in \mathscr{C}_H$ is OS. □

For our present purpose, OS has three implications. First, it ensures that the outcome of the COP is always an allocation and is independent of the choice of a precedence order (Hatfield et al., 2021b, Proposition 3); hence, we can omit the dependence on π and let $F^\star$

[16] This is equivalent to the following alternative definition: $C_H = (C_h)_{h \in H}$ is OS if every component C_h is so in the sense that the associated R_h meets $R_h(\{x_1, \ldots, x_{n-1}\}) \subseteq R_h(\{x_1, \ldots, x_n\})$ for any offer process $(x_1, \ldots, x_n)$ that is for h and is observable at C_h.

denote the *uniquely-defined* COM.[17] Second, it makes the COM the unique candidate for a stable and strategy-proof mechanism; i.e., if C_H is OS and some $F : \mathscr{P}_D \times \{C_H\} \to \mathscr{A}$ is both stable and strategy-proof, then F must be equal to $F^\star$ (Hatfield et al., 2021b, Theorem 1b). Third, when the choice functions are OS, the COM satisfies the duality in the sense of Proposition 1, without any further qualifications.

Fact 2. *Let C_H be an OS profile of choice functions. For any $\succ \in \mathscr{P}_D$ and $Y \subseteq X^G$, then, $F^\star\left(\succ, C_H^{-Y}\right) = F^\star\left(\succ_D^{-Y}, C_H\right)$.*[18]

In the case of an OS domain, we can strengthen Theorem 1 as follows. In contrast to Example 3 above, OS turns out to ensure that the COM should respect not only strong but weak improvements whenever it is strategy-proof. Together with Theorem 1 and the second implication of OS we mentioned above, this leads to Theorem 2 below. It would be worth emphasizing that the IUC is *not* redundant in this theorem: Even when the domain is OS and rich, the COM can respect improvements without being strategy-proof, as we will see in Example 4 below.

Theorem 2. *Suppose that $\mathscr{C}_H$ is an OS domain of profiles of choice functions and let $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ be a stable mechanism. Then, F respects weak improvements and satisfies the IUC if it is strategy-proof. When $\mathscr{C}_H$ is rich, the converse is also true: F respects weak improvements and satisfies the IUC (if and) only if it is strategy-proof.*

Example 4. Let $D = \{d_1, d_2, d_3\}$, $H = \{h\}$, and $X^G = \{x_i, y_i\}_{i\in\{1,2,3\}}$, where for each $i \in \{1, 2, 3\}$, x_i and y_i are two possible contracts between d_i and h. Let $\succ_h$ be a preference relation over allocations such that

$$\begin{aligned}\{x_1, x_2, x_3\} &\succ_h \{y_1, y_2, y_3\} \\ &\succ_h \{x_1, y_2\} \succ_h \{x_1, x_2\} \succ_h \{x_2, y_3\} \succ_h \{y_1, y_2\} \succ_h \{y_1, x_3\} \\ &\succ_h \text{[any other doubleton allocations]} \succ_h \text{[any singletons]} \succ_h \emptyset,\end{aligned}$$

where all tripletons except $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ are unacceptable, and the unspecified rankings among doubletons and among singletons are arbitrary. Let C_h be the choice

[17] For feasibility and order-independence of the COP, see also Flanagan (2014), Hatfield and Kominers (2019), Hirata and Kasuya (2014), Kominers and Sönmez (2016), and Zhang (2016).

[18] Given that C_H is OS, so is C_H^{-Y} for any Y as we demonstrate in Appendix E. Thus, the left-hand side of the equality is always well-defined.

function induced by $\succ_h$ and $\mathscr{C}_h := \{C_h^{-Y} : Y \subseteq X^G\}$ the minimal rich domain containing C_h. As we do in Appendix B, one can confirm that this $\mathscr{C}_H$ is an OS domain (which implies that the COM is uniquely defined and is stable) and that the COM defined on $\mathscr{P}_D \times \mathscr{C}_H$ respects strong improvements. Nevertheless, the COM is *not* strategy-proof on this domain, as it is not so at C_h. To see this, let $\succ_D$ be such that $y_1 \succ_{d_1} x_1 \succ_{d_1} \emptyset$, $x_2 \succ_{d_2} y_2 \succ_{d_2} \emptyset$, and $x_3 \succ_{d_3} \emptyset \succ_{d_3} y_3$. The outcome of the COM at $(\succ_D, C_H)$ is $\{y_1, y_2\}$, and d_3 is assigned the null contract. However, if d_3 reports $\rhd_d$ such that $y_3 \rhd_{d_3} x_3 \rhd_{d_3} \emptyset$, the outcome becomes $\{x_1, x_2, x_3\}$, which is strictly better in regard to his true preference $\succ_{d_3}$. □

The fact that the IUC is not redundant in Theorem 2 would lead to the following question: Under what condition does respect for improvements become fully equivalent to strategy-proofness? If such a condition is highly restrictive, there is the possibility for the use of non-strategy-proof mechanisms: If so, over a reasonably broad domain (that does not meet the restrictive condition), there can be a stable mechanism that respects improvements but is not strategy-proof. Such a mechanism might be appealing enough for those who judge that the IUC is not a desideratum by itself. Conversely, such a possibility is limited if the full equivalence holds under a mild condition: If so, no stable and non-strategy-proof mechanisms can respect improvements unless we restrict our attention to a narrow range of domains. As such, those mechanisms would be normatively undesirable, whether or not it is implementable. In the next two subsections, we investigate the above question and provide conditions for the full equivalence between respect for improvements and strategy-proofness.

5.3 Strong Observable Substitutability

In this subsection, we answer the above question while *restricting our candidate mechanism to the COM*. Although Theorem 3 below does not hold with arbitrary stable mechanisms, it would be a natural first step given the central role the COM has played in the literature. Indeed, we will heavily rely on it when we provide its counterpart for general stable mechanisms in the next subsection. To begin, we now define *strong observable substitutability* (for short, strong OS), strengthening the original OS.

Definition 8. A profile C_H of choice functions is said to be *strongly observably substitutable* (for short, strongly OS) if for any two processes $\mathbf{x}$ and $\mathbf{y}$ both observable at C_H, $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$

implies $R_H(\mathbb{X}(\mathbf{x})) \subseteq R_H(\mathbb{X}(\mathbf{y}))$.[19] A domain $\mathscr{C}_H$ of profiles of choice functions is said to be strongly OS if every $C_H \in \mathscr{C}_H$ is strongly OS. □

Strong OS is strong in comparing more pairs of offer processes than OS does. Namely, strong OS compares two observable processes even if they arise along different paths of COPs, whereas OS is relevant only when one is a subprocess of the other. For an example of a choice function that is OS but not strongly OS, refer back to Example 4 in the previous subsection. In that example, both $\mathbf{x} = (y_1, x_2, y_3)$ and $\mathbf{y} = (y_1, x_2, x_3, y_2, y_3)$ are observable at C_h, and $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ holds. However, the choice outcomes are $C_h(\{y_1, x_2, y_3\}) = \{x_2, y_3\}$ and $C_h(\{y_1, x_2, x_3, y_2, y_3\}) = \{y_1, y_2, y_3\}$. That is, C_h rejects y_1 from $\mathbb{X}(\mathbf{x})$ but not from $\mathbb{X}(\mathbf{y})$. Therefore, this choice function is not strongly OS, although it is OS as we demonstrate in Appendix B.

Concerning strong OS, two further remarks are in order. First, we could argue that it is relatively weak among the existing substitutability conditions: As we demonstrate in Appendix D, strong OS is weaker than unilateral substitutability (Hatfield and Kojima, 2010) and substitutable completability (Hatfield and Kominers, 2019); as a consequence, it is weak enough *not* to guarantee certain key structures of (the set of) stable allocations. Second, we can characterize the gap between OS and strong OS by whether or not the COM satisfies two monotonicity properties introduced by Kojima and Manea (2010): Given that a profile of choice functions is OS, it is also strongly OS if and only if the COM satisfies weak Maskin monotonicity, if and only if the COM satisfies IR monotonicity. For details including the definitions of the monotonicity properties, see Appendix G. See also Section 6.3 below for further discussions of substitutability conditions.

For our main purpose, the virtue of strong OS is in simplifying the condition for the COM to be strategy-proof. According to Proposition 2, a stable mechanism is strategy-proof if it is immune to "dropping" strategies and "adding" strategies. When choice functions are strongly OS, no agent can profitably manipulate the COM by adding a contract to her preference list, and hence, the COM is strategy-proof if it is not manipulable via dropping strategies. Combined with Fact 2 and Theorem 2, it then follows that strategy-proofness and respect for improvements (either weak or strong) are fully equivalent for the COM, without the presence of the IUC. We formally state these results as follows.

[19] This is equivalent to the following alternative definition: $C_H = (C_h)_{h \in H}$ is strongly OS if every component C_h is so in the sense that for any two offer processes $\mathbf{x}$ and $\mathbf{y}$ that are for h and are observable at C_h, $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ implies $R_h(\mathbb{X}(\mathbf{x})) \subseteq R_h(\mathbb{X}(\mathbf{y}))$.

Proposition 3. *Let C_H be a strongly OS profile of choice functions. Then, the cumulative offer mechanism at C_H, $F^\star(\cdot, C_H)$, is strategy-proof if and only if there are no $\succ_D \in \mathscr{P}_D$ and $x \in X^G$ such that $F^\star\left(\succ_D^{-x}, C_H\right) \succ_{\mathsf{d}(x)} F^\star(\succ_D, C_H)$.*

Theorem 3. *Let $\mathscr{C}_H$ be a rich and strongly OS domain of profiles of choice functions. Then, the following are all equivalent: (i) the cumulative offer mechanism $F^\star : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is strategy-proof, (ii) it respects weak improvements, and (iii) it respects strong improvements.*

5.4 Unilateral Substitutability

In this subsection, we establish the full equivalence between respect for improvements and strategy-proofness *for general stable mechanisms*. It should be noted we *cannot* replace the COM with an arbitrary stable mechanism in Theorem 3: Even when $\mathscr{C}_H$ is rich and strongly OS, a non-COM stable mechanism can respect weak improvements without being strategy-proof. Further, the same is possible even when the COM does not respect improvements; see Appendix I for such an example. Nevertheless, it becomes impossible once we strengthen strong OS to *unilateral substitutability* (Hatfield and Kojima, 2010).

Definition 9. A profile C_H of choice functions is said to be *unilaterally substitutable* (for short, US), if there are no $x, y \in X^G$ and $Z \subseteq X^G$ such that (i) $x \notin C_H(Z \cup \{x\})$, (ii) $x \in C_H(Z \cup \{x, y\})$, and (iii) $\mathsf{d}(x) \notin \mathsf{d}(Z)$.[20]

Theorem 4. *Let $\mathscr{C}_H$ be a rich and US domain of profiles of choice functions and $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ a stable mechanism. Then, the following are all equivalent: (i) F is strategy-proof, (ii) it respects strong improvements, and (iii) it respects weak improvements.*

To conclude our main analyses, let us now summarize Theorems 1–4. Theorems 1–2 essentially state that a stable mechanism respects improvements if and "almost" only if it is strategy-proof. They leave a general possibility that a stable and non-strategy-proof mechanism respects improvements when it fails to meet the IUC, which would be a natural invariance property. However, such a mechanism cannot be the COM when choice functions are strongly OS (Theorem 3). Moreover, the possibility is open only when choice functions are not unilaterally substitutable (Theorem 4). As a whole, those theorems would suggest the necessity of strategy-proofness for a stable mechanism to be normatively desirable.

[20] This is equivalent to the following alternative definition: $C_H = (C_h)_{h \in H}$ is US if every component C_h is so in the sense that there are no $x, y \in X^G$ and $Z \subseteq X^G$ that satisfy conditions (i)–(iii) with C_h instead of C_H.

6 Disussions and Extensions

6.1 Two-Dimensional Strategy-Proofness

Our definition of strategy-proofness implicitly assumes that the agents can manipulate a mechanism only through misreporting their preferences and that the choice functions are given and fixed for them. In some circumstances, however, the agents might be able to *deliberately disimprove* the choice functions. First, as we mentioned in the introduction, they could do so by hiding some relevant information. Suppose that a contract x represents a special arrangement for minority agents, but one's eligibility is based on *voluntary* disclosure of her ethnic background since the matching authority cannot force her to disclose such personal information for privacy reasons. Then, a minority agent can hide her eligibility for x if it is beneficial to do so, whereas it would be more difficult for a non-minority agent to feign the minority status. Second, the agents could manipulate the choice functions through *strategic incompetence*. Consider the classic model without contractual terms and suppose that the choice functions simply represent the institutions' preferences. Suppose further that so as to rank the agent, the institutions conduct interviews with them. In this scenario, an agent could *pretend to be incompetent* at her interview with h, thereby lowering her own position in h's preference. In particular, she could make herself unacceptable by doing sufficiently badly at the interview, although it would be difficult to lower her rank position exactly by an arbitrary number. In other words, agent d could make the choice function to be C_h^{-x}, where $x = (d, h)$, when it would be C_h if she does her best at her interview with h.

In those circumstances, then, agents could manipulate a mechanism not only by misreporting their preferences but also by disimproving the choice functions. Then, a natural requirement for a mechanism would be an immunity to such two-dimensional manipulations. For non-stable mechanisms, this requirement can be stronger than the standard strategy-proofness that precludes preference manipulations only. In light of Theorem 1, however, we can easily verify that the two requirements coincide for stable mechanisms.

Theorem 5. *For any domain $\mathscr{C}_H$ of profiles of choice functions, a stable mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is strategy-proof if and only if it is "two-dimensionally strategy-proof," in the sense that $F\left((\succ_d, \succ_{-d}), C_H\right) \succeq_d F\left((\rhd_d, \succ_{-d}), C_H^{-Y}\right)$ holds for any $d \in D$, $\succ_D \in \mathscr{P}_D$, $\rhd_d \in \mathscr{P}_d$, and $C_H^{-Y} \in \mathscr{C}_H$ such that $\mathsf{d}(Y) = \{d\}$.*

When the domain $\mathscr{C}_H$ is OS, we can strengthen the above theorem as follows: A stable mechanism F is strategy-proof if and only if $F((\succ_d, \succ_{-d}), C_H) \succeq_d F\left((\rhd_d, \succ_{-d}), C'_H\right)$ always holds for any $C_H, C'_H \in \mathscr{C}_H$ such that C_H is a weak Y-improvement and $\mathsf{d}(Y) = \{d\}$. That is, under OS, strategy-proofness rules out profitable two-dimensional manipulations, even if the agents can disimprove their priorities in the weak sense. We omit the proof for this statement since it is parallel to the one for Theorem 5.

6.2 Respect for Group Improvements

In this subsection, we examine collective improvements of priority structures *for groups of agents*. In our main analyses, we restricted our attention to improvements for a single agent, represented by Y-improvements (either strong or weak) such that $\mathsf{d}(Y)$ is a singleton. Practically speaking, however, most institutional changes affect multiple agents (i.e., $\mathsf{d}(Y)$ is not a singleton), and those changes are intended to make the target group $\mathsf{d}(Y)$ collectively better off. In what follows, we discuss three possible desiderata concerning the welfare consequences of such group improvements: Namely, as a result of a Y-improvement, the target group (i) should Pareto improve, (ii) should not weakly Pareto deteriorate , or (iii) should not strongly Pareto deteriorate.

The first requirement that every member of the target group $\mathsf{d}(Y)$ should be (weakly) better off is the strongest among the three and would be ideal for policymakers; however, it is apparently too strong. In general, the members of a target group compete with each other either directly or indirectly, and as a consequence, an improvement for one can have a negative effect on another within the same group. Notice that this remains the case even if the "relative ranking within the group" is well-defined and kept constant; see Kojima (2012, the proof of Theorem 2) for an example.

Actually, the second requirement is still too demanding. In Appendix F, we show that for a stable mechanism, this requirement is (almost) equivalent to *strong group strategy-proofness*, which necessitates that no group of agents can weakly Pareto improve by a joint manipulation. Since strong group strategy-proofness is almost impossible even in the classic environment (Ergin, 2002), precluding weak Pareto deterioration of target groups is also almost impossible. Unless we focus on a highly restrictive class of choice functions, there is a Y-improvement, either weak or strong, such that none of the target group $\mathsf{d}(Y)$ gets strictly better off while some of them get strictly worse off.

Therefore, a more reasonable requirement would be the third one that an improvement should not make all of the target group strictly worse off. In contrast to the other two, this requirement can be met with a broader class of choice functions. More specifically, the COM respects improvements for groups in this particular sense whenever it is strategy-proof and the choice functions are OS, as we formally state below:

Definition 10. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to *respect weak group improvements* if the following holds: For any $\succ_D \in \mathscr{P}_D$ and $C_H, C'_H \in \mathscr{C}_H$ such that C_H is a weak Y-improvement over C'_H, $F(\succ_D, C_H) \succeq_d F\left(\succ_D, C'_H\right)$ holds for some $d \in \mathsf{d}(Y)$. □

Theorem 6. *Let $\mathscr{C}_H$ be an OS domain of profiles of choice functions. If the cumulative offer mechanism $F^\star : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is strategy-proof, then, it respects weak group improvements.*

In relation to the literature, this theorem generalizes Theorem 2 of Hafalir et al. (2013). In the context of school choice, they compare the COM with and without minority reserves and show that an introduction of reserves makes at least one minority student weakly better off. It is easy to see that in their classic setup, the choice functions with and without reserves disagree only if there is a minority student who is admitted with the reserves but not without them. Therefore, introducing reserves is a weak Y-improvement with $\mathsf{d}(Y)$ being the set of all minority students. Combined with the fact that the COM is strategy-proof with and without reserves (Hafalir et al., 2013, Proposition 1), our Theorem 6 thus implies their Theorem 2. Actually, we can derive a stronger claim in their setup: At least one minority student gets weakly better by *any increase* in the number of reserved seats *from any initial numbers* (i.e., not necessarily from zeros), since any such increase is a weak improvement for minority students.

6.3 Sufficiency for Strategy-Proofness and Respect for Improvements

In this subsection, we consider the condition, *in terms of (the domain of) choice functions*, for a stable mechanism to respect improvements. In Sections 4–5, we identified the condition for respect for improvements in terms of another property of a mechanism, i.e., strategy-proofness. Concerning the condition in terms of choice functions, our results allow us to translate the existing results on strategy-proofness to those on respect for improvements:

On the one hand, it is well known that, depending on the choice functions, no stable mechanism may be strategy-proof. As a consequence, no stable mechanism may respect improvements, either. On the other hand, as we will discuss later in this subsection, the existing literature has established several sufficient conditions on choice functions that guarantee the strategy-proofness of the COM. According to our main results, the COM respects improvements, too, when choice functions satisfy those conditions. In addition to such immediate "translations," we also obtain a novel sufficient condition for the COM to respect improvements and to be strategy-proof, as a technical by-product of our main analyses.

As a subcondition in our sufficiency result, we introduce one more property for choice functions, *strong observable size-monotonicity* (for short, strong OSM). Parallel to the relation of strong OS to the original OS, strong OSM is a strengthening of the original observable size-monotonicity (for short, OSM) of Hatfield et al. (2021b). The original OSM only requires C_H to satisfy $\#C_H(\{x_1, \dots, x_{n-1}\}) \leq \#C_H(\{x_1, \dots, x_n\})$ for any observable process $(x_1, \dots, x_n)$ at C_H, whereas the following definition of strong OSM also compares $\mathbf{x}$ and $\mathbf{y}$ across different paths of offer processes.

Definition 11. A profile C_H of choice functions is said to be *strongly observably size-monotonic (for short, strongly OSM)* if for any two observable offer processes $\mathbf{x}$ and $\mathbf{y}$ at C_H, $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ implies $\#C_H(\mathbb{X}(\mathbf{x})) \leq \#C_H(\mathbb{X}(\mathbf{y}))$.[21] A domain $\mathscr{C}_H$ of profiles of choice functions is said to be strongly OSM if every $C_H \in \mathscr{C}_H$ is strongly OSM. □

As we formally state below, the combination of strong OS and strong OSM constitutes a sufficient condition for the COM to be strategy-proof and to respect weak improvements. It should be noted that strong OS and the original OSM are insufficient. This is because OSM is known to be insufficient for strategy-proofness even if it is combined with the original substitutability of Hatfield and Milgrom (2005), which is stronger than strong OS. See Hatfield et al. (2021b, Example 4) for such an example.[22]

[21] This is equivalent to the following alternative definition: $C_H = (C_h)_{h \in H}$ is strongly OSM if every component C_h is so in the sense that $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ implies $\#C_h(\mathbb{X}(\mathbf{x})) \leq \#C_h(\mathbb{X}(\mathbf{y}))$, for any two offer processes $\mathbf{x}$ and $\mathbf{y}$ that are for h and are observable at C_h.

[22] The choice function in their example indeed satisfies the Hatfield and Milgrom (2005) substitutability, even though they provide it as an example that meets OS and OSM.

Theorem 7. *Let $\mathscr{C}_H$ be a strongly OS and strongly OSM domain of profiles of choice functions. Then, the cumulative offer mechanism $F^\star : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is strategy-proof and respects weak improvements.*

Since the above theorem is applicable to the cases where $\mathscr{C}_H$ is a singleton, it provides a sufficient condition for the COM to be strategy-proof at a *fixed* profile of choice functions. As such, our condition is not the weakest in the existing literature: Hatfield et al. (2021b) provide the weakest sufficiency result to date that the COM is strategy-proof if each institution's choice function is OS, OSM, and *non-manipulable via contracual terms* (for short, NM).[23] Their condition as a whole is weaker than ours, because ours implies each of their three subconditions.

At the same time, our condition has several potential merits: First, ours is the weakest sufficient condition to date for the COM to be *group strategy-proof*. Second, even as a condition for individual strategy-proofness, ours is the weakest among those which are applicable even when there is only one institution, and the one-institution model could be more practically relevant than it appears. Third, in conjunction with Proposition 3, Theorem 7 crystallizes exactly what manipulations each subcondition does and needs to exclude. We discuss these potential merits in detail in Appendix C.

Before concluding this subsection, to be fair, we should note that the ideas of strong OS and strong OSM can be found in the previous literature: First, Hatfield et al. (2017) study a class of choice functions, which are induced by multiple divisions and *flexible allotments* in an institution, and they show that such choice functions satisfy all the three conditions of Hatfield et al. (2021b). In their proofs for OS and OSM, they actually show that those choice functions are strongly OS and strongly OSM, although not explicitly stated so. Second, Schlegel (2020) shows that OS and OSM are sufficient (without NM) for strategy-proofness in a generalized version of matching with salaries (Echenique, 2012; Kelso and Crawford, 1982; Schlegel, 2015), where contractual terms are linearly ordered and preferences are restricted to be monotonic in regard to that order. In his proof, he exploits the fact that OS and OSM become equivalent to strong OS and strong OSM, respectively, in his environment. To some extent, thus, the usefulness of strong OS and strong OSM should have been recognized already. As a condition for strategy-proofness, the technical

[23] For other existing sufficient conditions, see also Aygün and Turhan (2019),Hatfield and Kojima (2010),Hatfield and Kominers (2019), Hatfield and Milgrom (2005), Kominers and Sönmez (2016), Sönmez (2013), and Sönmez and Switzer (2013).

	Assumption	Proofs for
Section 7.1	None	Fact 1; Propositions 1 and 2; Theorems 1 and 5
Section 7.2	OS	Fact 2; Theorems 2 and 6
Section 7.3	Strong OS	Proposition 3; Theorems 3 and 7
Section 7.4	US	Theorem 4

Table 1: Organization of Section 7

contribution of Theorem 7 would lie in that we distill those two conditions as a separate property for general choice functions and in that we establish their sufficiency for strategy-proofness in the general model of matching with contracts (i.e., with the unrestricted, possibly non-monotonic preference domain).

7 Proofs

In this section, we provide the proofs for the Facts, Propositions, and Theorems that we have presented above. In doing so, we categorize the proofs by the assumptions we make on $\mathscr{C}_H$. In Section 7.1, we prove the results that hold on unstructured domains. In Sections 7.2–7.4, we present the proofs with OS, strong OS, and US, respectively. The organization of this section is summarized in Table 1. At the beginning of each subsection, we also present some additional definitions and lemmas that we use in the subsequent proofs. The proofs of those lemmas are all relegated to Appendix A.

7.1 Proofs with General Domains

In this subsection, we provide the proofs for Fact 1, Propositions 1 and 2, and Theorems 1 and 5. Although our results are stated in terms of a mechanism defined on $\mathscr{P}_D \times \mathscr{C}_H$, we will often work with its restrictions, taking C_H or $\succ_D$ as fixed. It is thus useful to define some terminology regarding those restrictions: We call a mapping $f : \mathscr{P}_D \to \mathscr{A}$ a *D-mechanism* and $\varphi : \mathscr{C}_H \to \mathscr{A}$ an *H-mechanism*. A D-mechanism f is said to be stable at C_H if $f(\succ_D)$ is stable at $(\succ_D, C_H)$ for all $\succ_D \in \mathscr{P}_D$. It is said to be strategy-proof if $f(\succ_d, \succ_{-d}) \succeq_d f(\rhd_d, \succ_{-d})$ for all $\succ_d$, $\rhd_d$, and $\succ_{-d}$. The definitions for an H-mechanism to be stable, to respect improvements, and to satisfy the IUC are analogous and thus omitted.

Before we proceed, let us also introduce two lemmas: The first is the equivalence, in terms of the set of stable allocations, between dropping of Y from preferences and a strong Y-disimprovement of choice functions. The second is an H-mechanism counterpart of the uniqueness result for D-mechanisms by Hirata and Kasuya (2017, Theorem 1). The proofs of these lemmas are relegated to Appendix A.

Lemma 1. *Let C_H be an arbitrary profile of choice functions. For any $\succ_D \in \mathscr{P}_D$ and $Y \subseteq X^G$, an allocation $X \in \mathscr{A}$ is stable at $\left(\succ_D^{-Y}, C_H\right)$ if and only if it is stable at $\left(\succ_D, C_H^{-Y}\right)$.*

Lemma 2. *Let $\mathscr{C}_H$ be a rich domain of profiles of choice functions and $\succ_D \in \mathscr{P}_D$ an arbitrary preference profile. If two H-mechanisms $\phi, \psi : \mathscr{C}_H \to \mathscr{A}$ are stable, respect strong improvements, and satisfy the IUC at $\succ_D$, then, $\phi(C_H) = \psi(C_H)$ for all $C_H \in \mathscr{C}_H$.*

7.1.1 Proof of Fact 1

This is an immediate corollary of the proof of Proposition 1 below. ■

7.1.2 Proof of Proposition 1

To show the "if" part, suppose that $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is stable and strategy-proof. Arbitrarily fix $C_H, C_H^{-Y} \in \mathscr{C}_H$. Define two D-mechanisms, f and g, by $f(\succ_D) := F\left(\succ_D, C_H^{-Y}\right)$ and $g(\succ_D) := F\left(\succ_D^{-Y}, C_H\right)$ for each $\succ_D \in \mathscr{P}_D$. By assumption, f is stable at C_H^{-Y} and strategy-proof. Since each $g(\succ_D)$ is stable at $\left(\succ_D^{-Y}, C_H\right)$ by assumption, it is so at $\left(\succ_D, C_H^{-Y}\right)$ by Lemma 1; i.e., g is a stable D-mechanism at C_H^{-Y}. Further, for any $d \in D$, $\succ_d, \rhd_d \in \mathscr{P}_d$, and $\succ_{-d} \in \prod_{d' \neq d} \mathscr{P}_{d'}$, the strategy-proofness of F implies

$$g(\succ_d, \succ_{-d}) = F\left(\left(\succ_d^{-Y}, \succ_{-d}^{-Y}\right), C_H\right) \succeq_d F\left(\left(\rhd_d^{-Y}, \succ_{-d}^{-Y}\right), C_H\right) = g\left(\rhd_d, \succ_{-d}\right),$$

since otherwise d could profitably manipulate F by reporting $\rhd_d^{-Y}$ when the true preference is $\succ_d^{-Y}$. That is, g is also strategy-proof. Then f and g must coincide, since at most one D-mechanism can be both stable and strategy-proof at an arbitrary profile of choice functions (Hirata and Kasuya, 2017, Theorem 1).[24] Since C_H and C_H^{-Y} are arbitrary,

[24] Strictly speaking, Hirata and Kasuya (2017) establish their theorem on the standard preference domain, where a preference also ranks unacceptable contracts. However, one can easily check that their proof remains

$F\left(\succ_D, C_H^{-Y}\right) = F\left(\succ_D^{-Y}, C_H\right)$ holds for any $\succ_D \in \mathscr{P}_D$ and any $C_H, C_H^{-Y} \in \mathscr{C}_H$.

To show the "only if" part, suppose that $\mathscr{C}_H$ is rich and that $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is stable, respects strong improvements, and satisfies the IUC. To begin with, arbitrarily fix $\succ_D \in \mathscr{P}_D$ and $Y \subseteq X^G$. By the richness assumption, $C_H^{-Y} \in \mathscr{C}_H$ for each $C_H \in \mathscr{C}_H$, and hence, we can define an H-mechanism $\phi : \mathscr{C}_H \to \mathscr{A}$ by $\phi(C_H) := F(\succ_D, C_H^{-Y})$ for each $C_H \in \mathscr{C}_H$. In what follows, we show that ϕ is stable, respects strong improvements, and satisfies the IUC at $\succ_D^{-Y}$. First, $\phi(C_H)$ is always stable at $(\succ_D^{-Y}, C_H)$, by the stability of F and Lemma 1. Second, to confirm that ϕ respects strong improvements at $\succ_D^{-Y}$, arbitrarily fix $d \in D$ and $Z \subseteq X^G$ such that $\mathsf{d}(Z) = \{d\}$. What we need to show is $\phi(C_H) \succeq_d^{-Y} \phi\left(C_H^{-Z}\right)$ for any $C_H \in \mathscr{C}_H$. On the one hand, since F is assumed to respect strong improvements,

$$\phi(C_H) = F\left(\succ_D, C_H^{-Y}\right) \succeq_d F\left(\succ_D, C_H^{-(Y\cup Z)}\right) = \phi\left(C_H^{-Z}\right),$$

for any $C_H \in \mathscr{C}_H$. On the other hand, since $\phi(C_H)$ must be disjoint from Y by the stability of F, two preferences $\succ_d$ and $\succ_d^{-Y}$ agree on the ranking between $\phi(C_H)$ and $\phi\left(C_H^{-Z}\right)$. These together entail $\phi(C_H) \succeq_d^{-Y} \phi\left(C_H^{-Z}\right)$ for any $C_H \in \mathscr{C}_H$. Lastly, to check the IUC, suppose

$$F\left(\succ_D, C_H^{-(Y\cup Z)}\right) = \phi\left(C_H^{-Z}\right) \neq_d \phi\left(C_H\right) = F\left(\succ_D, C_H^{-Y}\right),$$

for some $Z \subseteq X^G$ such that $\mathsf{d}(Z) = \{d\}$. By the IUC of F, the right-hand side must contain some $z \in Z$, and hence, ϕ satisfies the IUC.

We have thus far established that ϕ is stable, respects strong improvements, and satisfies the IUC at $\succ_D^{-Y}$. By the assumptions on F, the same is also true for $\psi : C_H \mapsto F(\succ_D^{-Y}, C_H)$. By Lemma 2, therefore, these two H-mechanisms must coincide; that is, $F\left(\succ_D, C_H^{-Y}\right) = F\left(\succ_D^{-Y}, C_H\right)$ for any $C_H \in \mathscr{C}_H$. ■

7.1.3 Proof of Proposition 2

Since the "only if" part is trivial, we only establish the "if" part. To simplify the notation, arbitrarily fix C_H and let $f(\cdot)$ denote $F(\cdot, C_H)$. Suppose that for any $\succ_D \in \mathscr{P}_D$ and $x \in X^G$, neither (1) nor (2) in the statement of this proposition holds; equivalently, both of the

valid with our current definition of preferences.

following do hold:

$$f(\succ_D) \succeq_{\mathsf{d}(x)} f\left(\succ_D^{-x}\right), \qquad \text{and} \tag{3}$$

$$\left[f\left(\succ_D^{-x}\right) \succeq_{\mathsf{d}(x)} f(\succ_D) \text{ or } x \in f(\succ_D)\right]. \tag{4}$$

We then have $x \notin f(\succ_D) \Rightarrow f\left(\succ_D^{-x}\right) =_{\mathsf{d}(x)} f(\succ_D)$ for any $\succ_D$ and x. Repeatedly applying the same argument, for any $\succ_D \in \mathscr{P}_D$, $d \in D$, and $Z \subseteq X^G$ such that $\mathsf{d}(Z) = \{d\}$, we have

$$\left[Z \cap f(\succ_D) = \emptyset\right] \Longrightarrow \left[f\left(\succ_d^{-Z}, \succ_{-d}\right) =_d f(\succ_D)\right]. \tag{5}$$

Now, arbitrarily fix $\succ_D \in \mathscr{P}_D$, $d \in D$, and $\rhd_d \in \mathscr{P}_d$, and define $y_\succ := \mathsf{x}(d, f(\succ_D))$ and $y_\rhd := \mathsf{x}(d, f(\rhd_d, \succ_{-d}))$. To establish strategy-proofness, it suffices to show $y_\succ \succeq_d y_\rhd$. This is immediate from individual rationality of f if $y_\rhd$ is null or unacceptable for $\succ_d$. In what follows, thus, assume $y_\rhd \succ_d \varnothing$. To begin, define $Y_1 := (\mathrm{Ac}(\succ_d) \cup \mathrm{Ac}(\rhd_d)) - \{y_\rhd\}$ so that $\succ_d^{-Y_1}$ and $\rhd_d^{-Y_1}$ coincide; namely, both preferences refer to the one such that only $y_\rhd$ is acceptable. Since $Y_1 \cap f(\rhd_d, \succ_{-d}) = \emptyset$ by definitions, equation (5) entails

$$f(\rhd_d, \succ_{-d}) =_d f\left(\rhd_d^{-Y_1}, \succ_{-d}\right) \equiv f\left(\succ_d^{-Y_1}, \succ_{-d}\right), \tag{6}$$

and hence, $y_\rhd \in f\left(\succ_d^{-Y_1}, \succ_{-d}\right)$. Next, let $Y_2 := \mathrm{Ac}(\succ_d) - \{y_\succ, y_\rhd\}$ so that $\succ_d^{-Y_2}$ is a preference such that only $y_\succ$ and $y_\rhd$ are acceptable. Since $Y_2 \cap f(\succ_D) = \emptyset$, equation (5) entails

$$f(\succ_D) =_d f\left(\succ_d^{-Y_2}, \succ_{-d}\right), \tag{7}$$

and hence, $y_\succ \in f\left(\succ_d^{-Y_2}, \succ_{-d}\right)$. Substituting $\left(\succ_d^{-Y_2}, \succ_{-d}\right)$ and $y_\succ$ into $\succ_D$ and x in equation (3) above, it follows that

$$f\left(\succ_d^{-Y_2}, \succ_{-d}\right) \succeq_d^{-Y_2} f\left(\left(\succ_d^{-Y_2}\right)^{-y_\succ}, \succ_{-d}\right) \equiv f\left(\succ_d^{-Y_1}, \succ_{-d}\right), \tag{8}$$

since by definitions, $Y_1 = Y_2 \cup \{y_\succ\}$ and hence $\left(\succ_d^{-Y_2}\right)^{-y_\succ} = \succ_d^{-Y_1}$. Combining equations (6)–(8), we obtain $y_\succ \succeq_d y_\rhd$ as desired. ■

7.1.4 Proof of Theorem 1

First, let F be a stable and strategy-proof mechanism, and arbitrarily fix $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$. The strategy-proofness of F implies that for any $Y \subseteq X^G$ such that $C_H^{-Y} \in \mathscr{C}_H$ and $\mathsf{d}(Y) = \{d\}$, we have

$$F(\succ_D, C_H) \succeq_d F\left(\left(\succ_d^{-Y}, \succ_{-d}\right), C_H\right) = F\left(\succ_D, C_H^{-Y}\right),$$

where the equality holds by Proposition 1; i.e., F respects strong improvements. To establish the IUC, suppose $C_H^{-Y} \in \mathscr{C}_H$, $Y \cap F(\succ_D, C_H) = \emptyset$, and $\mathsf{d}(Y) = \{d\}$ for some $Y \subseteq X^G$ and $d \in D$. For d not to benefit by reporting $\succ_d$ when the true preference is $\succ_d^{-Y}$, we must have

$$F(\succ_D, C_H) \preceq_d F\left(\left(\succ_d^{-Y}, \succ_{-d}\right), C_H\right) = F\left(\succ_D, C_H^{-Y}\right),$$

where the equality is again by Proposition 1. Since F respects strong improvements as shown above, $F\left(\succ_D, C_H^{-Y}\right) =_{\mathsf{d}(x)} F(\succ_D, C_H)$ must hold.

Next, let $\mathscr{C}_H$ be a rich domain and suppose that $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is stable, respects strong improvements, and satisfies the IUC. Arbitrarily fix $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$ and $x \in X^G$. Note that by the richness assumption, $C_H^{-x} \in \mathscr{C}_H$. Since F respects strong improvements,

$$F(\succ_D, C_H) \succeq_{\mathsf{d}(x)} F\left(\succ_D, C_H^{-x}\right) = F\left(\succ_D^{-x}, C_H\right), \tag{9}$$

where the equality follows from Proposition 1. As F also satisfies the IUC, this further implies

$$\left[F(\succ_D, C_H) \succ_{\mathsf{d}(x)} F\left(\succ_D^{-x}, C_H\right)\right] \Rightarrow x \in F(\succ_D, C_H). \tag{10}$$

Since $\succ_D$, C_H, and x are all arbitrary, equations (9)–(10) ensure via Proposition 2 the strategy-proofness of F. ■

7.1.5 Proof of Theorem 5

The "if" part is trivial by definition. To show the "only if" part, suppose that F is stable and strategy-proof. Then, F respects strong improvements by Theorem 1. For any $(\succ_D, C_H)$, d,

$\rhd_d$, and Y such that $C_H^{-Y} \in \mathscr{C}_H$ and $\mathsf{d}(Y) = \{d\}$, thus, we have

$$F(\succ_D, C_H) \succeq_d F\left(\succ_D, C_H^{-Y}\right) \succeq_d F\left((\rhd_d, \succ_{-d}), C_H^{-Y}\right),$$

where the first and second preferences follow, respectively, from respect for improvements and strategy-proofness. ■

7.2 Proofs with OS Domains

This subsection contains the proofs for Fact 2 and Theorems 2 and 6. In the proof of Theorem 6, we will exploit the following lemma, which can be seen as a very weak form of non-bossiness. Suppose that C_H is OS and the COM is strategy-proof. This lemma then states that for an agent to affect its outcome at all, she needs to misreport the upper contour set of what she obtains under truth-telling. Note that it is *not* necessarily immediate from strategy-proofness, because it excludes the possibility of affecting the outcomes of *others*. The proof of this lemma is relegated to Appendix A.

Lemma 3. *Let C_H be an OS profile of choice functions, and suppose that the cumulative offer mechanism at C_H, denoted by $f^\star(\cdot) := F^\star(\cdot, C_H)$, is strategy-proof. For any $\succ_D, \rhd_D \in \mathscr{P}_D$, then, $f^\star(\succ_D) = f^\star(\rhd_D)$ holds if there is $d \in D$ such that*

- *$f^\star(\succ_D) \unrhd_d \emptyset$,*
- *$\{x \in X^G : x \rhd_d f^\star(\succ_D)\} = \{x \in X^G : x \succ_d f^\star(\succ_D)\}$, and*
- *$\rhd_{d'} = \succ_{d'}$ for all $d' \in D - \{d\}$.*

Before we proceed, we need to introduce one more definition from Hatfield et al. (2021b): We say that an offer process is complete at $(\succ_D, C_H)$ if it is the outcome of the COP with $(\succ_D, C_H)$ and some (generalized) precedence order. More formally, $\mathbf{x} = (x_1, \ldots, x_n)$ is *complete at* $(\succ_D, C_H)$ if it is observable at C_H and satisfies the following:

- x_i is acceptable for $\succ_{\mathsf{d}(x_i)}$ for any $i \in \{1, \ldots, n\}$,
- $i < j \Leftrightarrow x_i \succ_d x_j$ for all $d \in D$ and $i, j \in \{1, \ldots, n\}$ with $\mathsf{d}(x_i) = \mathsf{d}\left(x_j\right) = d$, and
- $\mathrm{Ac}(\succ_d) \subset \{x_1, \ldots, x_n\}$ if d signs no contract at $C_h(\{x_1, \ldots, x_n\})$, for all $d \in D$.

7.2.1 Proof of Fact 2

The proof is immediate and thus is omitted. ■

7.2.2 Proof of Theorem 2

Recall that under OS, a stable mechanism is strategy-proof only if it is the COM (Hatfield et al., 2021b). Given Theorem 1, thus, it suffices to establish that the COM respects weak improvements assuming it is strategy-proof, which is a special case of Theorem 6. As we will establish this stronger claim below, we omit the proof here. ■

7.2.3 Proof of Theorem 6

Suppose that $\mathscr{C}_H$ is OS and the COM, $F^\star : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$, is strategy-proof. Towards a contradiction, suppose that $\left(\succ_D, C_H, C'_H, Y\right)$ is a counterexample against respect for weak group improvements; i.e., C_H is a weak Y-improvement over C'_H and $F^\star(\succ_D, C'_H) \succ_d F^\star(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$. Taking C'_H as given, define the "size" of a preference profile $\succ'_D \in \mathscr{P}_D$ by

$$\sigma(\succ'_D) := \sum_{d \in \mathsf{d}(Y)} \left|\left\{x \in X^G : x \succ_d F^\star\left(\succ'_D, C'_H\right)\right\}\right|.$$

Without any loss of generality, assume further that $\succ_D$ in our counterexample is "minimal" with respect to this σ; i.e., for any $\succ'_D \in \mathscr{P}_D$ such that $\left(\succ'_D, C_H, C'_H, Y\right)$ also constitutes a counterexample, we have $\sigma\left(\succ'_D\right) \geq \sigma(\succ_D)$. In what follows, let $\mathbf{x} = (x_1, \ldots, x_T)$ be a complete offer process at $\left(\succ_D, C'_H\right)$. Since $F^\star\left(\succ_D, C'_H\right) \neq F^\star(\succ_D, C_H)$, there must be the first step $t^* < T$ along the offer process $\mathbf{x}$ at which C_H and C'_H disagree; that is, $C_H(\{x_1, \ldots, x_\tau\}) = C'_H(\{x_1, \ldots, x_\tau\})$ for all $\tau \in \{1, \ldots, t^* - 1\}$, and $C_H(\{x_1, \ldots, x_{t^*}\}) \neq C'_H(\{x_1, \ldots, x_{t^*}\})$.

To begin, consider the case where $Z := \{x_1, \ldots, x_{t^*-1}\} - C'_H(\{x_1, \ldots, x_{t^*-1}\})$ is non-empty. Construct a subsequence $\widetilde{\mathbf{x}} = (\widetilde{x}_1, \ldots, \widetilde{x}_{T-|Z|})$ of $\mathbf{x}$ by removing the contracts in Z. More formally, $\widetilde{\mathbf{x}}$ is the unique sequence such that (i) $\mathbb{X}\left(\widetilde{\mathbf{x}}\right) = \mathbb{X}(\mathbf{x}) - Z$ and (ii) $\widetilde{x}_k = x_\ell$ and $\widetilde{x}_{k'} = x_{\ell'}$ imply $k > k' \Leftrightarrow \ell > \ell'$. Notice that $\widetilde{\mathbf{x}}$ is observable and complete at $\left(\succ_D^{-Z}, C'_H\right)$.[25]

[25] To check observability, we need to confirm $\mathsf{d}\left(\widetilde{x}_{\tau+1}\right) \notin \mathsf{d}\left(C'_H(\{\widetilde{x}_1, \ldots, \widetilde{x}_\tau\})\right)$ for all $\tau \leq T - |Z| - 1$. For $\tau < t^* - |Z|$, it vacuously holds because $\{\widetilde{x}_1, \ldots, \widetilde{x}_{\tau+1}\}$ contains at most one contract for each agent. For $\tau \geq t^* - |Z|$, first note that $\{\widetilde{x}_1, \ldots, \widetilde{x}_\tau\} \cup Z = \{x_1, \ldots, x_{\tau+|Z|}\}$ and $\widetilde{x}_{\tau+1} = x_{\tau+|Z|+1}$. Moreover, $C'_H(\{\widetilde{x}_1, \ldots, \widetilde{x}_\tau\}) = C'_H(\{x_1, \ldots, x_{\tau+|Z|}\})$ must follow from the IRC since, by OS, the right-hand side contains no contract from Z. Combining these observations, we can conclude $\mathsf{d}\left(\widetilde{x}_{\tau+1}\right) \notin \mathsf{d}\left(C'_H(\{\widetilde{x}_1, \ldots, \widetilde{x}_\tau\})\right)$, as otherwise it contradicts the observability of $\mathbf{x}$.

By the definition of Z and the assumption of OS, $C'_H(\mathbb{X}(\mathbf{x}))$ should contain no contract from Z, and thus, $C'_H(\mathbb{X}(\widetilde{\mathbf{x}})) = C'_H(\mathbb{X}(\mathbf{x}))$ by the IRC. Therefore, we must have

$$F^\star\left(\succ_D^{-Z}, C'_H\right) = C'_H\left(\mathbb{X}\left(\widetilde{\mathbf{x}}\right)\right) = C'_H(\mathbb{X}(\mathbf{x})) = F^\star\left(\succ_D, C'_H\right).$$

Moreover, since Z is also equal to $\{x_1, \ldots, x_{t^*-1}\} - C_H(\{x_1, \ldots, x_{t^*-1}\})$ by the definition of t^*, we can follow parallel arguments with C_H so that we obtain $F^\star\left(\succ_D^{-Z}, C_H\right) = F^\star\left(\succ_D, C_H\right)$. As a consequence, if Z is non-empty, $\left(\succ_D^{-Z}, C_H, C'_H, Y\right)$ constitutes another counterexample to the claim of the theorem, while $\sigma\left(\succ_D^{-Z}\right) < \sigma\left(\succ_D\right)$ clearly holds. However, these contradict the assumption that we have chosen $\succ_D$ to be a minimal counterexample.

Next, consider the case where Z is empty; i.e., C'_H rejects no contract from $\{x_1, \ldots, x_{t^*-1}\}$. Arbitrarily fix $y \in Y$ such that $y \in C_H(\{x_1, \ldots, x_{t^*}\}) - C'_H(\{x_1, \ldots, x_{t^*}\})$. Such y must exist since, by assumptions, $C_H(\{x_1, \ldots, x_{t^*}\}) \neq C'_H(\{x_1, \ldots, x_{t^*}\})$ and C_H is a weak Y-improvement over C'_H. Notice that $\mathsf{d}(y)$ should strictly prefer y to the final outcome, $C'_H(\mathbb{X}(\mathbf{x}))$, since y is rejected by C'_H along $\mathbf{x}$; combined with the assumption of $F^\star\left(\succ_D, C'_H\right) \succ_d F^\star(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$, thus,

$$y \succ_{\mathsf{d}(y)} F^\star\left(\succ_D, C'_H\right) \succ_{\mathsf{d}(y)} F^\star(\succ_D, C_H).$$

We can then construct a distinct preference profile $\rhd_D$ from $\succ_D$, by lowering the ranking of y to somewhere between (the contracts $\mathsf{d}(y)$ signs at) $F^\star\left(\succ_D, C'_H\right)$ and $F^\star\left(\succ_D, C_H\right)$. More formally, $\rhd_D$ is a preference profile such that

- $\mathrm{Ac}\left(\rhd_{\mathsf{d}(y)}\right) = \mathrm{Ac}\left(\succ_{\mathsf{d}(y)}\right)$,
- $w \rhd_{\mathsf{d}(y)} w' \Leftrightarrow w \succ_{\mathsf{d}(y)} w'$ for any $w, w' \in \mathrm{Ac}\left(\rhd_{\mathsf{d}(y)}\right) - \{y\}$,
- $F^\star(\succ_D, C'_H) \rhd_{\mathsf{d}(y)} y \rhd_{\mathsf{d}(y)} F^\star(\succ_D, C_H)$, and
- $\rhd_{d'} = \succ_{d'}$ for all $d' \neq \mathsf{d}(y)$.

Note that $\succ_D$ and $\rhd_D$ satisfy all the conditions in Lemma 3 with $d = \mathsf{d}(y)$, and hence, $F^\star(\rhd_D, C_H) = F^\star(\succ_D, C_H)$. Then, a contradiction occurs if $F^\star\left(\rhd_D, C'_H\right) = F^\star\left(\succ_D, C'_H\right)$: If so, $\left(\rhd_D, C_H, C'_H, Y\right)$ also constitutes a counterexample. Furthermore, the same equality also implies $\sigma\left(\rhd_D\right) = \sigma\left(\succ_D\right) - 1$, contradicting the minimality assumption. To complete the proof, thus, it suffices to establish $F^\star\left(\rhd_D, C'_H\right) = F^\star\left(\succ_D, C'_H\right)$.

To do so, let $\tau^* \in \{1, \ldots, t^* - 1\}$ be the step at which $y = x_{\tau^*}$ is offered along the above-defined process $\mathbf{x}$, which is complete at $(\succ_D, C'_H)$. Define $\mathbf{z} = (z_1, \dots, z_{T-1})$ be the subsequence of $\mathbf{x}$ such that $\mathbb{X}(\mathbf{z}) = \mathbb{X}(\mathbf{x}) - \{y\}$; i.e., $z_t = x_t$ if $t < \tau^*$, and $z_t = x_{t+1}$ otherwise. In what follows, we confirm that $\mathbf{z}$ is observable, i.e.,$\mathsf{d}(z_{t+1}) \notin \mathsf{d}\left(C'_H(\{z_1, \dots, z_t\})\right)$ for all $t \in \{1, \dots, T-2\}$, and that it is complete at $(\rhd_D, C'_H)$. First, recall that C'_H rejects no contract from $\{x_1, \dots, x_{t^*-1}\}$ by assumption, and hence, $\{z_1, \dots, z_{t^*-1}\} = \{x_1, \dots, x_{t^*}\} - \{y\}$ contains at most one contract for each agent. For any $t < t^* - 1$, thus, $\mathsf{d}(z_{t+1}) \notin \mathsf{d}\left(C'_H(\{z_1, \dots, z_t\})\right)$ trivially holds. For $t \geq t^* - 1$, the same follows from the observability of $\mathbf{x}$, because we have $z_{t+1} = x_{t+2}$ and $C'_H(\{z_1, \dots, z_t\}) = C'_H(\{x_1, \dots, x_{t+1}\})$.[26] That is, $\mathbf{z}$ is observable. Next, note that the original assumption of $F^\star\left(\succ_D, C'_H\right) \succ_{\mathsf{d}(y)} F^\star(\succ_D, C_H)$ implies that $\mathsf{d}(y)$ holds a non-null contract at $F^\star\left(\succ_D, C'_H\right) = C'_H(\mathbb{X}(\mathbf{x}))$. Since $C'_H(\mathbb{X}(\mathbf{z})) = C'_H(\mathbb{X}(\mathbf{x}))$ by the IRC, the completeness of $\mathbf{z}$ at $\left(\rhd_D, C'_H\right)$ is immediate from that of $\mathbf{x}$ at $\left(\succ_D, C'_H\right)$. The observability and completeness of $\mathbf{z}$ entails $F^\star\left(\rhd_D, C'_H\right) = C'_H(\mathbb{X}(\mathbf{z}))$. Combined with $C'_H(\mathbb{X}(\mathbf{z})) = C'_H(\mathbb{X}(\mathbf{x})) = F^\star\left(\succ_D, C'_H\right)$, we obtain $F^\star\left(\rhd_D, C'_H\right) = F^\star\left(\succ_D, C'_H\right)$ as desired. ■

7.3 Proofs with Strongly OS Domains

In this subsection, we prove Proposition 3 and Theorems 3 and 7. In doing so, we rely on the following two lemmas. The first is a collection of simple algorithmic properties of the COM under OS. The second is another weak form of non-bossiness of the COM: Under strong OS, no group of agents can harm any agent by dropping strategies, unless they drop what they are assigned under truth-telling. It should be noted that this property does *not* generally hold under OS, even when the COM is strategy-proof.[27] The proofs of these lemmas are given in Appendix A.

Lemma 4. *Let C_H be an OS profile of choice functions and $f^\star(\cdot) := F^\star(\cdot, C_H)$ denote the cumulative offer mechanism at C_H. For any preference profile $\rhd_D \in \mathscr{P}_D$ and non-null contract $w \in X^G$, the following hold:*

(a) if $\mathsf{d}(w)$ prefers $f^\star(\rhd_D)$ to w (i.e., if $f^\star(\succ_D) \rhd_{\mathsf{d}(w)} w$), then $f^\star\left(\rhd_D^{-w}\right) = f^\star(\rhd_D)$; and

(b) if w is chosen at $\rhd_D$ (i.e., if $w \in f^\star(\rhd_D)$), then $f^\star(\rhd_D) \rhd_{\mathsf{d}(w)} f^\star(\rhd_D^{-w})$.

[26] To see $C'_H(\{z_1, \dots, z_t\}) = C'_H(\{x_1, \dots, x_{t+1}\})$, recall that by definition, C'_H does not choose $y = x_{\tau^*}$ from $\{x_1, \dots, x_{t^*}\}$. By OS, thus, $y = x_{\tau^*} \notin C'_H(\{x_1, \dots, x_{t+1}\})$ for any $t \geq t^* - 1$. Since $\{x_1, \dots, x_{t+1}\} = \{z_1, \dots, z_t\} \cup \{y\}$, then, $C'_H(\{z_1, \dots, z_t\}) = C'_H(x_1, \dots, x_{t+1}\})$ follows from the IRC.

[27] See Example 9 in Appendix I for a counterexample.

Lemma 5. *Let C_H be a strongly OS profile of choice functions and $f^\star(\cdot) = F^\star(\cdot, C_H)$ denote the cumulative offer mechanism at C_H. For any $\rhd_D \in \mathscr{P}_D$ and $Z \subseteq X^G$ such that $Z \cap f^\star(\rhd_D) = \emptyset$, then, $f^\star\left(\rhd_D^{-Z}\right) \trianglerighteq_d f^\star(\rhd_D)$ holds for all $d \in D$.*

7.3.1 Proof of Proposition 3

Given Proposition 2, it suffices to show that there are no $\succ \in \mathscr{P}_D$ and $x \in X^G$ such that both $F^\star(\succ_D, C_H) \succ_{\mathsf{d}(x)} F^\star(\succ_D^{-x}, C_H)$ and $x \notin F^\star(\succ_D, C_H)$. This non-existence is an immediate corollary of Lemma 5 above. ■

7.3.2 Proof of Theorem 3

Suppose that $\mathscr{C}_H$ is rich and strongly OS. By Theorem 2, the COM respects weak improvements if it is strategy-proof. By definitions, it respects strong improvements if it respects weak improvements. To complete the proof, suppose that the COM respects strong improvements; specifically, $F^\star(\succ_D, C_H) \succeq_{\mathsf{d}(x)} F^\star\left(\succ_D, C_H^{-x}\right)$ for all $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$ and $x \in X^G$. Recall that $F^\star\left(\succ_D^{-x}, C_H\right) = F^\star\left(\succ_D, C_H^{-x}\right)$ always holds under the assumption of strong OS (Fact 2). Therefore, the respect for strong improvements entails $F^\star(\succ_D, C_H) \succeq_{\mathsf{d}(x)} F^\star\left(\succ_D^{-x}, C_H\right)$ for all $\succ_D \in \mathscr{P}_D$, $C_H \in \mathscr{C}_H$, and $x \in X^G$. By Proposition 3, this ensures the strategy-proofness of $F^\star$. ■

7.3.3 Proof of Theorem 7

Suppose that $\mathscr{C}_H$ is strongly OS and strongly OSM. Given Theorem 2, it suffices to establish the strategy-proofness of the COM. By Proposition 3, then, we only need to demonstrate that $F^\star(\succ_D^{-x}, C_H) \succ_{\mathsf{d}(x)} F^\star(\succ_D, C_H)$ never holds. Towards a contradiction, suppose otherwise that it holds for some $(x, \succ_D, C_H)$. In what follows, fix such $(x, \succ_D, C_H)$ and let $f^\star(\cdot) := F^\star(\cdot, C_H)$. Assume further that taking C_H as given, $(x, \succ_D)$ is "minimal" in the following sense: For any $x' \in X^G$ and $\rhd_D \in \mathscr{P}_D$,

$$f^\star\left(\rhd_D^{-x'}\right) \succ_{\mathsf{d}(x')} f^\star(\rhd_D) \Longrightarrow \sum_{d \in D} |\mathrm{Ac}(\rhd_d)| \geq \sum_{d \in D} |\mathrm{Ac}(\succ_d)|. \tag{11}$$

This assumption is without loss of generality, because X^G is finite, and hence, so is $\mathscr{P}_D$. Lastly, let $\mathbf{y}$ and $\mathbf{y}^-$ be a complete offer process at $(\succ_D, C_H)$ and at $(\succ_D^{-x}, C_H)$, respectively.

To derive a contradiction, first note that $x \notin f^\star(\succ_D)$ follows from Lemma 4 (b) and the assumption of $f^\star\left(\succ_D^{-x}\right) \succ_{\mathsf{d}(x)} f^\star(\succ_D)$. Then, Lemma 5 entails that

$$f^\star\left(\succ_D^{-x}\right) \succeq_d f^\star\left(\succ_D\right) \text{ for all } d \in D, \tag{12}$$

which further leads to the following observations: On the one hand, for equation (12) to hold true, weakly more agents (in the set sense) should sign a non-null contract at $f^\star\left(\succ_D^{-x}\right)$ than at $f^\star\left(\succ_D\right)$. On the other hand, equation (12) also implies $\mathbb{X}(\mathbf{y}^-) \subseteq \mathbb{X}(\mathbf{y})$ under the OS assumption. By strong OSM, thus, each institution signs a weakly greater number of non-null contracts at $\succ_D$ than at $\succ_D^{-x}$. For these observations to be valid simultaneously, each agent must sign a non-null contract at $f^\star\left(\succ_D\right)$ if and only if so does she at $f^\star\left(\succ_D^{-x}\right)$. In particular, $\mathsf{d}\left(x\right)$ signs two non-null contracts, say z and z^-, at $\succ_D$ and $\succ_D^{-x}$, respectively.[28] However, this contradicts the minimality assumption for the following reason: By assumptions, $z^- \succ_{\mathsf{d}(x)} z \in f^\star(\succ_D)$. Thus, parts (a) and (b) of Lemma 4 imply, respectively, $f^\star\left(\succ_D^{-\{x,z\}}\right) = f^\star\left(\succ_D^{-x}\right)$ and $f^\star\left(\succ_D\right) \succ_{\mathsf{d}(x)} f^\star\left(\succ_D^{-z}\right)$. Combined with the original assumption of $f^\star(\succ_D^{-x}) \succ_{\mathsf{d}(x)} f^\star(\succ_D)$, these together imply $f^\star\left(\succ_D^{-\{x,z\}}\right) \succ_{\mathsf{d}(x)} f^\star\left(\succ_D^{-z}\right)$. This contradicts equation (11) with $(x', \rhd_D) = \left(x, \succ_D^{-z}\right)$, since $\left|\mathrm{Ac}\left(\succ_{\mathsf{d}(x)}^{-z}\right)\right| = \left|\mathrm{Ac}\left(\succ_{\mathsf{d}(x)}\right)\right| - 1$ and $\succ_d = \succ_d^{-z}$ for all $d \neq \mathsf{d}\left(x\right)$. ■

7.4 Proof with US Domains

In this subsection, we prove Theorem 4. Actually, it is immediate once we establish the following lemma, the proof of which is provided in Appendix A. This lemma can be seen as a counterpart of Hirata and Kasuya (2017, Theorem 2) and Hatfield et al. (2021b, Theorem 1b), who establish similar results with strategy-proofness.

Lemma 6. *Let $\mathscr{C}_H$ be a rich and US domain of profiles of choice functions. If a stable mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ respects strong improvements, then it is the cumulative offer mechanism.*

---

[28] Remember that $\mathsf{d}\left(x\right)$ should sign a non-null contract at $\succ_D^{-x}$ by the original assumption that she prefers her assignment at $f^\star\left(\succ_D^{-x}\right)$ to the one at $f^\star\left(\succ_D\right)$.

7.4.1 Proof of Theorem 4

This is an immediate corollary of Theorem 3 and Lemma 6 because US implies strong OS as we demonstrate as Proposition 4 in Appendix D. ■

References

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

Afacan, M. O. (2016): "Characterizations of the Cumulative Offer Process," *Social Choice and Welfare*, 47, 531–542.

——— (2017): "Some Further Properties of the Cumulative Offer Process," *Games and Economic Behavior*, 104, 656–665.

Afacan, M. O. and B. Turhan (2015): "On Relationships between Substitutes Conditions," *Economics Letters*, 126, 10–12.

Alcalde, J. and S. Barberà (1994): "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems," *Economic Theory*, 4, 417–435.

Artemov, G., Y.-K. Che, and Y. He (2021): "Strategic Mistakes: Implications for Market Design Research," *mimeo*.

Avataneo, M. and B. Turhan (2021): "Slot-specific Priorities with Capacity Transfers," *Games and Economic Behavior*, 129, 536–548.

Aygün, O. and I. Bó (2021): "College Admission with Multidimensional Privileges: The Brazilian Affirmative Action Case," *American Economic Journal–Microeconomics*, 13, 1–28.

Aygün, O. and T. Sönmez (2013): "Matching with Contracts: Comments," *American Economic Review*, 103, 2050–2051.

Aygün, O. and B. Turhan (2019): "Dynamic Reserves in Matching Markets," *Journal of Economic Theory*, 188, 105069.

——— (2020): "Designing Direct Matching Mechanisms for India with Comprehensive Affirmative Action," *mimeo*.

Aziz, H. and F. Brandl (2021): "Efficient, Fair, and Incentive-Compatible Healthcare Rationing," *mimeo*.

Balinski, M. and T. Sönmez (1999): "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84, 73–94.

Bando, K. and K. Imamura (2016): "A Necessary and Sufficient Condition for Weak Maskin Monotonicity in an Allocation Problem with Indivisible Goods," *Social Choice and Welfare*, 47, 589–606.

Barberà, S., D. Berga, and B. Moreno (2016): "Group Strategy-Proofness in Private Good Economies," *American Economic Review*, 106, 1073–1099.

Biró, P., A. Hassidim, A. Romm, R. I. Shorrer, and S. Sóvágó (2021): "Need versus Merit: The Large Core of College Admissions Markets," *mimeo*.

Carroll, G. (2012): "When Are Local Incentive Constraints Sufficient?" *Econometrica 80*, 661–686.

Chun, Y. and K. Yun (2020): "Upper-Contour Strategy-Proofness in the Probabilistic Assignment Problem," *Social Choice and Welfare*, 54, 667–687.

Dimakopoulos, P. D. and C.-P. Heller (2019): "Matching with Waiting Times: The German Entry-Level Labour Market for Lawyers," *Games and Economic Behavior*, 115, 289–313.

Dubins, L. E. and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley Algorithm," *American Mathematical Monthly*, 88, 485–491.

Echenique, F. (2012): "Contracts versus Salaries in Matching," *American Economic Review*, 102, 594–601.

Echenique, F. and M. B. Yenmez (2015): "How to Control Controlled School Choice," *American Economic Review*, 105, 2679–2694.

Ehlers, L. (2004): "In Search of Advice for Participants in Matching Markets which Use the Deferred-Acceptance Algorithm," *Games and Economic Behavior*, 48, 249–270.

Ehlers, L., I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim (2014): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," *Journal of Economic Theory*, 153, 648–683.

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

Flanagan, F. X. (2014): "Relaxing the Substitutes Condition in Matching Markets with Contracts," *Economics Letters*, 123, 113–117.

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

Greenberg, K., A. Pathak Parag, and T. Sönmez (2021): "Mechanism Design meets Priority Design: Redesigning the US Army's Branching Process," *NBER Working Paper*, 28911.

Haeringer, G. and F. Klijn (2009): "Constrained School Choice," *Journal of Economic Theory*, 144, 1921–1947.

Hafalir, I. E., F. Kojima, and M. B. Yenmez (2022): "Interdistrict School Choice: A Theory of Student Assignment," *Journal of Economic Theory*, 201, article 105441.

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2013): "Effective Affirmative Action in School Choice," *Theoretical Economics*, 8, 325–363.

Hassidim, A., A. Romm, and R. I. Shorrer (2017): "Redesigning the Israeli Psychology Master's Match," *American Economic Review*, 107, 205–209.

Hatfield, J. W. and F. Kojima (2009): "Group Incentive Compatibility for Matching with Contracts," *Games and Economic Behavior*, 67, 745–749.

——— (2010): "Substitutes and Stability for Matching with Contracts," *Journal of Economic Theory*, 145, 1704–1723.

Hatfield, J. W., F. Kojima, and S. D. Kominers (2021a): "Strategy-Proofness, Investment Efficiency, and Marginal Returns: An Equivalence," *mimeo*.

Hatfield, J. W. and S. D. Kominers (2019): "Hidden Substitutes," *mimeo*.

Hatfield, J. W., S. D. Kominers, and A. Westkamp (2017): "Stable and Strategy-Proof Matching with Flexible Allotments," *American Economic Review Papers and*, Proceedings: 107, 214–219.

——— (2021b): "Stability, Strategy-Proofness, and Cumulative Offer Mechanisms," *Review of Economic Studies*, 88, 1457–1502.

Hatfield, J. W. and P. R. Milgrom (2005): "Matching with Contracts," *American Economic Review*, 95, 913–935.

Hirata, D. and Y. Kasuya (2014): "Cumulative Offer Process is Order-Independent," *Economics Letters*, 124, 37–40.

——— (2017): "On Stable and Strategy-Proof Rules in Matching Markets with Contracts," *Journal of Economic Theory*, 168, 27–43.

Imamura, K. (2020): "Meritocracy versus Diversity," *mimeo*.

Iwase, Y., S. Tsuruta, and A. Yoshimura (2022): "Nash Implementation on the Basis of

General Priorities," *Games and Economic Behavior*, 132, 368–379.

KADAM, S. V. (2017): "Unilateral Substitutability Implies Substitutable Completability in Many-to-One Matching with Contracts," *Games and Economic Behavior*, 102, 56–68.

KASUYA, Y. (2021a): "Group Incentive Compatibility and Welfare for Matching with Contracts," *Economics Letters*, 202, article 109824.

——— (2021b): "Unilateral Substitutability is Necessary for Doctor-Optimal Stability," *Economics Letters*, 207, article 110047.

KELSO, A. S. AND V. P. CRAWFORD (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 50, 1483–1504.

KOJIMA, F. (2012): "School Choice: Impossibilities for Affirmative Action," *Games and Economic Behavior*, 75, 685–693.

KOJIMA, F. AND M. MANEA (2010): "Axioms for Deferred Acceptance," *Econometrica*, 78, 633–653.

KOJIMA, F. AND P. A. PATHAK (2009): "Incentives and Stability in Large Two-Sided Matching Markets," *American Economic Review*, 99, 608–627.

KOMINERS, S. D. AND T. SÖNMEZ (2016): "Matching with Slot-Specific Priorities: Theory," *Theoretical Economics*, 11, 683–710.

KUMANO, T. (2017): "Nash Implementation of Constrained Efficient Stable Matchings under Weak Priorities," *Games and Economic Behavior*, 104, 230–240.

KUMANO, T. AND K. MARUTANI (2021): "Matching with Interdependent Choices," *mimeo*.

MORRILL, T. (2013a): "An Alternative Characterization of the Deferred Acceptance Algorithm," *International Journal of Game Theory*, 42, 19–28.

——— (2013b): "An Alternative Characterization of Top Trading Cycles," *Economic Theory*, 54, 181–197.

PATHAK, P. A. AND T. SÖNMEZ (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98, 1636–1652.

PATHAK, P. A., T. SÖNMEZ, M. U. ÜNVER, AND M. B. YENMEZ (2020): "Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing," *mimeo*.

ROTH, A. E. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7, 617–628.

——— (1984): "Stability and Polarization of Interests in Job Matching," *Econometrica*, 52, 47–58.

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

Roth, A. E. and J. H. Vande Vate (1991): "Incentives in Two-sided Matching with Random Stable Mechanisms," *Economic Theory*, 1, 31–44.

Roy, S. and S. Sadhukhan (2022): "On the Equivalence of Strategy-Proofness and Upper Contour Strategy-Proofness for Randomized Social Choice Functions," *Journal of Mathematical Economics*, 99, article 102593.

Sato, S. (2013a): "Strategy-Proofness and the Reluctance to Make Large Lies: The Case of Weak Orders," *Social Choice and Welfare*, 40, 479–494.

——— (2013b): "A Sufficient Condition for the Equivalence of Strategy-Proofness and Nonmanipulability by Preferences Adjacent to the Sincere One," *Journal of Economic Theory*, 148, 259–278.

Schlegel, J. C. (2015): "Contracts versus Salaries in Matching: A General Result," *Journal of Economic Theory*, 159, 552–573.

——— (2020): "Equivalent Choice Functions and Stable Mechanisms," *Games and Economic Behavior*, 123, 41–53.

Sönmez, T. (2013): "Bidding for Army Career Specialties: Improving the ROTC Branching Mechanism," *Journal of Political Economy*, 121, 186–219.

Sönmez, T. and T. B. Switzer (2013): "Matching with (Branch-of-Choice) Contracts at the United States Military Academy," *Econometrica*, 81, 451–488.

Sönmez, T. and M. B. Yenmez (2022): "Affirmative Action in India via Vertical, Horizontal, and Overlapping Reservations," *Econometrica*, 90, 1143–1176.

Takamiya, K. (2001): "Coalition Strategy-Proofness and Monotonicity in Shapley-Scarf Housing Markets," *Mathematical Social Sciences*, 41, 201–213.

——— (2007): "Domains of Social Choice Functions on Which Coalition Strategy-Proofness and Maskin Monotonicity are Equivalent," *Economics Letters*, 95, 348–354.

Tomoeda, K. (2019): "Efficient Investments in the Implementation Problem," *Journal of Economic Theory*, 182, 247–278.

Westkamp, A. (2013): "An Analysis of the German University Admissions System," *Economic Theory*, 53, 561–589.

Zhang, J. (2016): "On Sufficient Conditions for the Existence of Stable Matchings with

Contracts," *Economics Letters*, 145, 230–234.

A Proofs of the Lemmas

A.1 Proof of Lemma 1

It is immediate to confirm that by definition, X is individually rational at $\left(\succ_D^{-Y}, C_H\right)$ if and only if it is so at $(\succ_D, C_H^{-Y})$. Suppose that at $(\succ_D^{-Y}, C_H)$, an allocation X is individually rational (and hence $X \cap Y = \emptyset$) but blocked by (h, X'). Then by definition, (h, X') should also block X at $(\succ_D, C_H^{-Y})$. Conversely, if $(h,, X')$ blocks an individually rational X at $(\succ_D, C_H^{-Y})$, then it must also block it at $(\succ_D^{-Y}, C_H)$. ∎

A.2 Proof of Lemma 2

Let ϕ, ψ be two H-mechanisms satisfying all the assumptions. Towards a contradiction, suppose that there exists $C_H \in \mathscr{C}_H$ such that $\phi(C_H) \neq \psi(C_H)$. Let $Y \subseteq X^G$ be such that

$$\phi\left(C_H^{-Y}\right) \neq \psi\left(C_H^{-Y}\right), \text{ and}$$
$$\phi\left(C_H^{-Y'}\right) \neq \psi\left(C_H^{-Y'}\right) \implies |Y'| \leq |Y|, \text{ for all } Y' \subseteq X^G.$$

Such Y must exist because X^G is finite, while it may be the empty set. Since $\phi\left(C_H^{-Y}\right)$ and $\psi\left(C_H^{-Y}\right)$ are two distinct stable allocations, some $d \in D$ must sign two distinct non-null contracts (Hirata and Kasuya, 2017, Lemma 2); that is,

$$\varnothing \neq \mathsf{x}\left(d, \phi\left(C_H^{-Y}\right)\right) \neq \mathsf{x}\left(d, \psi\left(C_H^{-Y}\right)\right) \neq \varnothing.$$

Without loss of generality, assume $\phi\left(C_H^{-Y}\right) \succ_d \psi\left(C_H^{-Y}\right)$. Let $z := \mathsf{x}\left(d, \psi\left(C_H^{-Y}\right)\right)$ denote the non-null contract d signs at $\psi\left(C_H^{-Y}\right)$ and $Y' := Y \cup \{z\}$. On the one hand, it follows from ψ's respect for improvements that $\psi\left(C_H^{-Y}\right) \succ_d \psi\left(C_H^{-Y'}\right)$, since C_H^{-Y} is a strong z-improvement over $C_H^{-Y'}$. Notice that the preference must be strict, since d must not sign z at $\psi\left(C_H^{-Y'}\right)$. On the other hand, since $z \notin \phi\left(C_H^{-Y}\right)$ by assumption, the IUC implies $\phi\left(C_H^{-Y'}\right) =_d \phi\left(C_H^{-Y}\right)$. These together imply $\phi\left(C_H^{-Y'}\right) \neq \psi\left(C_H^{-Y'}\right)$, but this contradicts the definition of Y, as $z \notin Y$ and hence $|Y'| = |Y| + 1$ by the definition of Y'.[29] ■

A.3 Proof of Lemma 3

Taking d as arbitrarily fixed, suppose towards a contradiction that $(\succ_D, \rhd_D)$ is a counterexample; i.e., the three conditions on $\succ_D$ and $\rhd_D$ are satisfied while $f^\star(\rhd_D) \neq f^\star(\succ_D)$. Without any loss of generality, suppose further that it is "minimal" in the following sense: For any other counterexample $(\succ'_D, \rhd'_D)$,

$$\min\left\{\sum_{d'\in D}\left|\mathrm{Ac}\left(\succ'_{d'}\right)\right|, \sum_{d'\in D}\left|\mathrm{Ac}\left(\rhd'_{d'}\right)\right|\right\} \geq \min\left\{\sum_{d'\in D}\left|\mathrm{Ac}\left(\succ_{d'}\right)\right|, \sum_{d'\in D}\left|\mathrm{Ac}\left(\rhd_{d'}\right)\right|\right\}. \quad (13)$$

To complete the proof, then, it suffices to construct a non-empty Y such that $(\succ'_D, \rhd'_D) = \left(\succ_D^{-Y}, \rhd_D^{-Y}\right)$ forms a counterexample violating this inequality.

To begin with, note that $f^\star(\rhd_D) =_d f^\star(\succ_D)$ should hold by the assumption of strategy-proofness: If $f^\star(\rhd_D) \succ_d f^\star(\succ_D)$, then d would have an incentive to report $\rhd_d$ when the true preference is $\succ_d$. If $f^\star(\succ_D) \succ_d f^\star(\rhd_D)$, then $f^\star(\succ_D) \rhd_d f^\star(\rhd_D)$ follows from the second assumption for $(\succ_D, \rhd_D)$, i.e., $\{x \in X^G : x \rhd_d f^\star(\succ_D)\} = \{x \in X^G : x \succ_d f^\star(\succ_D)\}$. Thus, d could benefit by reporting $\succ_d$ when the true preference is $\rhd_d$.

Next, we confirm that there should be some $d^* \in D$ who signs distinct non-null contracts at $f^\star(\succ_D)$ and $f^\star(\rhd_D)$; i.e., there should exist $x^*_\succ \in f^\star(\succ_D)$ and $x^*_\rhd \in f^\star(\rhd_D)$ such that $\mathsf{d}\left(x^*_\succ\right) = \mathsf{d}\left(x^*_\rhd\right) = d^*$ and $x^*_\succ \neq x^*_\rhd$. By Lemma 2 of Hirata and Kasuya (2017), such d^* is guaranteed to exist if $f^\star(\succ_D)$ is stable at $(\rhd_D, C_H)$. For some (h, X) to block $f^\star(\succ_D)$ at $(\rhd_D, C_H)$ but not at $(\succ_D, C_H)$, we must have $C_h(X) \rhd_{d'} f^\star(\succ_D) \succ_{d'} C_h(X)$ for some $d' \in D$. However, this is clearly impossible under our assumptions; whether $d' = d$ or not, $\succ_{d'}$ and $\rhd_{d'}$ share the upper contour set of (the contract d' signs at) $f^\star(\succ_D)$. Therefore, $f(\succ_D)$ is stable at $(\rhd_D, C_H)$ and d^* should exist. Note that $d^* \neq d$ and thus $\succ_{d^*} = \rhd_{d^*}$, because d must be indifferent between $f(\rhd_D)$ and $f(\succ_d)$ as seen above.

Now, suppose for a moment that $x^*_\succ \succ_{d^*} x^*_\rhd$ and let

$$Y := \left\{y \in X^G : \mathsf{d}\left(y\right) = d^* \text{ and } x^*_\succ \succ_{d^*} y\right\} \ni x^*_\rhd.$$

[29] Since ψ is stable and thus individually rational, $z \notin Y$ follows from $z \in \psi\left(C_H^{-Y}\right)$.

Note that the contracts in Y are never offered along the COP at $\succ_D$ even though they are acceptable. We thus have $f\left(\succ_D^{-Y}\right) = f(\succ_D)$, which further leads to two observations: First, it is immediate to check that $\left(\succ_D^{-Y}, \rhd_D^{-Y}\right)$ meets the three conditions in the statement of this lemma. Second, it also follows that $f\left(\succ_D^{-Y}\right) \neq f\left(\rhd_D^{-Y}\right)$ for the following reason: On the one hand, $x_\succ^* \in f\left(\succ_D^{-Y}\right)$ because $x_\succ^* \in f(\succ_D)$ by definition and $f(\succ_D) = f\left(\succ_D^{-Y}\right)$ as seen above. On the other hand, strategy-proofness implies $x_\succ^* \notin f\left(\rhd_D^{-Y}\right)$, as otherwise d^* can profitably manipulate by reporting $\rhd_{d^*}^{-Y}$ when the true preference is $\rhd_{d^*}$.[30] That is, $\left(\succ_D^{-Y}, \rhd_D^{-Y}\right)$ constitutes a counterexample to the claim of this lemma. This, however, contradicts the minimality assumption we have imposed on $(\succ_D, \rhd_D)$: Since $d^* \neq d$ as seen above, we have $\succ_{d^*} = \rhd_{d^*}$, and hence, $\left|\text{Ac}\left(\succ_{d^*}^{-Y}\right)\right| = \left|\text{Ac}\left(\rhd_{d^*}^{-Y}\right)\right|$ is strictly smaller than $|\text{Ac}(\succ_{d^*})| = |\text{Ac}(\rhd_{d^*})|$. For any $d' \neq d^*$, $\succ_{d'}^{-Y} = \succ_{d'}$ and $\rhd_{d'}^{-Y} = \rhd_{d'}$. Thus, $(\succ'_D, \rhd'_D) = \left(\succ_D^{-Y}, \rhd_D^{-Y}\right)$ violates inequality (13).

The case of $x_\rhd^* \succ_{d^*} x_\succ^*$ is perfectly symmetric with $Y := \{y \in X^G : x_\rhd^* \succ_{d^*} y\}$, and the proof is complete. ■

A.4 Proof of Lemma 4

First, suppose that $\mathsf{d}(w)$ signs at $f^\star(\rhd_D)$ a non-null contract $z \rhd_{\mathsf{d}(w)} w$, and let $\mathbf{y} = (y_1, \ldots, y_T)$ be a complete offer process at $(\rhd_D, C_H)$. Then, $\mathbb{X}(\mathbf{y})$ cannot contain w for the following reason: For w to be offered along the COP, z must be rejected beforehand. Under the assumption of OS, then, z must be also rejected from $\mathbb{X}(\mathbf{y})$, which contradicts its definition. Given $w \notin \mathbb{X}(\mathbf{y})$, it is immediate to see that $\mathbf{y} = (y_1, \ldots, y_T)$ is also complete at $\left(\rhd_D^{-w}, C_H\right)$ and hence, $f^\star\left(\rhd_D^{-w}\right) = f^\star(\rhd_D)$.

Second, suppose $w \in f^\star(\rhd_D)$, and let $\mathbf{y} = (y_1, \ldots, y_T)$ be a complete offer process at $(\rhd_D, C_H)$. Apparently, there exists some t such that $y_t = w$. By rerunning the COP from step t with $\rhd_D^{-w}$, then, we can obtain an offer process $\mathbf{y}' = \left(y_1, \ldots, y_{t-1}, y'_t, \ldots, y'_{T'}\right)$ that is complete at $\left(\rhd_D^{-w}, C_H\right)$. By definitions, any contract better than w for $\mathsf{d}(w)$, with respect to either $\rhd_{\mathsf{d}(w)}$ or $\rhd_{\mathsf{d}(w)}^{-w}$, must be an element of and be rejected from $\{y_1, \ldots y_{t-1}\}$. Under the assumption of OS, it must be also rejected from $\mathbb{X}(\mathbf{y}')$. Therefore, we obtain $f^\star(\rhd_D) \rhd_{\mathsf{d}(w)} f^\star\left(\rhd_D^{-w}\right)$. ■

[30] Notice that $x_\succ^* \rhd_{d^*} x_\rhd^*$ follows from $x_\succ^* \succ_{d^*} x_\rhd^*$, since $\rhd_{d^*} = \succ_{d^*}$ as we have mentioned above.

A.5 Another Lemma for the Proof of Lemma 5

The next lemma compares two offer processes, **x** and **y**, such that some contract is rejected along the former whereas it is not along the latter. Under the assumption of strong OS, this requires $\mathbb{X}(\mathbf{x}) \nsubseteq \mathbb{X}(\mathbf{y})$, i.e., some contract should be offered along **x** but not along **y**. Moreover, the lemma states that there needs to be a certain kind of preference reversal between the preference profiles underlying **x** and **y**.

Lemma 7. *Suppose that C_H is a strongly OS profile of choice functions. Let* **x** *and* **y** *be a complete offer process at $(\succ_D, C_H)$ and $(\rhd_D, C_H)$, respectively. Suppose further that $\Delta_R := R_H(\mathbb{X}(\mathbf{x})) - R_H(\mathbb{X}(\mathbf{y}))$ is non-empty.*[31] *Then, there exists $x^* \in \mathbb{X}(\mathbf{x}) - \mathbb{X}(\mathbf{y})$ such that either [1] $x^* \notin \mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}\right)$ or [2] $x^* \succ_{\mathsf{d}(x^*)} y^*$ and $y^* \rhd_{\mathsf{d}(x^*)} x^*$, where y^* is the (non-null) contract $\mathsf{d}(x^*)$ signs at $C_H(\mathbb{X}(\mathbf{y}))$.*

Proof. Suppose Δ_R is non-empty. Then, there exists the first step n at which any contract in Δ_R is rejected during the process $\mathbf{x} = (x_1, \ldots, x_n)$; that is, n is such that $R_H(\{x_1, \ldots x_{n-1}\}) \cap \Delta_R$ is empty while $R_H(\{x_1, \ldots, x_n\}) \cap \Delta_R$ is not. The latter implies $R_H(\{x_1, \ldots, x_n\}) \nsubseteq R_H(\mathbb{X}(\mathbf{y}))$ by the definition of Δ_R. This further entails $\{x_1, \ldots, x_n\} \nsubseteq \mathbb{X}(\mathbf{y})$ by the assumption of strong OS. That is, there exists $k \leq n$ such that $x_k \in \{x_1, \ldots, x_n\} - \mathbb{X}(\mathbf{y})$.

Now let $x^* := x_k$ and $d := \mathsf{d}(x^*)$. If $x^* \in \mathrm{Ac}(\rhd_d)$, then d should sign some (non-null) contract y^* at $F^\star(\rhd_D, C_H)$ and $y^* \rhd_d x^*$; otherwise, $x^* \notin \mathbb{X}(\mathbf{y})$ contradicts the assumption that **y** is complete at $(\rhd_D, C_H)$. Furthermore, $x^* \succ_d y^*$ should hold for the following reason: If $y^* \succ_d x^*$, then y^* must be offered and rejected *before* $x^* = x_k$ is offered at step $k < n$ of the process **x**. By the assumption of (strong) OS, it then follows that $y^* \in R_H(\{x_1, \ldots, x_{n-1}\}) \subseteq R_H(\mathbb{X}(\mathbf{x}))$, which further entails $y^* \in R_H(\{x_1, \ldots, x_{n-1}\}) \cap \Delta_R$ since $y^* \notin R_H(\mathbb{X}(\mathbf{y}))$ by its definition. This, however, contradicts the definition of n. As we have shown that $x^* \in \mathrm{Ac}(\rhd_d)$ implies $y^* \rhd_d x^*$ and $x^* \succ_d y^*$, the proof is complete. ■

A.6 Proof of Lemma 5

Let **x** and **y** be a complete offer process at $\left(\rhd_D^{-Z}, C_H\right)$ and $(\rhd_D, C_H)$, respectively. By definition, $f^\star\left(\rhd_D^{-Z}\right) = C_H(\mathbb{X}(\mathbf{x}))$ and $f^\star(\rhd_D) = C_H(\mathbb{X}(\mathbf{y}))$. Towards a contradiction, suppose $Z \cap f^\star(\rhd_D) = \emptyset$ and $f^\star(\rhd_D) \rhd_d f^\star\left(\rhd_D^{-Z}\right)$ for some $d \in D$. Then, d should

[31] Recall that for each $X \subseteq X^G$, $R_H(X)$ is defined to be $X - C_H(X)$.

sign some non-null contract y at $f^\star(\rhd_D)$. Since $y \notin Z$ by definitions, we should have $y \rhd_d^{-Z} f^\star\left(\rhd_D^{-Z}\right)$. Therefore, y should be offered (and rejected) along the process $\mathbf{x}$; i.e., $\Delta_R = R_H(\mathbb{X}(\mathbf{x})) - R_H(\mathbb{X}(\mathbf{y}))$ contains y, and hence, it is non-empty. Substituting $\left(\rhd_D^{-Z}, \rhd_D\right)$ in this proof into $(\succ_D, \rhd_D)$ in Lemma 7 shown in A.5 above, there should exist $x^* \in \mathbb{X}(\mathbf{x})$ such that [1] $x^* \notin \mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}\right)$ or [2] $y^* \rhd_{\mathsf{d}(x^*)} x^*$ but not $y^* \rhd_{\mathsf{d}(x^*)}^{-Z} x^*$. However, neither case is possible: The first case is impossible, because $x^* \in \mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}^{-Z}\right)$ is necessary for it to be offered along $\mathbf{x}$ and $\mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}^{-Z}\right)$ is a subset of $\mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}\right)$ by definition. The second case is impossible, either, since $\rhd_{\mathsf{d}(x^*)}$ and $\rhd_{\mathsf{d}(x^*)}^{-Z}$ fully agree on the rankings among $\mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}^{-Z}\right)$. ■

A.7 Proof of Lemma 6

Suppose that $\mathscr{C}_H$ is US and rich, and let $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ be a stable mechanism that respects strong improvements. Towards a contradiction, assume that there is $(\succ_D, C_H) \in \mathscr{P}_D \times \mathscr{C}_H$ such that $F(\succ_D, C_H) \neq F^\star(\succ_D, C_H)$. Taking $\succ_D$ as fixed, assume further that for any $w \in X^G$,

$$[C_H^{-w} \neq C_H] \Rightarrow \left[F\left(\succ_D, C_H^{-w}\right) = F^\star\left(\succ_D, C_H^{-w}\right)\right]. \tag{14}$$

This assumption is without loss of generality because X^G is finite. By Lemma 2 of Hirata and Kasuya (2017), there must be an agent d who signs two distinct non-null contracts, say x and $x^\star$, at $F(\succ_D, C_H)$ and $F^\star(\succ_D, C_H)$. Note that agent d should strictly prefer $x^\star$ to x because US implies $F^\star(\succ_D, C_H)$ is agent-optimally stable (Hatfield and Kojima, 2010, Theorem 5). This leads to two further observations: First, d also strictly prefers $x^\star$ to $F\left(\succ_D, C_H^{-x}\right)$ because $F(\succ_D, C_H) \succeq_d F\left(\succ_D, C_H^{-x}\right)$ by the assumption of respect for improvements. Second, $F^\star(\succ_D, C_H) = F^\star\left(\succ_D, C_H^{-x}\right)$ because x is not offered along the COP at $(\succ_D, C_H)$. These observations entail $F\left(\succ_D, C_H^{-x}\right) \neq F^\star\left(\succ_D, C_H^{-x}\right)$. It then follows from equation (14) that $C_H^{-x} = C_H$; consequently, $F\left(\succ_D, C_H^{-x}\right) = F\left(\succ_D, C_H\right)$, and hence, $x \in F\left(\succ_D, C_H^{-x}\right)$. However, this contradicts the stability of F. ■

B Omitted Examples

B.1 Stability and Strategy-Proofness in Example 3

In this appendix, we will confirm the claim we have made in Example 3: $F^\star_\pi$ is a stable and strategy-proof mechanism, where π is the precedence order such that $(\pi(d_1), \pi(d_2), \pi(d_3)) = (3, 2, 1)$. Recall that $X^G = \{x_1, x_2, x_3, y_1, y_2\}$, where each x_i and y_i are two contracts between d_i and h, and that C_h and C'_h are induced by

$$\{x_1, x_2, x_3\} \succ_h \{y_1, y_2\} \succ_h \{y_1, x_3\} \succ_h \mathbf{\{y_2, x_3\}} \succ_h \mathbf{\{x_1\}} \succ_h \emptyset, \text{ and}$$
$$\{x_1, x_2, x_3\} \succ'_h \{y_1, y_2\} \succ'_h \{y_1, x_3\} \succ'_h \mathbf{\{x_1\}} \succ'_h \mathbf{\{y_2, x_3\}} \succ'_h \emptyset,$$

where all the subsets of X^G unspecified above are unacceptable. Since all the other cases are straightforward, we only check the stability and strategy-proofness of $F^\star_\pi$ at C_h and C'_h.

Stability: Remember that by definition, an outcome of a COP with a precedence order is stable if its outcome is an allocation. Therefore, we only need to check that the outcomes of $F^\star_\pi$ are always an allocation. When x_3 is acceptable for agent d_3, the outcomes of $F^\star_\pi(\succ_D, C_h)$ and $F^\star_\pi(\succ_D, C'_h)$ are listed in Table 2 below. When x_3 is unacceptable, the outcomes of $F^\star_\pi$ can be written as

$$F^\star_\pi(\succ_D, C_h) = F^\star_\pi(\succ_D, C'_h) = \max_{\succ_{d_1}} \left\{Y \in \left\{\{x_1\}, \{y_1, y_2\}, \emptyset\right\} : Y \text{ is acceptable for } d_2\right\}. \quad (15)$$

In either case, the outcome of $F^\star_\pi$ is an allocation for any $\succ_D$; hence, it is a stable mechanism both at C_h and C'_h.

Strategy-Proofness: To begin with, note that d_3 never has an incentive to misreport, since his unique non-null contract x_3 may be chosen only when $x_3 \succ_{d_3} \varnothing$. When x_3 is unacceptable for d_3, it is easy to see that d_1 and d_2 have no incentive to misreport, given that the values of $F^\star_\pi$ can be rewritten as (15). To check the incentives for d_1 and d_2 when $x_3 \succ_{d_3} \varnothing$, we consider four subcases:

- First, suppose that y_1 is the best contract for d_1. The outcomes in this case with C_h and C'_h are listed in the second and fourth rows of Table 2 (a) and (b), respectively.

	$x_2, y_2, \varnothing$	$y_2, x_2, \varnothing$	$x_2, \varnothing$	$y_2, \varnothing$	$\varnothing$
$x_1, y_1, \varnothing$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{y_1, y_2\}}$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{y_1, y_2\}}$	$\{x_1\}$
$y_1, x_1, \varnothing$	$\{y_1, y_2\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$
$x_1, \varnothing$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{y_2, x_3\}}$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{y_2, x_3\}}$	$\{x_1\}$
$y_1, \varnothing$	$\{y_1, y_2\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$
$\varnothing$	$\{y_2, x_3\}$	$\{y_2, x_3\}$	$\emptyset$	$\{y_2, x_3\}$	$\emptyset$

(a) The Case of C_h

	$x_2, y_2, \varnothing$	$y_2, x_2, \varnothing$	$x_2, \varnothing$	$y_2, \varnothing$	$\varnothing$
$x_1, y_1, \varnothing$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{x_1, x_2, x_3\}}$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{x_1\}}$	$\{x_1\}$
$y_1, x_1, \varnothing$	$\{y_1, y_2\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$
$x_1, \varnothing$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{x_1, x_2, x_3\}}$	$\{x_1, x_2, x_3\}$	$\boldsymbol{\{x_1\}}$	$\{x_1\}$
$y_1, \varnothing$	$\{y_1, y_2\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$	$\{y_1, y_2\}$	$\{y_1, x_3\}$
$\varnothing$	$\{y_2, x_3\}$	$\{y_2, x_3\}$	$\emptyset$	$\{y_2, x_3\}$	$\emptyset$

(b) The Case of C'_h

Table 2: The outcomes of $F^\star_\pi$ in Example 3 when $x_3 \succ_{d_3} \varnothing$. The rows and columns represent the preferences of agent d_1 and d_2, represented as ordered lists. (For instance, "$x_1, y_1, \varnothing$" denotes $\succ_{d_1}$ such that $x_1 \succ_{d_1} y_1 \succ_{d_1} \varnothing$.) The cells with colored, bold fonts are the points where $F^\star_\pi(\cdot, C_h)$ and $F^\star_\pi(\cdot, C'_h)$ disagree.

Notice that either with C_h or C'_h, the outcome is $\{y_1, y_2\}$ if y_2 is acceptable for d_2, and it is $\{y_1, x_3\}$ otherwise. With this observation, it is immediate to see that there is no room to manipulate $F^\star_\pi$ in this case.

- Second, consider the case where x_1 is the best for d_1 and the choice function is C_h. The outcomes in this case are listed in the first and third rows of Table 2 (a). In this case, d_1 fails to obtain x_1 only if y_2 is the best contract for d_2. And if so, d_1 is assigned y_1 if she reports it acceptable and the null contract otherwise. Therefore, d_1 has no incentive to misreport. The incentive compatibility for d_2 is immediate, as she always gets her best contract.
- Next, consider the case where x_1 is the best for d_1 and the choice function is C'_h. The outcomes in this case are listed in the first and third rows of Table 2 (b). Note that d_1 always signs her best contract, x_1, and thus has no incentive to misreport. From d_2's perspective, she obtains x_2 if she reports it acceptable and the null contract otherwise. Thus, d_2 has no incentive to misreport, either.
- Lastly, suppose that no contract is acceptable for d_1. Then d_1 clearly has no incentive to manipulate. Further, d_2 has no incentive to misreport, either, no matter if the choice function is C_h or C'_h. This is because she obtains y_2 if she reports it acceptable and the null contract otherwise; see the fifth row of Table 2 (a)–(b).

Therefore, $F^\star_\pi$ is strategy-proof both at C_h and C'_h.

B.2 OS and Respect for Improvements in Example 4

In this appendix, we will confirm the claim we have made in Example 4: The domain $\mathscr{C}_H$ is OS and the COM respects weak improvements. Remember that $X^G = \{x_i, y_i\}_{i\in\{1,2,3\}}$, where x_i and y_i are two possible contracts between d_i and h. Recall also that $\mathscr{C}_h = \{C_h^{-Y} : Y \subseteq X^G\}$ and that C_h is induced by $\succ_h$ such that

$$\begin{aligned}\{x_1, x_2, x_3\} &\succ_h \{y_1, y_2, y_3\} \\ &\succ_h \{x_1, y_2\} \succ_h \{x_1, x_2\} \succ_h \{x_2, y_3\} \succ_h \{y_1, y_2\} \succ_h \{y_1, x_3\} \\ &\succ_h \text{ [any other doubleton allocations] } \succ_h \text{ [any singletons] } \succ_h \emptyset,\end{aligned}$$

where all tripletons except $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ are unacceptable, and the unspecified rankings among doubletons and among singletons are arbitrary.

$\mathbb{X}\left(\mathbf{w}^3\right)$	$R_h\left(\mathbb{X}\left(\mathbf{w}^3\right)\right)$	w_4	$R_h\left(\mathbb{X}\left(\mathbf{w}^4\right)\right)$	w_5	$R_h\left(\mathbb{X}\left(\mathbf{w}^5\right)\right)$
$\{x_1, x_2, x_3\}$	$\emptyset$				
$\{x_1, x_2, y_3\}$	$\{y_3\}$	x_3	$\{y_3\}$		
$\{x_1, y_2, x_3\}$	$\{x_3\}$	y_3	$\{x_3, y_3\}$		
$\{x_1, y_2, y_3\}$	$\{y_3\}$	x_3	$\{x_3, y_3\}$		
$\{y_1, x_2, x_3\}$	$\{x_2\}$	y_2	$\{x_2, x_3\}$	y_3	$\boldsymbol{\{x_2, x_3\}}$
$\{y_1, x_2, y_3\}$	$\boldsymbol{\{y_1\}}$	x_1	$\{y_1, y_3\}$	x_3	$\{y_1, y_3\}$
$\{y_1, y_2, x_3\}$	$\{x_3\}$	y_3	$\{x_3\}$		
$\{y_1, y_2, y_3\}$	$\emptyset$				

Table 3: Observable offer processes in Example 4.

Observable Substitutability: Since strong disimprovements preserve OS (Fact 4 in Appendix E), we only need to check that C_h satisfies OS. To begin with, let $\mathbf{w}^t = (w_1, \ldots, w_t)$ denote a generic observable process at C_h, and for each $\tau < t$, $\mathbf{w}^\tau = (w_1, \ldots, w_\tau)$ the sub-process of $\mathbf{w}^t$ with length τ. Two observations follow from the fact that C_h accepts any first two offers. First, $t > 3$ is necessary for C_h to violate OS along $\mathbf{w}^t$, i.e., to have $R_h\left(\mathbb{X}(\mathbf{w}^{t-1})\right) \not\subseteq R_h\left(\mathbb{X}(\mathbf{w}^t)\right)$. Second, $\{w_1, w_2, w_3\}$ must contain one contract from each agent; therefore, we have only six possible cases of $\mathbb{X}\left(\mathbf{w}^3\right)$ for an observable offer process, as listed in Table 3. Consider, for instance, the case of $\{w_1, w_2, w_3\} = \{y_1, x_2, y_3\}$. In this case, only y_1 is rejected from $\mathbb{X}(\mathbf{w}^3)$ and hence, $w_4 = x_1$ is necessary for $\mathbf{w}^4 = (w_1, w_2, w_3, w_4)$ to be observable. From $\mathbb{X}(\mathbf{w}^4) = \{y_1, x_1, x_2, y_3\}$, then, C_h chooses $\{x_1, x_2\}$ and rejects $\{y_1, y_3\} \supseteq \{y_1\}$. Thus, the only possibility for w_5 is x_3, and from $\{w_1, \ldots, w_5\} = \{y_1, x_1, x_2, y_3, x_3\}$, C_h chooses $\{x_1, x_2, x_3\}$ while rejecting $\{y_1, y_3\}$ again. Since every agent holds a non-null contract, there is no w_6 such that $\mathbf{w}^6 = (w_1, \ldots, w_6)$ becomes observable. That is, C_h satisfies $R_h(\{w_1, \ldots, w_{t-1}\}) \subseteq R_h(\{w_1, \ldots, w_t\})$ along any $\mathbf{w}^t$ such that $\{w_1, w_2, w_3\} = \{y_1, x_2, y_3\}$. With Table 3, one can check the other cases in a similar way.

Respect for Strong Improvements: Now we show that the COM defined on $\mathscr{P}_D \times \mathscr{C}_H$ respects strong improvements. By Fact 2 in Section 5.2, our task reduces to checking that

$$F^\star(\succ_D, C_h) \succeq_{d_i} F^\star\left(\succ_D^{-w_i}, C_h\right) \tag{16}$$

holds for all $\succ_D \in \mathscr{P}_D$, $d_i \in D$ and $w_i \in X^G$ such that $\mathsf{d}(w_i) = d_i$. This is because, by the definition of $\mathscr{C}_H$ here, any $\widetilde{C}_h \in \mathscr{C}_H$ is equal to C_h^{-Y} for some Y. If respect for strong improvements is violated, thus, $F^\star\left(\rhd_D, C_h^{-Y\cup\{w\}}\right) \rhd_{\mathsf{d}(w)} F^\star\left(\rhd_D, C_h^{-Y}\right)$ for some $\rhd_D$, Y, and w; by Fact 2, this is equivalent to $F^\star\left(\succ_D^{-w}, C_h\right) \rhd_{\mathsf{d}(w)} F^\star(\succ_D, C_h)$ where $\succ_D = \rhd_D^{-Y}$.

To see (16) indeed holds true, arbitrarily fix $\succ_D$, d_i, and w_i such that $\mathsf{d}(w_i) = d_i$. For (16) to fail to hold, d_i must sign a non-null contract, say w_i', at $F^\star\left(\succ_D^{-w_i}, C_h\right)$. Further, $F^\star(\succ_D, C_h) = F^\star\left(\succ_D^{-w_i}, C_h\right)$ should follow if $w_i' \succ_{d_i} w_i$, simply by the definition of the COP.[32] Therefore, (16) fails to hold only if $w_i \succ_{d_i} w_i'$; since d_i has only two non-null contracts in this example, this is equivalent to

$$F^\star(\succ_D, C_h) =_{d_i} \varnothing \text{ and } F^\star\left(\succ_D^{-w_i}, C_h\right) =_{d_i} w_i', \tag{17}$$

where $\{w_i, w_i'\} = \{x_i, y_i\}$ and both of them are acceptable for $\succ_{d_i}$.[33] For $d_i = d_1$ and d_2, (17) never holds for the following reason: Notice that x_1 and y_2 are never rejected along any observable paths of the COP, as one can confirm with Table 3. Thus, $F^\star(\succ_D, C_h) =_{d_i} \varnothing$ and $\mathrm{Ac}(\succ_{d_i}) = \{x_i, y_i\}$ are incompatible with each other.

What remains to consider is the case of $d_i = d_3$ and $\mathrm{Ac}(\succ_{d_3}) = \{x_3, y_3\}$. Note, again with Table 3, that $F(\succ_D, C_h) = \{x_1, y_2\}$ is necessary for both x_3 and y_3 to be rejected along the COP with a preference profile $\succ_D$. Then, it is immediate to see that the outcome of the COP should remain the same even when d_3 stops offering either x_3 or y_3. That is, (17) cannot hold for d_3, and hence, for any d_i. As a consequence, the COM respects improvements in this market.

[32] Under OS, if w_i' is chosen at $F^\star\left(\succ_D^{-w_i}, C_h\right)$, it is never rejected during the COP with $\succ_D^{-w_i}$. Then, d_i has no chance to offer w_i, which is assumed to be less preferred to w_i', even if it is acceptable. Thus, the COP with $\succ_D$ should run exactly the same as with $\succ_D$.

[33] First, $w_i \in \mathrm{Ac}(\succ_{d_i})$ is necessary for the COM outcomes to differ between $\succ_D$ and $\succ_D^{-w_i}$. Second, $w_i' \in \mathrm{Ac}(\succ_{d_i})$ is necessary fo $w_i' \in F^\star\left(\succ_D^{-w_i}, C_h\right)$ since the COM is individually rational.

Respect for Weak Improvements: Lastly, we check that the COM respects not only strong but also weak improvements. To do so, let $X, Y, Z \subseteq X^G$ be such that C_h^{-X} is a weak Z-improvement over C_h^{-Y} and $\mathsf{d}(Z) = \{d\}$ for some $d \in D$. First, then, $X \subseteq Y$ should hold: If there is $x \in X - Y$, we have $C_h^{-X}(\{x\}) = \emptyset$ and $C_h^{-Y}(\{x\}) = \{x\}$. This means C_h^{-X} cannot be a weak Z-improvement over C_h^{-Y}, since $C_h^{-X}(\{x\}) \neq C_h^{-Y}(\{x\})$ and $C_h^{-X}(\{x\}) - C_h^{-Y}(\{x\}) = \emptyset$. Second, $Y \subseteq X \cup Z$ should also hold: For any $y \in Y - X$, we have $C_h^{-X}(\{y\}) = \{y\}$ and $C_h^{-Y}(\{y\}) = \emptyset$. For C_h^{-X} to be a weak Z-improvement over C_h^{-Y}, hence, $y \in Z$ is necessary. Combining the two observations, we obtain $X \subseteq Y \subseteq X \cup Z$. This means C_h^{-X} is a strong W-improvement over C_h^{-Y}, where $W = Y - X$ and $\mathsf{d}(W) = \{d\}$. That is, over this $\mathscr{C}_H$, any weak improvement for agent d is also a strong improvement for d. Since the COM respects strong improvements as we have seen above, it also respects weak improvements.

C Merits of Theorem 7

In this appendix, we detail the potential merits of Theorem 7 that we mentioned in the main text. To begin, our condition is actually sufficient for the COM to be *group strategy-proof* (for short, gSP), which requires that no group of agents can strongly Pareto improve by a joint manipulation. This is because strong OS and strategy-proofness jointly imply gSP as we establish in Appendix G. This is in contrast to the fact that the condition by Hatfield et al. (2021b) is insufficient for gSP (Kasuya, 2021a). Further, as a condition for gSP, ours is strictly weaker than the ones by Hatfield and Kojima (2010) and Hatfield and Kominers (2019). Specifically, as we demonstrate in Appendix D, our strong OS is strictly weaker than their unilateral substitutability and substitutable completability. From a technical point of view, this implies that ours is weak enough *not* to ensure such key structures as the "rural hospital" theorem and the existence of the doctor-optimal stable matching.[34] As a result, not only is our condition the weakest to date for gSP, but it also requires us to depart from the canonical line of proof that exploits those structures.

Next, even as a condition for individual strategy-proofness, ours is the weakest among those which are applicable even when there is only one institution. In such a special case, the condition by Hatfield et al. (2021b) becomes null, because one of their subconditions, NM,

[34]See Example 8 in Appendix I and Kasuya (2021b).

turns into a restatement of strategy-proofness. Our condition would thus be informative at least as a new sufficient condition for their NM. Furthermore, a single-institution market would not be as extreme as it appears: First, technically speaking, having only one institution is without loss of generality, because a multi-institution model (D, H, X^G) can always be rewritten into a single-institution model $\left(D, \left\{\widetilde{h}\right\}, \widetilde{X}^G\right)$, by defining

$$\widetilde{X}^G := \left\{\widetilde{x} = \left(d, \widetilde{h}, (h, \theta)\right) : h \in H \text{ and } (d, h, \theta) \in X^G\right\}. \tag{18}$$

That is, we can identify a contract x between d and h in the original market with a contract $\widetilde{x}$ between d and $\widetilde{h}$ in the single-institution market, by treating h as a part of the contractual term as if h is a branch or subentity of $\widetilde{h}$. Note that $\widetilde{h}$ would have a natural interpretation in many applications, where the matching market is governed by a central authority (e.g., an education authority in a city that governs its school choice system).

Second, the above transformation might be helpful to accommodate *interdependent priority structures*. In the multi-institution case, the IRC implies that each C_h is independent of contracts available to the other institutions; i.e., $C_h(X) = C_h(X')$ for any $X, X' \subseteq X^G$ such that $\{x' \in X : \mathsf{h}(x') = h\} = \{x'' \in X' : \mathsf{h}(x'') = h\}$. While it might appear innocuous, this independence imposes a non-trivial restriction on stability as no justified envy: In the school choice context, for instance, whether a student's claim for a seat at a school h is justified or not can depend only on the assignment of the seats at h, independently of those at other schools. However, the central authority might want to adopt a more flexible criterion of justifiability; e.g., they might judge the student's claim reasonable if she has no other school to attend, but not if she is attending to another decent school. As long as we maintain the IRC, a multi-institution model cannot accommodate such flexibility. In the single-institution counterpart we specified above, contrastingly, $C_{\widetilde{h}}$ is allowed to accommodate interdependency across original h's without violating the IRC, as any contract $\widetilde{x} \in \widetilde{X}^G$ involves $\widetilde{h}$. Thus, the single-institution formulation could be useful to generalize the concept of stability as no justified envy, and our sufficient condition is relevant even in such a direction. While the study of interdependent choice functions is beyond the scope of the present paper, see Kumano and Marutani (2021) for a pioneering work.

Lastly, another potential virtue of our sufficiency result would be in crystallizing the correspondence between each sub-condition and the possible manipulations it excludes. Proposition 3 and Theorem 7 tell us that strong OS is sufficient to preclude profitable adding strategies, while strong OSM is to ensure that dropping strategies also become non-profitable. In the literature, it has been shown that substitutability conditions are sufficient to eliminate certain classes of possible manipulations and that additional conditions are needed for *the rest*.[35] To our knowledge, however, we are the first to formally show exactly what a size-monotonicity condition needs to and actually does rule out. It is our new reduction of strategy-proofness (Proposition 2) that enables this crystallization.

D Relations among Substitutability Conditions

Since Hatfield and Milgrom (2005), a variety of substitutability conditions has been defined and investigated in the literature. In this appendix, we pick three conditions and examine their logical relation to our new concept of strong OS: US, bilateral substitutability, and substitutable completability.[36] In doing so, for simplicity, we consider choice functions rather than profiles of them. The definitions of strong OS and US for an individual choice function are parallel to those for a profile and given in footnotes 19 and 20. The definitions of the other two properties are as follows.

Definition 12 (Hatfield and Kojima, 2010)**.** A choice function C_h for institution h satisfies satisfies *bilateral substitutability* (for short, *BS*) if there are no $x, y \in X^G$ and $Z \subseteq X^G$ such that (i) $x \notin C_H(Z \cup \{x\})$, (ii) $x \in C_H(Z \cup \{x, y\})$, and (iii) $\mathsf{d}(x), \mathsf{d}(y) \notin \mathsf{d}(Z)$. □

Definition 13 (Hatfield and Kominers, 2019)**.** Given a choice function C_h for institution h, a function $C_h^+ : 2^{X^G} \to 2^{X^G}$ is called a *completion* of C_h if it satisfies for all $X \subseteq X^G$, (i) $C_h^+(X) \subseteq X$ and (ii) $C_h^+(X) \in \mathscr{A} \Rightarrow C_h^+(X) = C_h(X)$.[37] A completion C_h^+ of C_h is *substitutable* if $x \notin C_h^+(\{x\} \cup Y)$ and $Y \subseteq Z$ imply $x \notin C_h^+(\{x\} \cup Z)$ for any $x \in X^G$ and $Y, Z \subseteq X^G$. A choice function C_h for institution h is called *substitutably completable* (for short, *SC*) if it has a completion that is substitutable and satisfies the IRC. □

Comparing those three conditions with our strong OS, we can establish the following.

Proposition 4. *Strong OS is strictly weaker than US and SC, and it is logically independent of BS.*

[35] See, e.g., Hatfield and Milgrom (2005, Theorem 10) and Hatfield et al. (2021b, Lemma B.4).

[36] For logical relations among the existing substitutability conditions, see also Afacan and Turhan (2015), Flanagan (2014), Hatfield et al. (2021b), Kadam (2017), and Zhang (2016).

[37] Note that C_h^+ may not be a choice function in our sense, since $C_h^+(X)$ is allowed to be not an allocation.

Proof. To establish the entire claim, it suffices to show that strong OS is strictly weaker than SC for the following reasons: First, it is known that US implies SC (Kadam, 2017; Zhang, 2016); thus, US implies strong OS if SC does. Second, it is also known that both OS and SC are independent of BS (Hatfield and Kominers, 2019; Hatfield et al., 2021b). Therefore, once we confirm that strong OS is weaker than SC (while it is stronger than OS by definition), it should be independent of BS as well.

To establish that strong OS is weaker than SC, suppose that C_h has a completion C_h^+ that is substitutable and satisfies the IRC. Let $\mathbf{x}$ and $\mathbf{y}$ be two observable offer processes for h. By Hatfield and Kominers (2019, Theorem A.2) and Zhang (2016, Lemma 1), then, $C_h(\mathbb{X}(\mathbf{x})) = C_h^+(\mathbb{X}(\mathbf{x}))$ and $C_h(\mathbb{X}(\mathbf{y})) = C_h^+(\mathbb{X}(\mathbf{y}))$. When $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ holds, the substitutability of C_h^+ entails

$$\begin{aligned}\mathbb{X}(\mathbf{x}) - C_h(\mathbb{X}(\mathbf{x})) &= \mathbb{X}(\mathbf{x}) - C_h^+(\mathbb{X}(\mathbf{x})) \\ &\subseteq \mathbb{X}(\mathbf{y}) - C_h^+(\mathbb{X}(\mathbf{y})) = \mathbb{X}(\mathbf{y}) - C_h(\mathbb{X}(\mathbf{y})).\end{aligned}$$

That is, C_h is strongly OS when it is SC.

Next we show by example that SC is not implied by strong OS. Suppose that $D = \{d_x, d_y, d_z\}$, $H = \{h\}$, and $X^G = \{x, \hat{x}, y, z, \hat{z}\}$, where $\mathsf{d}(x) = \mathsf{d}(\hat{x}) = d_x$, $\mathsf{d}(y) = d_y$, and $\mathsf{d}(z) = \mathsf{d}(\hat{z}) = d_z$. Define $\succ_h$ to be the preference for h given by

$$\begin{aligned}&\{\hat{x}, z\} \succ_h \{x, \hat{z}\} \succ_h \{y, \hat{z}\} \succ_h \{\hat{x}, y\} \succ_h \{x, y\} \succ_h \{y, z\} \\ &\quad \succ_h \{\hat{x}, \hat{z}\} \succ_h \{x, z\} \succ_h \{y\} \succ_h \{\hat{z}\} \succ_h \{\hat{x}\} \succ_h \{x\} \succ_h \{z\} \succ_h \emptyset,\end{aligned} \tag{19}$$

and let C_h be the choice function induced by $\succ_h$. Hatfield et al. (2021b, Examples 5–6) show that this C_h is OS but not SC.

To show that this C_h is indeed strongly OS, let $\mathbf{w}^t = (w_1, \ldots, w_t)$ denote a generic observable offer process of length t. Since C_h satisfies OS, which implies the order independence of the COP, we can restrict our attention to offer processes such that $\mathsf{d}(w_1) = d_x$ and $\mathsf{d}(w_2) = d_z$. Note that for any such offer process, $C_h(\{w_1, w_2\}) = \{w_1, w_2\}$. For $\mathbf{w}^3$ to be observable, thus, $w_3 = y$ is necessary. Now suppose for a moment that $\mathbf{w}^3 = (x, z, y)$. Then, C_h chooses $\{x, y\}$ and rejects $\{z\}$ from $\mathbb{X}(\mathbf{w}^3)$. The only possible w_4 that makes $\mathbf{w}^4$ observable is $w_4 = \hat{z}$. From $\mathbb{X}(\mathbf{w}^4)$, C_h rejects $\{z, y\}$ and there is no w_5 with which $\mathbf{w}^5$ is observable. Similarly, we can check all possible paths along which an observable process

$\mathbb{X}(\mathbf{w}^3)$	$R_h\left(\mathbb{X}\left(\mathbf{w}^3\right)\right)$	$\mathbb{X}(\mathbf{w}^4)$	$R_h\left(\mathbb{X}\left(\mathbf{w}^4\right)\right)$
$\{x, y, z\}$	$\{z\}$	$\{x, z, \hat{z}, y\}$	$\{z, y\}$
$\{x, y, \hat{z}\}$	$\{y\}$		
$\{\hat{x}, y, z\}$	$\{y\}$		
$\{\hat{x}, y, \hat{z}\}$	$\{\hat{x}\}$	$\{x, \hat{x}, \hat{z}, y\}$	$\{\hat{x}, y\}$

Table 4: Observable offer processes for C_h induced by (19)

evolves, as summarized in Table 4. With this table, it is immediate to see that C_h is not only OS but also strongly OS. ■

E Properties Preserved by Strong Disimprovements

In this appendix, we show that strong disimprovements preserve various properties on institutions' choice functions. These imply that our richness assumption is compatible with the other assumptions on the domain of choice functions. We begin with the IRC, which we have assumed throughout the paper.

Fact 3. *Let C_h be a choice function for institution h. For any $w \in X^G$, then, C_h^{-w} satisfies the IRC if C_h satisfies it.*

Proof. Suppose that $x \notin C_h^{-w}(X \cup \{x\})$. If $x = w$, then by definitions, $(X \cup \{x\}) - \{w\} = X - \{w\}$ and thus, $C_h^{-w}(X \cup \{x\}) = C_h^{-w}(X)$. If $x \neq w$, then, $(X \cup \{x\}) - \{w\} = (X - \{w\}) \cup \{x\}$. Therefore, we have

$$\begin{aligned} C_h^{-w}(X \cup \{x\}) &\equiv C_h\Big((X \cup \{x\}) - \{w\}\Big) \\ &= C_h\Big((X - \{w\}) \cup \{x\}\Big) = C_h(X - \{w\}) \equiv C_h^{-w}(X), \end{aligned}$$

where the second equality holds by the assumption that C_h satisfies the IRC. ■

Next, we confirm that strong disimprovements preserve (strong) observable substitutability and (strong) observable size-monotonicity. To see the point, suppose that

$\mathbf{x} = (x_1, \ldots, x_n)$ is an offer process for institution h. Given $\mathbf{x}$, let $\mathbf{x}^{-w}$ denote the subsequence of $\mathbf{x}$ that we can obtain by removing w if $\mathbb{X}(\mathbf{x})$ contains it. More formally, $\mathbf{x}^{-w} = (x_1^{-w}, \ldots, x_{\widetilde{n}}^{-w})$ is given as follows:

- if $x_k \neq w$ for any $k \in \{1, \ldots, n\}$, then, $\widetilde{n} = n$ and $x_t^{-w} = x_t$ for all $t \in \{1, \ldots, \widetilde{n}\}$; and
- if $x_k = w$ for some $k \in \{1, \ldots, n\}$, then, $\widetilde{n} = n-1$, $x_t^{-w} = x_t$ for each $t \in \{1, \ldots, k-1\}$, and $x_t^{-w} = x_{t+1}$ for each $t \in \{k, \ldots, \widetilde{n}\}$.

It is then immediate to see that by definitions, $\mathbf{x}^{-w}$ is observable for C_h if $\mathbf{x}$ is observable for C_h^{-w}. With this observation, it is straightforward to establish the following fact:

Fact 4. *Let C_h be a choice function for institution h. For any $w \in X^G$, then, C_h^{-w} satisfies OS, strong OS, OSM, and strong OSM, respectively if C_h satisfies the same condition(s).*

Proof. First, let $\mathbf{x} = (x_1, \ldots, x_n)$ be an offer process for h that is observable at C_h^{-w}. Then, $\mathbf{x}^{-w} = \left(x_1^{-w}, \ldots, x_{\widetilde{n}}^{-w}\right)$ as defined above is observable at C_h. If $x_n = w$, $C_h^{-w}(\{x_1, \ldots, x_n\}) = C_h^{-w}(\{x_1, \ldots, x_{n-1}\})$ holds by definitions, and hence, C_h^{-w} cannot violate OS or OSM at this $\mathbf{x}$. If $x_n \neq w$, then it immediately follows from definitions that $C_h^{-w}(\{x_1, \ldots, x_{n-1}\}) = C_h\left(\left\{x_1^{-w}, \ldots, x_{\widetilde{n}-1}^{-w}\right\}\right)$ and $C_h^{-w}(\{x_1, \ldots, x_n\}) = C_h\left(\left\{x_1^{-w}, \ldots, x_{\widetilde{n}}^{-w}\right\}\right)$. Therefore, if C_h^{-w} violates OS (resp. OSM) at $\mathbf{x}$, then C_h violates OS (resp. OSM) with respect to $\mathbf{x}^{-w}$.

Next, let $\mathbf{x}$ and $\mathbf{y}$ be two observable offer processes at C_h^{-w} such that $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$. Then, $\mathbf{x}^{-w}$ and $\mathbf{y}^{-w}$ are observable at C_h, and $\mathbb{X}(\mathbf{x}^{-w}) \subseteq \mathbb{X}(\mathbf{y}^{-w})$. Since $C_h^{-w}(\mathbb{X}(\mathbf{x})) = C_h(\mathbb{X}(\mathbf{x}^{-w}))$ and $C_h^{-w}(\mathbb{X}(\mathbf{y})) = C_h(\mathbb{X}(\mathbf{y}^{-w}))$ by definitions, if C_h^{-w} violates strong OS (resp. strong OSM) with respect to $\mathbf{x}$ and $\mathbf{y}$, then C_h violates strong OS (resp. strong OSM) with respect to $\mathbf{x}^{-w}$ and $\mathbf{y}^{-w}$. ■

Lastly, strong disimprovements also preserve the substitutability conditions we consider in Appendix D.

Fact 5. *Let C_h be a choice function for institution h. For any $w \in X^G$, then, C_h^{-w} satisfies US, BS, and SC, respectively if C_h satisfies the same condition(s).*

Proof. To check US and BS, suppose that $x \notin C_h^{-w}(Z \cup \{x\})$ and $x \in C_h^{-w}(Z \cup \{x, y\})$. These require $w \notin \{x, y\}$ for the following reasons: If $w = x$, then $x \in C_h^{-w}(Z \cup \{x, y\})$ is impossible by the definition of $C_h^{-w} = C_h^{-x}$. If $w = y$, we must have $(Z \cup \{x\}) - \{w\} = (Z \cup \{x, y\}) - \{w\}$, and hence, $C_h^{-w}(Z \cup \{x\}) = C_h^{-w}(Z \cup \{x, y\})$. Given $w \notin \{x, y\}$, then, it follows from the definition of C_h^{-w} that $x \notin C_h\left(\widetilde{Z} \cup x\right)$ and $x \in C_h\left(\widetilde{Z} \cup \{x, y\}\right)$, where

$\widetilde{Z} := Z - \{w\}$. Moreover, $\left(x, y, \widetilde{Z}\right)$ meets the third condition for US and BS (i.e., condition (iii) of Definitions 9 and 12) whenever (x, y, Z) does, since $\widetilde{Z}$ is a subset of Z. That is, C_h violates US (resp. BS) if C_h^{-w} violates US (resp. BS).

To complete the proof, next suppose that C_h has a completion C_h^+ that is substitutable and meets the IRC. Define $(C_h^+)^{-w} : 2^{X^G} \to 2^{X^G}$ by $(C_h^+)^{-w}(Y) := C_h^+(Y - \{w\})$ for each $Y \subseteq X^G$. Following the same way as of the proofs of Facts 3 and 4, it is then immediate to check $(C_h^+)^{-w}$ is a completion of C_h^{-w} and inherits from C_h^+ both substitutability and the IRC. Hence, C_h^{-w} is SC if C_h is SC. ■

F Strong Respects for Group Improvements

In this appendix, we consider stronger notions of respects for group improvements than the one we consider in Section 6.2. Specifically, we investigate the possibility of a stable mechanism respecting group improvements in the following two senses. Notice that both of them preclude not only strong but also *weak Pareto deterioration* for a group of agents when their priorities improve.

Definition 14. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to *strongly respect strong group improvements* if there are no $\succ_D \in \mathscr{P}_D$, $C_H \in \mathscr{C}_H$, and $Y \subseteq X^G$ such that

- $F\left(\succ_D, C_H^{-Y}\right) \succeq_d F(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$, and
- $F\left(\succ_D, C_H^{-Y}\right) \succ_{d^*} F(\succ_D, C_H)$ for some $d^* \in \mathsf{d}(Y)$. □

Definition 15. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to *strongly respect weak group improvements* if there are no $\succ_D \in \mathscr{P}_D$, $C_H, C'_H \in \mathscr{C}_H$, and $Y \subseteq X^G$ such that

- C_H is a weak Y-improvement over C'_H,
- $F\left(\succ_D, C'_H\right) \succeq_d F(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$, and
- $F\left(\succ_D, C'_H\right) \succ_{d^*} F(\succ_D, C_H)$ for some $d^* \in \mathsf{d}(Y)$. □

Let us also introduce a slightly stronger version of the IUC, which is given as follows. The only difference from the original IUC is that it requires the outcome to be invariant for *all agents* rather than only for the agent whose priority changes.

Definition 16. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to satisfy *the strong irrelevance* *of unchosen contracts* (for short, strong IUC) if

$$[Y \cap F(\succ_D, C_H) = \emptyset] \Longrightarrow F(\succ_D, C_H) = F\left(\succ_D, C_H^{-Y}\right), \tag{20}$$

for all $\succ_D \in \mathscr{P}_D$, $C_H \in \mathscr{C}_H$, and $Y \subseteq X^G$ such that $\mathsf{d}(Y)$ is a singleton and $C_H^{-Y} \in \mathscr{C}_H$. □

In what follows, we demonstrate that strong respect for group improvements (as defined in Definitions 14 and 15 above) is related to the strong version of group strategy-proofness we define next. It should be noted that the literature sometimes refers to this definition as "group strategy-proofness" and to what we call group strategy-proofness (Definition 21 in Appendix G.2 below) as "weak group strategy-proofness."

Definition 17. A mechanism $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is said to be *strongly group strategy-proof* if there are no $\succ_D, \rhd_D \in \mathscr{P}_D$, and $C_H \in \mathscr{C}_H$ such that

- $F(\rhd_D, C_H) \succeq_d F(\succ_D, C_H)$ for all $d \in E := \{d' \in D : \succ_{d'} \neq \rhd_{d'}\}$, and
- $F(\rhd_D, C_H) \succ_{d^*} F(\succ_D, C_H)$ for some $d^\star \in D$. □

On the unstructured domain, we can establish the following equivalence theorem, which is an analogue of Theorem 1. Remember that even in the classic setup, the deferred acceptance mechanism (and hence, any stable mechanism) does *not* satisfy strong group strategy-proofness, except for certain special cases (Ergin, 2002; Roth, 1982). Thus, this theorem could be seen as a negative result that it is almost impossible to design a stable mechanism that strongly respects group improvements.

Theorem 8. *Let $\mathscr{C}_H$ be an arbitrary domain of choice functions and $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ a stable mechanism. Then, F strongly respects strong group improvements and satisfies the strong IUC if it is strongly group strategy-proof. When $\mathscr{C}_H$ is rich, the converse is also true: F strongly respects strong group improvements and satisfies the strong IUC (if and) only if it is strongly group strategy-proof.*

Proof. See Appendix F.1 below. ■

In the case of the COM under OS, we can tighten the above theorem in two ways: First, the two definitions of strong respect for group improvements become equivalent. Specifically, when the COM is strongly group strategy-proof, it strongly respects not only strong but also weak group improvements. Second, OS and strong respect for group improvements jointly imply the strong IUC of the COM. This contrasts with the case of individual improvements; as we have seen in Example 4, OS and respect for (individual) improvements *do not* ensure the IUC of the COM. These two implications of OS lead to the following theorem.

Theorem 9. *Let $\mathscr{C}_H$ be an OS domain of profiles of choice functions. Then, the cumulative offer mechanism $F^\star : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ strongly respects weak group improvements if it is strongly group strategy-proof. When $\mathscr{C}_H$ is rich, the converse is also true: $F^\star$ strongly respects weak group improvements (if and) only if it is strongly group strategy-proof.*

Proof. See Appendix F.2 below. ∎

F.1 Proof of Theorem 8

First, suppose that $F : \mathscr{P}_D \times \mathscr{C}_H \to \mathscr{A}$ is stable and strongly group strategy-proof. To show that it strongly respects strong group improvements, arbitrarily fix $(\succ_D, C_H)$ and Y such that $F\left(\succ_D, C_H^{-Y}\right) \succeq_d F(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$. By Proposition 1, it is equivalent to assume $F\left(\succ_D^{-Y}, C_H\right) \succeq_d F(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$. Since $\succ_{d'}^{-Y} = \succ_{d'}$ for each $d' \notin \mathsf{d}(Y)$, strong group strategy-proofness then requires that $F\left(\succ_D^{-Y}, C_H\right) =_d F(\succ_D, C_H)$ for all $d \in D$. This establishes strong respect for strong group improvements. Next, to establish the strong IUC, arbitrarily fix $(\succ_D, C_H)$, d, and Y such that $Y \cap F(\succ_D, C_H) = \emptyset$, $\mathsf{d}(Y) = \{d\}$, and $C_H^{-Y} \in \mathscr{C}_H$. For d not to have an incentive to misreport, we must have $F(\succ_D, C_H) =_d F\left(\succ_D^{-Y}, C_H\right)$. If there is $d^* \neq d$ who strictly prefers $F\left(\succ_D^{-Y}, C_H\right)$ to $F(\succ_D, C_H)$ with respect to $\succ_{d^*} = \succ_{d^*}^{-Y}$, it contradicts the assumption of strong group strategy-proofness with $\rhd_D = \succ_D^{-Y}$; thus, there should be no such d^*. Since $F(\succ_D, C_H) \succ_{d^*} F\left(\succ_d^{-Y}, C_H\right)$ with $d^* \neq d$ is also impossible by the symmetric argument, $F(\succ_D, C_H) = F\left(\succ_D^{-Y}, C_H\right)$ must hold. Applying Proposition 1, we obtain $F(\succ_D, C_H) = F\left(\succ_D, C_H^{-Y}\right)$ as desired.

Second, suppose that $\mathscr{C}_H$ is rich, F strongly respects strong group improvements, and it satisfies the strong IUC. Suppose that $F(\rhd_D, C_H) \succeq_d F(\succ_D, C_H)$ holds for all $d \in E := \{d' : \succ_{d'} \neq \rhd_{d'}\}$. What we need to establish is $F(\rhd_D, C_H) = F(\succ_D, C_H)$. To do so, let $Y := \{y \in X^G : \mathsf{d}(y) \in E \text{ and } y \notin F(\rhd_D, C_H)\}$. Note that $\rhd_D^{-Y} = \succ_D^{-Y}$ by definition: For $d \in E$, the preference $\rhd_d^{-Y} = \succ_d^{-Y}$ is such that [i] only $\mathsf{x}(d, F(\rhd_D, C_H))$ is acceptable if it is non-null and [ii] no non-null contract is acceptable otherwise. With Proposition 1, we therefore obtain

$$F(\rhd_D, C_H) = F(\rhd_D, C_H^{-Y}) = F\left(\rhd_D^{-Y}, C_H\right) = F\left(\succ_D^{-Y}, C_H\right) = F\left(\succ_D, C_H^{-Y}\right), \tag{21}$$

where the first equality is obtained by repeatedly applying the strong IUC.[38] By the assumption of $F(\rhd_D, C_H) \succeq_d F(\succ_D, C_H)$, it follows that $F\left(\succ_D, C_H^{-Y}\right) \succeq_d F(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$. Then, strong respect for strong group improvements entails $F\left(\succ_D, C_H^{-Y}\right) =_d F(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$, and hence, $Y \cap F(\succ_D, C_H) = \emptyset$. Repeatedly applying the strong IUC, thus, we obtain $F\left(\succ_D, C_H^{-Y}\right) = F(\succ_D, C_H)$. Combined with (21), we can conclude $F(\rhd_D, C_H) = F(\succ_D, C_H)$, and the proof is complete. ■

F.2 Proof of Theorem 9

First, suppose that the COM $F^\star$ is strongly group strategy-proof on $\mathscr{P}_D \times \mathscr{C}_H$. It is indeed without loss of generality to assume $\mathscr{C}_H$ is rich. This is because if the COM is strongly group strategy-proof on $\mathscr{C}_H$, then it is so on $\mathscr{C}_H^*$, where $\mathscr{C}_H^*$ is the minimal rich domain containing $\mathscr{C}_H$.[39] To establish strong respect for weak group improvements, arbitrarily fix $\left(\succ_D, C_H, C'_H\right)$ such that $F^\star\left(\succ_D, C'_H\right) \succeq_d F^\star(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$ and C_H is a weak Y-improvement over C'_H. What we need to show is $F^\star\left(\succ_D, C'_H\right) =_d F^\star(\succ_D, C_H)$ for all $d \in \mathsf{d}(Y)$. Let $Z := X^G - \left(F^\star(\succ_D, C'_H) \cup F^\star(\succ_D, C_H)\right)$. Since $F^\star$ satisfies the strong IUC by Theorem 8, we have

$$F^\star\left(\succ_D, (C'_H)^{-Z}\right) = F^\star(\succ_D, C'_H) \quad \text{and} \quad F^\star\left(\succ_D, C_H^{-Z}\right) = F^\star(\succ_D, C_H).$$

[38] More specifically, the first equality is obtained as follows: Partition Y into $Y_1, \dots, Y_K$ so that $\mathsf{d}(Y_k) = \{d_k\}$ for each $k \in \{1, \dots, K\}$. Then, the strong IUC entails that

$$F(\rhd_D, C_H) = F\left(\rhd_D, C_H^{-Y_1}\right) = F\left(\rhd_D, C_H^{-Y_1 \cup Y_2}\right) = \cdots = F\left(\rhd_D, C_H^{-Y}\right),$$

since for each k, $F(\rhd_D, C_H) = F\left(\rhd_D, C_H^{-Y_1 \cup \cdots \cup Y_k}\right)$ implies $Y_{k+1} \cap F\left(\rhd_D, C_H^{-Y_1 \cup \cdots \cup Y_k}\right) = \emptyset$, and hence, the strong IUC is applicable with Y_{k+1}.

[39] This can be easily confirmed as follows: First, if a mechanism $F = F^\star$ defined on $\mathscr{C}_H$ is strongly group strategy-proof, then we can construct an extension $\widetilde{F}$ defined on $\mathscr{C}_H^*$ maintaining strong group strategy-proofness, as we did in the proof of Proposition 1. Second, the COM is well-defined and should coincide with this extension, even for $C_H \in \mathscr{C}_H^* - \mathscr{C}_H$, because $\mathscr{C}_H^*$ inherits OS from $\mathscr{C}_H$ (Fact 4 in Appendix E).

By Proposition 1, these further entail

$$F^\star\left(\succ_D^{-Z}, C'_H\right) = F^\star\left(\succ_D, C'_H\right) \quad \text{and} \quad F^\star\left(\succ_D^{-Z}, C_H\right) = F^\star(\succ_D, C_H),$$

respectively. Thus, it suffices to establish $F^\star\left(\succ_D^{-Z}, C'_H\right) = F^\star\left(\succ_D^{-Z}, C_H\right)$.

Let $\mathbf{x}' = (x'_1, \ldots, x'_T)$ be a complete offer process at $\left(\succ_D^{-Z}, C'_H\right)$. Recall that by our original assumption, any $d \in \mathsf{d}(Y)$ weakly prefers $F^\star\left(\succ_D^{-Z}, C'_H\right) = F^\star\left(\succ_D, C'_H\right)$ to $F^\star\left(\succ_D^{-Z}, C_H\right) = F^\star(\succ_D, C_H)$. By the definition of Z and the assumption of OS, therefore, $d \in \mathsf{d}(Y)$ offers some x_t if and only if $x_t \in F^\star\left(\succ_D^{-Z}, C'_H\right)$.[40] That is, any contract in Y is never rejected along the process $\mathbf{x}'$. By the definition of weak improvements, it follows that $C_H(\{x_1, \ldots, x_t\}) = C'_H(\{x_1, \ldots, x_t\})$ for any $t \in \{1, \ldots, T\}$. That is, $\mathbf{x}'$ is a complete offer process also at $\left(\succ_D^{-Z}, C_H\right)$, and hence, $F^\star\left(\succ_D^{-Z}, C'_H\right) = F^\star\left(\succ_D^{-Z}, C_H\right)$.

To show the converse, suppose that $\mathscr{C}_H$ is rich. Given Theorem 8, it suffices to show that strong respect for strong group improvements implies the strong IUC. To establish the contraposition, suppose that $F^\star$ violates the strong IUC; i.e., there exists $(\succ_D, C_H)$ and x such that $x \notin F^\star(\succ_D, C_H)$ and $F^\star(\succ_D, C_H) \neq F^\star\left(\succ_D, C_H^{-x}\right)$. In what follows, we confirm that $F^\star$ violates strong respect for strong group improvements.

To begin with, fix a precedence order $\pi(\cdot)$ such that $\pi(\mathsf{d}(x)) = |D|$; i.e., π allows $\mathsf{d}(x)$ to make an offer only when no other agent has an offer to make. Let $\mathbf{x} = (x_1, \ldots, x_T)$ be *the* complete offer process at $(\succ_D, C_H)$ induced by this particular $\pi(\cdot)$. For the assumption of $F^\star(\succ_D, C_H) \neq F^\star\left(\succ_D, C_H^{-x}\right)$ to hold, x must be offered at some step, i.e., $x = x_{t^*}$ for some $t^* \in \{1, \ldots, T\}$. Further, we must have $x = x_{t^*} \in C_H(\{x_1, \ldots, x_{t^*}\})$; otherwise, $\mathbf{x}$ is also complete at $\left(\succ_D, C_H^{-x}\right)$ and hence, $F^\star(\succ_D, C_H) = F^\star\left(\succ_D, C_H^{-x}\right)$ should hold. By the assumption of $x \notin F^\star(\succ_D, C_H)$, however, x must be (firstly) rejected at some later step $t^{**} > t^*$; that is,

$$t^{**} = \min\left\{t > t^* : x \notin C_H(\{x_1, \ldots, x_t\})\right\},$$

is well-defined. Let $Y := \{x_{t^*}, \ldots, x_{t^{**}}\}$ and $Z := X^G - \{x_1, \ldots, x_{t^{**}}\}$. By definition, $\{x_1, \ldots, x_{t^*-1}\} = X^G - (Y \cup Z)$ and $\{x_1, \ldots, x_{t^{**}}\} = X^G - Z$. Therefore, $(x_1, \ldots, x_{t^*-1})$

[40] Namely, $d \in \mathsf{d}(Y)$ cannot offer $w \in F^\star(\succ_D, C_H) - F^\star(\succ_D, C'_H)$: For w to be offered, the contract that d signs at $F^\star(\succ_D, C'_H)$ must be rejected beforehand, but if so, it would not be chosen from $\mathbb{X}(\mathbf{x}')$ by the OS assumption.

and $(x_1, \ldots, x_{t^{**}})$ are complete, respectively, at $\left(\succ_D^{-Y\cup Z}, C_H\right)$ and $\left(\succ_D^{-Z}, C_H\right)$. Then, also by definitions, $\mathsf{d}(x)$ should be assigned the null contract both at $F^\star\left(\succ_D^{-Y\cup Z}, C_H\right) = C_H(\{x_1, \ldots, x_{t^*-1}\})$ and at $F^\star\left(\succ_D^{-Z}, C_H\right) = C_H(\{x_1, \ldots, x_{t^{**}}\})$.[41] By repeatedly applying Fact 2, we can translate this observation into

$$F^\star\left(\succ_D^{-Z}, C_H^{-Y}\right) =_{\mathsf{d}(x)} F^\star\left(\succ_D^{-Z}, C_H\right). \tag{22}$$

Now, recall that by the definition of π, agent $\mathsf{d}(x)$ makes an offer only when all the other agents either hold a contract or have offered all acceptable contracts. For any x_τ with $\tau > t^*$ and $\mathsf{d}(x_\tau) \neq \mathsf{d}(x)$, thus, $\mathsf{d}(x_\tau)$ must hold a better contract at $C_H(\{x_1, \ldots, x_{t^*-1}\})$. Hence,

$$F^\star\left(\succ_D^{-Y\cup Z}, C_H\right) = C_H(\{x_1, \ldots, x_{t^*-1}\}) \succ_d^{-Z} C_H(\{x_1, \ldots, x_{t^{**}}\}) = F^\star\left(\succ_D^{-Z}, C_H\right),$$

holds for any $d \in \mathsf{d}(Y) - \{\mathsf{d}(x)\}$. Since we have $F^\star\left(\succ_D^{-Z}, C_H^{-Y}\right) = F^\star\left(\succ_D^{-Y\cup Z}, C_H\right)$ by repeatedly applying Fact 2, this can be translated into

$$F^\star\left(\succ_D^{-Z}, C_H^{-Y}\right) \succ_d^{-Z} F^\star\left(\succ_D^{-Z}, C_H\right) \text{ for all } d \in \mathsf{d}(Y) - \mathsf{d}(x). \tag{23}$$

Since $\mathsf{d}(Y) - \{\mathsf{d}(x)\} \neq \emptyset$ follows from $t^* < t^{**}$, equations (22)–(23) form a violation of strong respects for strong group improvements, and the proof is complete. ■

G More on Strong Observable Substitutability

In this appendix, we present further implications of strong OS. First, we show that strong OS is necessary and sufficient for the COM to satisfy two monotonicity conditions by Kojima and Manea (2010). Second, we also demonstrate that under strong OS, group strategy-proofness of the COM reduces to individual strategy-proofness.[42]

[41] First, for $\mathsf{d}(x)$ to offer $x = x_{t^*}$ at step t of the process $\mathbf{x}$, she should hold no non-null contract at $C_H(\{x_1, \ldots, x_{t^*-1}\})$. Second, under the assumption of OS, she should hold no contract at $C_H(\{x_1, \ldots, x_{t^{**}}\})$ because x is rejected at step t^{**} by definition.

[42] In the classic setup without contracts and for non-wasteful allocation mechanisms, Bando and Imamura (2016) show a close connection between one of the two monotonicity properties (Definition 18 below) and group strategy-proofness (Definition 21 below). See also Takamiya (2001, 2007) for related results.

G.1 Monotonicity Properties

Kojima and Manea (2010) define the following two properties for matching mechanisms, as an axiom for their characterization of the deferred acceptance mechanism in the classic setup.[43]

Definition 18. A preference $\succ_d \in \mathscr{P}_d$ for agent d is called a *monotonic transformation* of another preference $\rhd_d \in \mathscr{P}_d$ at $x \in X^G \cup \{\varnothing\}$ if

$$\left\{w \in X^G \cup \{\varnothing\} : w \succ_d x\right\} \subseteq \left\{w \in X^G \cup \{\varnothing\} : w \rhd_d x\right\}. \tag{24}$$

A D-mechanism $f : \mathscr{P}_D \to \mathscr{A}$ is said to be *weakly Maskin monotonic* if the following holds for any $\succ_D, \rhd_D \in \mathscr{P}_D$: If $\succ_d$ is a monotonic transformation of $\rhd_d$ at $\mathsf{x}\left(d, f(\rhd_D)\right)$ for each $d \in D$, then $f(\succ_D) \succeq_d f(\rhd_D)$ holds for all $d \in D$. □

Definition 19. A preference $\succ_d \in \mathscr{P}_d$ for agent d is called an *IR monotonic transformation* of another preference $\rhd_d \in \mathscr{P}_d$ at $x \in X^G \cup \{\varnothing\}$ if

$$\{w \in \mathrm{Ac}(\succ_d) : w \succ_d x\} \subseteq \{w \in X^G \cup \{\varnothing\} : w \rhd_d x\}. \tag{25}$$

A D-mechanism $f : \mathscr{P}_D \to \mathscr{A}$ is said to be *IR monotonic* if the following holds for any $\succ_D, \rhd_D \in \mathscr{P}_D$: If $\succ_d$ is an IR monotonic transformation of $\rhd_d$ at $\mathsf{x}\left(d, f(\rhd_D)\right)$ for each $d \in D$, then $f(\succ_D) \succeq_d f(\rhd_D)$ holds for all $d \in D$. □

Note that IR monotonicity is stronger than weak Maskin monotonicity for the following reason: Comparing equations (24) and (25), the left-hand side for a monotonic transformation, $\{w \in X^G \cup \{\varnothing\} : w \succ_d x\}$, is a superset of the one for an IR monotonic transformation, $\{w \in \mathrm{Ac}(\succ_d) : w \succ_d x\}$. Hence, a monotonic transformation of $\rhd_d$ at x is an IR monotonic transformation of $\rhd_d$ at x, but the converse does not necessarily hold true. As a consequence, some mechanisms satisfy weak Maskin monotonicity but not IR monotonicity; a notable example is the top trading cycles mechanism (Kojima and Manea, 2010; Morrill, 2013b).

Actually, when applied to the COM, either of the above two properties turns out to characterize the gap between OS and strong OS: Our Theorem 10 below implies that an OS

[43] See also Afacan (2016) and Morrill (2013a) for related characterizations.

profile C_H of choice functions is strongly OS, if and only if the COM with C_H satisfies weak Maskin monotonicity, if and only if it satisfies IR monotonicity. Note that we cannot simply drop OS in this statement, because without any qualification, the COM is not well-defined as a unique mechanism. The same characterization continues to hold true, however, even if we weaken OS to the following condition, which is also by Hatfield et al. (2021b).

Definition 20. A choice function C_h is said to be *observably substitutable across agents* (for short, *OSaA*) if for any offer process $(x_1, \ldots, x_n)$ for h that is observable at C_h,

$$\left[x \in R_h(\{x_1, \ldots, x_{n-1}\}) - R_h(\{x_1, \ldots, x_n\})\right] \Rightarrow \left[\mathsf{d}(x) \in \mathsf{d}\left(C_h(\{x_1, \ldots, x_{n-1}\})\right)\right]. \quad (26)$$

A profile $C_H = (C_h)_{h \in H}$ of choice functions is said to satisfy OSaA if every C_h satisfies it. Equivalently, C_H is OSaA if and only if

$$\left[x \in R_H(\{x_1, \ldots, x_{n-1}\}) - R_H(\{x_1, \ldots, x_n\})\right] \Rightarrow \left[\mathsf{d}(x) \in \mathsf{d}\left(C_H(\{x_1, \ldots, x_{n-1}\})\right)\right],$$

for any observable process $(x_1, \ldots, x_n)$ at C_H. □

Lemma 8. *Suppose that C_H is a profile of OS choice functions and let $f^\star(\cdot) := F^\star(\cdot, C_H)$ denote the cumulative offer mechanism at C_H. For any preference profile $\rhd_D \in \mathscr{P}_D$ and non-null contract $w \in X^G$, if w is the worst acceptable contract for $\mathsf{d}(w)$ (i.e., if $w' \trianglerighteq_{\mathsf{d}(w)} w$ for all $w' \in \mathrm{Ac}(\rhd_{\mathsf{d}(w)})$), then $f^\star\left(\rhd_D^{-w}\right) \trianglerighteq_d f^\star(\rhd_D)$ for all $d \neq \mathsf{d}(w)$.*

Proof. Suppose $w' \trianglerighteq_{\mathsf{d}(w)} w$ for all $w' \in \mathrm{Ac}(\rhd_{\mathsf{d}(w)})$. If $\mathsf{d}(w)$ signs a non-null contract w' at $f^\star\left(\rhd_D^{-w}\right)$, then, $f^\star\left(\rhd_D^{-w}\right) = f^\star(\rhd_D)$ follows from Lemma 4 (a). To complete the proof, let $\mathbf{y}^- = \left(y_1^-, \ldots, y_T^-\right)$ be a complete offer process at $(\rhd_D^{-w}, C_H)$, suppose that $\mathsf{d}(w)$ signs no non-null contract at $f^\star\left(\rhd_D^{-w}\right) = \cup_h C_h(\mathbb{X}(\mathbf{y}^-))$. Then, we can restart the COM from step $T+1$ by letting $\mathsf{d}(w)$ offer $x_{T+1} = w$, so as to obtain an offer process $\mathbf{y}'' = \left(y_1^-, \ldots, y_T^-, y''_{T+1}, \ldots, y''_{T''}\right)$ that is complete at $(\rhd_D, C_H)$. By the OS assumption, any contract rejected from $\mathbb{X}(\mathbf{y}^-)$ must be also rejected from $\mathbb{X}(\mathbf{y}'')$. For any $d \neq \mathsf{d}(w)$, thus, $f^\star\left(\rhd_D^{-w}\right) \trianglerighteq_d f^\star(\rhd_D)$. ■

Theorem 10. *Let C_H be an OSaA profile of choice functions, so that the cumulative offer mechanism at C_H, denoted by $f^\star(\cdot) = F^\star(\cdot, C_H)$, is well-defined. Then, the following are all equivalent: (1) $f^\star$ is IR monotonic, (2) $f^\star$ is weakly Maskin monotonic, and (3) C_H is strongly OS.*

Proof of (1) ⇒ (2). As we argued earlier, this part is immediate from the definitions of the two monotonicity properties. ■

Proof of (2) ⇒ (3). The proof is in two steps: Assuming OSaA, we first demonstrate that weak Maskin monotonicity implies OS, and then, we extend it to strong OS. To show the first part by contraposition, suppose that C_H is OSaA but not OS; i.e., there exists an observable process $(x_1, \ldots, x_M)$ and $m^* \in \{1, \ldots, M-1\}$ such that $x_{m^*} \notin C_H(\{x_1, \ldots, x_{M-1}\})$ and $x_{m^*} \in C_H(\{x_1, \ldots, x_M\})$. Without loss of generality, assume further that M is the first step at which OS is violated along this process; i.e., $R_H(\{x_1, \ldots, x_{m-1}\}) \subseteq R_H(\{x_1, \ldots, x_m\})$ for any $m < M$. Our goal is to show that $f^\star$ should violate weak Maskin monotonicity.

Note that $d^* := \mathsf{d}(x_{m^*})$ makes at least one offer between step m^* and step M of $(x_1, \ldots, x_M)$: On the one hand, agent d^* must hold some non-null contract at $C_H(\{x_1, \ldots, x_{M-1}\})$ by OSaA. On the other hand, $C_H(\{x_1, \ldots, x_{M-1}\})$ cannot contain any contract that d^* offers before step m^*; this is because (i) for d^* to offer x_{m^*}, any such contract must be once rejected before step m^*, and (ii) we have assumed OS is not violated until step M. Let x_{m° be the first offer d^* makes after step m^*; i.e., $m^\circ \in \{m^*+1, \ldots, M-1\}$, $\mathsf{d}(x_{m^\circ}) = d^*$, and $\mathsf{d}(x_m) \neq d^*$ for all $m \in \{m^*+1, \ldots, m^\circ - 1\}$. Note that this implies d^* holds no non-null contract at $C_H(\{x_1, \ldots, x_{m^\circ-1}\})$.

Now, let $\rhd_D$ be the minimal preference profile such that $\{x_1, \ldots, x_M\}$ is complete at $(\rhd_D, C_H)$. That is, for each d,

- $\mathrm{Ac}(\rhd_d) := \{x_m \in \{x_1, \ldots, x_M\} : \mathsf{d}(x_m) = d\}$, and
- $x_m \rhd_d x_{m'} \Leftrightarrow m < m'$ for any $x_m, x_{m'} \in \mathrm{Ac}(\rhd_d)$.

Also let $\succ_{d^*}$ be the truncation of $\rhd_{d^*}$ at x_{m^*}; i.e.,

- $\mathrm{Ac}(\succ_{d^*}) := \{x_m \in \mathrm{Ac}(\rhd_{d^*}) : x_m \trianglerighteq_{d^*} x_{m^*}\}$, and
- $x_m \succ_{d^*} x_{m'} \Leftrightarrow x_m \rhd_{d^*} x_{m'}$ for any $x_m, x_{m'} \in \mathrm{Ac}(\succ_{d^*})$.

Notice that $\succ_{d^*}$ is a monotonic transformation of $\rhd_{d^*}$ at $x_{m^*} \in f^\star(\rhd_D) = C_H(\{x_1, \ldots, x_M\})$. If d^* is assigned the null contract at $f^\star(\succ_{d^*}, \rhd_{-d^*})$, thus, it is a violation of weak Maskin monotonicity as desired.

To confirm d^* indeed signs no non-null contract at $f^\star(\succ_{d^*}, \rhd_{-d^*})$, note that up to step $m^\circ - 1$, the COP with $(\succ_{d^*}, \rhd_{-d^*})$ runs exactly the same as it does with $\succ_D$. We can thus construct an offer process $(x_1, \ldots, x_{m^\circ-1}, z_{m^\circ}, \ldots, z_T)$ so that it is complete at $(\succ_D, C_H)$. At the end of step $m^\circ - 1$ of this process, (i) d^* holds no non-null contract by the definition of m° as we noted above, and (ii) she has no more contract to offer, since x_{m^*} is the worst acceptable contract for $\vartriangleright_{d^*}$. Recursively applying OSaA, then, C_H should never (re)choose any contract for d^* afterwards; at the end of the process, thus, d^* holds no non-null contract at $C_H(\{x_1, \ldots, x_{m^\circ-1}, z_{m^\circ}, \ldots, z_T\}) = f^\star(\succ_{d^*}, \vartriangleright_{-d^*})$.

What remains to demonstrate is that taking the OS of C_H as given, weak Maskin monotonicity further implies strong OS. To establish the contraposition, suppose C_H is OS but not strongly OS: Suppose $\mathbf{x} = (x_1, \ldots, x_M)$ and $\mathbf{y} = (y_1, \ldots, y_N)$ are two observable processes at C_H such that $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$, and suppose $x^* \in \Delta_R := R_H(\mathbb{X}(\mathbf{x})) - R_H(\mathbb{X}(\mathbf{y}))$ for some $x^* \in \mathbb{X}(\mathbf{x})$. Note that these imply $x^* \in C_H(\mathbb{X}(\mathbf{y}))$. Under the assumption of OS, we can assume without loss of generality that $\mathsf{d}(x^*)$ is assigned the null contract at $C_H(\mathbb{X}(\mathbf{x}))$ for the following reason: Let m^* be the step along $\mathbf{x}$ at which x^* is firstly rejected; i.e., it is the smallest index such that $x^* \in R_H(\{x_1, \ldots, x_{m^*}\})$. Then, OS implies $C_H(\{x_1, \ldots, x_{m^*}\})$ contains no non-null contract for $\mathsf{d}(x^*)$.[44] Even if $\mathsf{d}(x^*)$ signs a non-null contract at $C_H(\mathbb{X}(\mathbf{x}))$, thus, we can redefine $\mathbf{x}$ to be $(x_1, \ldots, x_{m^*})$ maintaining $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ and $x^* \in \Delta_R$. In what follows, we show that $f^\star(\cdot)$ violates weak Maskin monotonicity, assuming $\mathsf{x}(\mathsf{d}(x^*), C_H(\mathbb{X}(\mathbf{x}))) = \varnothing$.

To begin with, let $\vartriangleright_D$ be the minimal preference profile such that $\mathbf{y} = (y_1, \ldots, y_N)$ becomes a complete process at $(\vartriangleright_D, C_H)$. Specifically, define $\vartriangleright_d$ for each $d \in D$ as follows:

- $\mathrm{Ac}(\vartriangleright_d) := \{y \in \mathbb{X}(\mathbf{y}) : \mathsf{d}(y) = d\}$, and
- $y_n \vartriangleright_d y_{n'} \Leftrightarrow n < n'$ for any $y_n, y_{n'} \in \mathrm{Ac}(\vartriangleright_d)$.

Moreover, construct $\succ_d$ from $\vartriangleright_d$ for each $d \in D$ as follows:

- $\mathrm{Ac}(\succ_d) := \left\{y \in \mathbb{X}(\mathbf{y}) : \mathsf{d}(y) = d \text{ and } y \in C_H(\mathbb{X}(\mathbf{x})) \cup C_H(\mathbb{X}(\mathbf{y}))\right\}$, and
- $y \succ_d y' \Leftrightarrow y \vartriangleright_d y'$ for any $y, y' \in \mathrm{Ac}(\succ_d)$.

Notice that $|\mathrm{Ac}(\succ_d)| \leq 2$ for any d by construction. Further, for any $x \in C_H(\mathbb{X}(\mathbf{x}))$, it is the best (or only) acceptable contract for $\succ_{\mathsf{d}(x)}$.[45] This leads to two additional observations: First, x^* is the only acceptable contract for $\succ_{\mathsf{d}(x^*)}$, because (i) $x^* \in C_H(\mathbb{X}(\mathbf{y}))$ by assumptions as we mentioned above, and (ii) we have chosen $\mathbf{x}$ so that $\mathsf{d}(x^*)$ signs no non-null contract at $C_H(\mathbb{X}(\mathbf{x}))$. Second, for each $d \in D$, $\succ_d$ is a monotonic transformation of $\vartriangleright_d$ at $\mathsf{x}(d, f^\star(\vartriangleright_D))$, where $f^\star(\vartriangleright_D) = C_H(\mathbb{X}(\mathbf{y}))$ by the definition of $\vartriangleright_D$.

---

[44] Note that OSaA is insufficient here. This is why we need the first half of this proof.

[45] To see this, suppose $\mathrm{Ac}(\succ_d) = \{x, y\}$, $x \in C_H(\mathbb{X}(\mathbf{x}))$, $y \in C_H(\mathbb{X}(\mathbf{y}))$, and $x \neq y$. Since $\mathbb{X}(\mathbf{x}) \subseteq \mathbb{X}(\mathbf{y})$ by assumption, x should be offered along the process $\mathbf{y}$. Under the assumption of OS, y is never rejected along $\mathbf{y}$, and hence, x should be offered before y. Therefore, $x \vartriangleright_d y$ holds by the construction of $\vartriangleright_d$, which is followed by $x \succ_d y$.

To complete the proof, let $\mathbf{z} = (z_1, \dots, z_T)$ be a complete offer process at $(\succ_D, C_H)$. Since each d has an element of $C_H(\mathbb{X}(\mathbf{x}))$ to offer first or has nothing to offer, it is without loss of generality to assume $\{z_1, \dots, z_t\} = C_H(\mathbb{X}(\mathbf{x}))$ and $z_{t+1} = x^*$, where $t = |C_H(\mathbb{X}(\mathbf{x}))|$.[46] Since $\{z_1, \dots, z_{t+1}\} \subseteq \mathbb{X}(\mathbf{x})$ by definition, then, $C_H(\{z_1, \dots, z_{t+1}\}) = C_H(\mathbb{X}(\mathbf{x}))$ by the IRC, and hence, $x^* \notin C_H(\{z_1, \dots, z_{t+1}\})$. This further leads to $x^* \notin C_H(\mathbb{X}(\mathbf{z})) = f^\star(\succ_D)$, because C_H is assumed to be OS. Since x^* is the only acceptable contract for $\succ_{\mathsf{d}(x^*)}$ as mentioned above, $\mathsf{d}(x^*)$ signs no non-null contract at $f^\star(\succ_D)$. Combined with $x^* \in C_H(\mathbb{X}(\mathbf{y})) = f^\star(\rhd_D)$, we can conclude $f^\star(\rhd_D) \rhd_{\mathsf{d}(x^*)} f^\star(\succ_D)$, despite each $\succ_d$ being a monotonic transformation of $\rhd_d$ at $\mathsf{x}(d, f^\star(\rhd_D))$; i.e., the COM violates weak Maskin monotonicity. ■

Proof of (3) ⇒ (1). The proof is two-fold: We first show that strong OS implies weak Maskin monotonicity and then extends it to IR monotonicity. For the first half, suppose that C_H is strongly OS and arbitrarily fix $\succ_D, \rhd_D \in \mathscr{P}_D$ such that $\succ_d$ is a monotonic transformation of $\rhd_d$ at $\mathsf{x}(d, f^\star(\rhd_D))$ for each $d \in D$. Below we show that $f^\star(\succ_D) \succeq_d f^\star(\rhd_D)$ for all $d \in D$.

To begin with, let us consider a special case where $\mathrm{Ac}(\succ_d) \subseteq \mathrm{Ac}(\rhd_d)$ for all $d \in D$. Towards a contradiction, suppose further that that $f^\star(\rhd_D) \succ_{d^\circ} f^\star(\succ_D)$ for some $d^\circ \in D$.[47] Since $f^\star(\succ_D) \succeq_{d^\circ} \varnothing$ by the individual rationality of the COM, this implies $f^\star(\rhd_D) \succ_{d^\circ} \varnothing$, and hence, d° should sign some non-null contract y° at $f^\star(\rhd_D)$. Then, it should be also offered but rejected along the COP with $\succ_D$; that is, $\Delta_R := R_H(\mathbb{X}(\mathbf{x})) - R_H(\mathbb{X}(\mathbf{y})) \ni y^\circ$ is non-empty, where $\mathbf{x}$ and $\mathbf{y}$ are a complete offer process at $(\succ_D, C_H)$ and $(\rhd_D, C_H)$, respectively. Applying Lemma 7, there should exist $x^* \in \mathbb{X}(\mathbf{x})$ such that (i) $x^* \notin \mathrm{Ac}(\rhd_{\mathsf{d}(x^*)})$ or (ii) $x^* \succ_{\mathsf{d}(x^*)} y^*$ and $y^* \rhd_{\mathsf{d}(x^*)} x^*$, where $y^* = \mathsf{x}(\mathsf{d}(x^*), f^\star(\rhd_D))$. The first case is impossible under the assumption of $\mathrm{Ac}(\succ_{\mathsf{d}(x^*)}) \subseteq \mathrm{Ac}(\rhd_{\mathsf{d}(x^*)})$. The second case is also impossible, by the assumption that $\succ_{\mathsf{d}(x^*)}$ is a monotonic transformation of $\rhd_{\mathsf{d}(x^*)}$ at $\mathsf{x}(\mathsf{d}(x^*), f^\star(\rhd_D)) = y^*$. To avoid a contradiction, therefore, we must have $f^\star(\succ_D) \succeq_d f^\star(\rhd_D)$ for all $d \in D$, as long as $\mathrm{Ac}(\succ_d) \subseteq \mathrm{Ac}(\rhd_d)$ for all $d \in D$.

To complete the first part of the proof, now consider the general case where $\mathrm{Ac}(\succ_d) \subseteq \mathrm{Ac}(\rhd_d)$ may fail to hold. Let $Z := \{z \in X^G : f^\star(\rhd_D) \succ_d z \succ_d \varnothing \text{ for some } d \in D\}$. For each $d \in D$, then, $\succ_d^{-Z}$ remains to be a monotonic transformation of $\rhd_d$ at $\mathsf{x}(d, f(\rhd_D))$,

[46] Recall that the COP is order-independent under the assumption of OS.

[47] As $f^\star(\rhd_D) \succeq_{d^\circ} \varnothing$ holds, $f^\star(\rhd_D) \succ_{d^\circ} f^\star(\succ_D)$ is equivalent to $f^\star(\succ_D) \not\succeq_{d^\circ} f^\star(\rhd_D)$, although unacceptable contracts are incomparable in our definition of preferences.

while we regain $\mathrm{Ac}\left(\succ_d^{-Z}\right) \subseteq \mathrm{Ac}(\rhd_d)$. Therefore, the conclusion of the previous paragraph entails that $f^\star\left(\succ_D^{-Z}\right) \succeq_d f^\star(\rhd_D)$ for all $d \in D$. This further implies that a complete offer process at $\left(\succ_D^{-Z}, C_H\right)$ is also complete at $(\succ_D, C_H)$. Thus, we should have $f^\star(\succ_D) = f^\star\left(\succ_D^{-Z}\right) \succeq_d f^\star(\rhd_D)$ for all $d \in D$, as desired.

Now we proceed to the proof of IR monotonicity. Continue assuming C_H is strongly OS. Let $\succ_D$ and $\rhd_D$ be an arbitrary pair of preference profiles such that each $\succ_d$ is an IR monotonic transformation of $\rhd_d$ at $\mathsf{x}\left(d, f^\star(\rhd_D)\right)$. What we need to show is $f^\star\left(\succ_D\right) \succeq_d f^\star\left(\rhd_D\right)$ for all $d \in D$.

To begin, for each $d \in D$, define an "intermediate" preference $\succ_d^0$ so that (i) $\succ_d^0$ is monotonic transformation of $\rhd_d$ at $\mathsf{x}\left(d, f^\star(\rhd_D)\right)$, and (ii) either $\succ_d = \succ_d^0$ or $\succ_d = \left(\succ_d^0\right)^{-x_d}$, where $x_d = \mathsf{x}\left(d, f^\star(\rhd_D)\right) \neq \varnothing$. More specifically, for each d,

- if $\succ_d$ is a monotonic transformation of $\rhd_d$ at $\mathsf{x}\left(d, f^\star(\rhd_D)\right)$, then $\succ_d^0 := \succ_d$; and
- otherwise, define $\succ_d^0$ by $\mathrm{Ac}(\succ_d^0) := \left\{w \in X^G : w \succeq_d x_d\right\}$, and $w \succ_d^0 w' \Leftrightarrow w \succ_d w'$ for all $w, w' \in \mathrm{Ac}(\succ_d^0)$.

Note that the second case arises only if d signs some non-null contract x_d at $f(\rhd_D)$, since an IR monotonic transformation at $\varnothing$ is always a monotonic transformation at $\varnothing$ by definitions. When $\succ_d^0 \neq \succ_d$, moreover, x_d is the least preferred acceptable contract for $\succ_d^0$, and hence, $\succ_d = \left(\succ_d^0\right)^{-x_d}$.[48] Since $\succ_D^0$ is a monotonic transformation of $\rhd_D$ at $f^\star(\rhd_D)$, it follows from weak Maskin monotonicity we have already established above,

$$f^\star\left(\succ_D^0\right) \succeq_d^0 f^\star(\rhd_D) \text{ for all } d \in D. \tag{27}$$

Now, arbitrarily label the agents as $\{d_1, \ldots, d_T\} = D$, where $T := |D|$, and construct a sequence of preference profile, $\succ_D^1 \ldots, \succ_D^T$, such that for each $i, t \in \{1, \ldots, T\}$, we have $\succ_{d_i}^t := \succ_{d_i}$ if $i \leq t$ and $\succ_{d_i}^t := \succ_{d_i}^0$ otherwise. For each $t \in \{1, \ldots, T\}$, we can then show that

$$f^\star\left(\succ_D^t\right) \succeq_d^t f^\star(\succ_D^{t-1}) \text{ for all } d \in D, \tag{28}$$

as follows. Recall that either $\succ_D^t = \succ_D^{t-1}$ or $\succ_D^t = \left(\succ_D^{t-1}\right)^{-x_{d_t}}$ holds by constructions, where x_{d_t} is the non-null contract d_t signs at $f^\star\left(\rhd_D\right)$. If $\succ_D^t = \succ_D^{t-1}$ and hence $f^\star\left(\succ_D^t\right) =$

[48] Note that if $x_d \in \mathrm{Ac}(\succ_d)$, then the IR monotonic transformation $\succ_d$ should be a monotonic transformation.

$f^\star\left(\succ_D^{t-1}\right)$, then (28) is trivial. Otherwise, for any $d \neq d_t$, $f^\star\left(\succ_D^{t}\right) \succeq_d^t f^\star\left(\succ_D^{t-1}\right)$ follows from Lemma 8, since $\succ_D^t = \left(\succ_D^{t-1}\right)^{-x_{d_t}}$ and x_{d_t} is the least acceptable contract for $\succ_{d_t}^{t-1}$. Moreover, $f^\star\left(\succ_D^{t}\right) \succeq_{d_t}^t f^\star\left(\succ_D^{t-1}\right)$ also holds for the following reasons:

- Agent d_t should sign some non-null contract at $f^\star\left(\succ_D^{t-1}\right)$, since

  $$f^\star\left(\succ_D^{t-1}\right) \succeq_{d_t}^0 f^\star\left(\succ_D^{t-2}\right) \succeq_{d_t}^0 \cdots \succeq_{d_t}^0 f^\star\left(\succ_D^{0}\right) \succeq_{d_t}^0 f^\star(\rhd_D) \ni x_{d_t},$$

  where for each $\tau < t$, the ranking between τ and $\tau - 1$ holds either by $\succ_D^\tau = \succ_D^{\tau-1}$ or by Lemma 8, as we have just argued above.
- If d_t signs x_{d_t} at $f^\star\left(\succ_D^{t-1}\right)$, then she should be assigned the null contract at $f^\star\left(\succ_D^{t}\right)$. This is because $f^\star\left(\succ_D^{t-1}\right) \succ_{d_t}^{t-1} f^\star\left(\succ_D^{t}\right)$ by Lemma 4 (b) with $\succ_D^t = \left(\succ_D^{t-1}\right)^{-x_{d_t}}$, and $x_{d_t} \in f^\star\left(\succ_D^{t-1}\right)$ is the least-preferred acceptable contract for $\succ_{d_t}^{t-1}$. Nevertheless, this implies $f^\star\left(\succ_D^{t}\right) \succ_{d_t}^t f^\star\left(\succ_D^{t-1}\right)$, since x_{d_t} is unacceptable for $\succ_{d_t}^t = \left(\succ_{d_t}^{t-1}\right)^{-x_{d_t}}$,
- If she signs another contract y_{d_t} at $f^\star\left(\succ_D^{t-1}\right)$, then $y_{d_t} \succ_{d_t}^{t-1} x_{d_t}$ because x_{d_t} is the least preferred acceptable. Hence, Lemma 4 (a) implies $f^\star\left(\succ_D^{t}\right) = f^\star\left(\succ_D^{t-1}\right)$.

Therefore, we should have (28) for all $d \in D$ and all $t \in \{1, \ldots, T\}$.

Now we are ready to establish $f^\star(\succ_D) \succ_d f^\star(\rhd_D)$ for all $d \in D$. Combining (27) and (28) across t's, we obtain

$$f^\star(\succ_D) \equiv f^\star\left(\succ_D^{T}\right) \succeq_d^T f^\star\left(\succ_D^{T-1}\right) \succeq_d^{T-1} \cdots \succeq_d^1 f^\star\left(\succ_D^{0}\right) \succeq_d^0 f^\star(\rhd_D).$$

By the definition of $\succ_d^t$'s, this particularly implies that for each $\tau \in \{1, \ldots, T\}$,

$$f^\star(\succ_D) \succeq_{d_\tau} f^\star\left(\succ_D^{\tau-1}\right) \succeq_{d_\tau}^0 f^\star(\rhd_D). \tag{29}$$

For $f^\star(\succ_D) \succeq_{d_\tau} f^\star(\rhd_D)$ fail to hold, thus, $\succ_{d_\tau}$ and $\succ_{d_\tau}^0$ must disagree on the ranking between $f^\star\left(\succ_D^{\tau-1}\right)$ and $f^\star(\rhd_D)$. By definitions, however, $\succeq_{d_\tau} \neq \succeq_{d_\tau}^0$ is possible only when $\succeq_{d_\tau} = \left(\succeq_{d_\tau}^0\right)^{-x_{d_\tau}}$ and $x_{d_\tau} \in f^\star(\rhd_D)$. That is to say, $f^\star\left(\succ_D^{\tau-1}\right) \succeq_{d_\tau}^0 f^\star(\rhd_D)$ and $f^\star\left(\succ_D^{\tau-1}\right) \not\succeq_{d_\tau} f^\star(\rhd_D)$ cannot simultaneously hold. We can thus conclude from (29) that $f^\star(\succ_D) \succeq_{d_\tau} f^\star(\rhd_D)$ for each $d_\tau \in D$, and the proof is complete. ■

G.2 Group Strategy-Proofness

Next, we consider group strategy-proofness, formally defined in our setup as follows:

Definition 21. A D-mechanism $f : \mathscr{P}_D \to \mathscr{A}$ is *group strategy-proof* if there are no $\succ_D, \rhd_D, \in \mathscr{P}_D$ such that $f(\rhd_D) \succ_d f(\succ_D)$ for all $d \in \{d' : \succ_{d'} \neq \rhd_{d'}\}$. □

As we demonstrate below, group strategy-proofness reduces to strategy-proofness for the COM, when the choice functions are strongly OS. Combined with Theorem 7, it follows that the COM is group strategy-proof if the choice functions satisfy both strong OS and strong OSM. This generalizes the results by Hatfield and Kojima (2009, 2010) and Hatfield and Kominers (2019), who establish group strategy-proofness of the COM under stronger substitutability conditions.[49]

Theorem 11. *Let C_H be a strongly OS profile of choice functions. Then, the cumulative offer mechanism at C_H, denoted by $f^\star(\cdot) = F^\star(\cdot, C_H)$, is group strategy-proof if and only if it is strategy-proof.*[50]

Proof. As the "only if" part is immediate by definition, we only establish the "if" part. Suppose towards a contradiction that $f^\star(\cdot)$ is strategy-proof and that there are $\succ^\circ_D, \rhd_D \in \mathscr{P}_D$ and $E \subseteq D$ such that $f^\star(\rhd_D) \succ^\circ_d f^\star\left(\succ^\circ_D\right)$ for all $d \in E$ and $\rhd_{d'} = \succ^\circ_{d'}$ for all $d' \in D - E$. Also assume $\mathrm{Ac}(\succ^\circ_d) \subseteq \mathrm{Ac}(\rhd_d)$ for all $d \in E$. This is without loss of generality for the following reason: Suppose $w \in \mathrm{Ac}(\succ^\circ_d) - \mathrm{Ac}(\rhd_d)$ for some $d \in E$. Let $\rhd'_d$ be the preference obtained by adding w to the "bottom" of the list of acceptable contracts; that is, (i) $z \rhd'_d z' \Leftrightarrow z \rhd_d z'$ for all $z, z' \in \mathrm{Ac}(\rhd_d)$, (ii) $z \rhd'_d w$ for all $z \in \mathrm{Ac}(\rhd_d)$, and (iii) $w \rhd'_d \varnothing$. Recall that d should sign a non-null contract at $f^\star(\rhd_D)$ by the assumption of $f^\star(\rhd_D) \succ^\circ_d f^\star\left(\succ^\circ_D\right)$. During the COP, thus, w is never offered no matter if it is acceptable or not; i.e., $f^\star\left(\rhd'_d, \rhd_{-d}\right) = f^\star(\rhd_D)$. Repeating the same argument, we can construct $\rhd'_d, \rhd''_d, \ldots, \rhd^{(n)}_d$ so that $f^\star(\rhd_D) = f^\star\left(\rhd'_d, \rhd_{-d}\right) = \cdots = f^\star\left(\rhd^{(n)}_d, \rhd_{-d}\right)$ and $\mathrm{Ac}\left(\succ^\circ_d\right) \subseteq \mathrm{Ac}\left(\rhd^{(n)}_D\right)$. By redefining $\rhd_D$ to be $\rhd^{(n)}$, we can always guarantee $\mathrm{Ac}(\succ^\circ_d) \subseteq \mathrm{Ac}(\rhd_d)$ without changing the outcome of the COM.

[49] See also Barberà et al. (2016) for the relation between individual and group strategy-proofness in a general environment beyond matching market.

[50] Since the COM is the unique candidate for a stable and strategy-proof mechanism when C_H is OS (Hatfield et al., 2021b), we can rephrase the conclusion as follows: A stable mechanism $f(\cdot)$ is group strategy-proof if and only if it is strategy-proof.

To begin, construct another preference profile $\succ_D$ from $\succ_D^\circ$ as follows: For each $d \in E$ and for each w such that $\mathsf{d}(w) = d$ and $w \succ_d^\circ f^\star(\rhd_D) \rhd_d w$, lower the "position" of w to anywhere between the (possibly null) contracts that d signs at $f^\star(\rhd_D)$ and at $f^\star(\succ_D^\circ)$. More formally, $\succ_D$ is a preference profile such that

- $\left\{z : z \succeq_d f^\star\left(\succ_D^\circ\right)\right\} = \left\{z : z \succeq_d^\circ f^\star\left(\succ_D^\circ\right)\right\} \subseteq \left(\mathrm{Ac}(\succ_d) \cup \{\varnothing\}\right)$ for all $d \in E$,
- $\{z : z \succ_d f^\star(\rhd_D)\} \subseteq \{z : z \rhd_d f^\star(\rhd_D)\}$ for all $d \in E$, and
- $\succ_{d'} = \succ_{d'}^\circ = \rhd_{d'}$ for all $d' \in D - E$.

By construction, for any $d \in D$, the ranking between $f^\star(\rhd_D)$ and $f^\star(\succ_D)$ remains unchanged either with $\succ_d^\circ$ or with $\succ_d$. Moreover, we also have $f^\star(\succ_D) = f^\star(\succ_D^\circ)$ by repeatedly applying Lemma 3.[51] These observations together imply $f^\star(\rhd_D) \succ_d f^\star(\succ_D)$ for each $d \in E$.

Now we are ready to derive a contradiction. For any $d \in E$, it follows from $f^\star(\rhd_D) \succ_d f^\star(\succ_D)$ that she should sign a non-null contract at $f^\star(\rhd_D)$, and moreover, this contract should be offered and rejected along the COP with $\succ_D$. That is, $R_H(\mathbb{X}(\mathbf{x})) - R_H(\mathbb{X}(\mathbf{y}))$ is non-empty, where $\mathbf{x}$ and $\mathbf{y}$ are the complete offer processes at $(\succ_D, C_H)$ and $(\rhd_D, C_H)$, respectively. Applying Lemma 7, there should exist $x^* \in \mathbb{X}(\mathbf{x})$ such that either [1] $x^* \notin \mathrm{Ac}\left(\rhd_{\mathsf{d}(x^*)}\right)$ or [2] $x^* \succ_{\mathsf{d}(x^*)} y^*$ and $y^* \rhd_{\mathsf{d}(x^*)} x^*$, where y^* is the (non-null) contract $\mathsf{d}(x^*)$ signs at $f^\star(\rhd_D)$. If $\mathsf{d}(x^*) \notin E$, either case clearly contradicts the construction that $\succ_{d'} = \rhd_{d'}$ for all $d' \notin E$. Even if $\mathsf{d}(x^*) \in E$, the first case contradicts the assumption of $\mathrm{Ac}(\succ_d) \subseteq \mathrm{Ac}(\rhd_d)$. The second case also contradicts (the second condition for) the construction of $\succ_D$, which ensures $f^\star(\rhd_D) \rhd_d x^* \Rightarrow f^\star(\rhd_D) \succ_d x^*$ for any $d \in E$. ■

H More on the Definition of an Improvement

In this appendix, we discuss alternative definitions of improvements of priority structures and compare them to our concept of weak improvements. First, we revisit the definition of unambiguous improvements of Kominers and Sönmez (2016). Second, we consider the definition by Afacan (2017).

[51] More precisely, we can establish this equality as follows: Arbitrarily order the members of E as $d_1, \ldots, d_n$, and for each $k \in \{1, \ldots, n\}$, let $\succ_D^k$ to be $\succ_d^k = \succ_d$ for all $d \in \{d_1, \ldots, d_k\}$ and $\succ_{d'}^k = \succ_{d'}^\circ$ for all the others. Then, Lemma 3 implies $f^\star\left(\succ_D^\circ\right) = f^\star\left(\succ_D^1\right) = \cdots = f^\star\left(\succ_D^n\right) \equiv f^\star(\succ_D)$.

H.1 Unambiguous Improvements of Kominers and Sönmez (2016)

Reconsider the case of slot-specific priorities (Example 2 in Section 5), where each choice function C_h is induced by a quota q_h and an ordered list $\mathbf{P} = (P_s)_{s=1,\ldots,q_h}$ of priority orders. Kominers and Sönmez (2016, Section 3.4.2) originally define an unambiguous improvement as follows: A list of slot-specific priorities $\mathbf{P} = (P_s)_{s=1,\ldots,q_h}$ is an unambiguous improvement over $\mathbf{Q} = (Q_s)_{s=1,\ldots,q_h}$ for agent d, if for any $s \in \{1, \ldots, q_h\}$,

- for any $x \in X_d$ and $y \in X_{-d}$, $x \; Q_s \; y \Rightarrow x \; P_s \; y$ and $x \; Q_s \; \varnothing \Rightarrow x \; P_s \; \varnothing$; and
- for any $z, w \in X_{-d}$, we have $z \; Q_s \; w \Leftrightarrow z \; P_s \; w$,

where $X_d := \{x \in X^G : \mathsf{d}(x) = d\}$ and $X_{-d} := X^G - X_d$. Note that this original definition does not require $z \; Q_s \; \varnothing \Leftrightarrow z \; P_s \; \varnothing$. This leads to two consequences, which we demonstrate in the example below: First, an unambiguous improvement for d may not be a weak $Y_{d,h}$-improvement in our sense, where $Y_{d,h} := \{y \in X^G : \mathsf{d}(y) = d \text{ and } \mathsf{h}(y) = h\}$, although the converse remains true. Second, the COM does *not* generally respect unambiguous improvements, even though Theorem 4 of Kominers and Sönmez (2016) claims it does.

Example 5. Let $D := \{d_1, d_2\}$, $H = \{h\}$, and $X^G = \{x_1, x_2\}$, where x_i is a contract between d_i and h. Suppose that h has $q_h = 2$ slots and consider two lists of slot-specific priorities, $\mathbf{P} = (P_1, P_2)$ and $\mathbf{Q} = (Q_1, Q_2)$ defined as follows:

$$\varnothing \; P_1 \; x_1 \; P_1 \; x_2, \qquad x_1 \; P_2 \; x_2 \; P_2 \; \varnothing,$$
$$x_1 \; Q_1 \; \varnothing \; Q_1 \; x_2, \text{ and } x_1 \; Q_2 \; x_2 \; Q_2 \; \varnothing.$$

Notice that $\mathbf{P}$ is an unambiguous improvement over $\mathbf{Q}$ for d_2 according to the original definition by Kominers and Sönmez (2016). However, the choice function $C_{\mathbf{P}}$ induced by $\mathbf{P}$ is *not* a weak x_2-improvement over $C_{\mathbf{Q}}$ induced by $\mathbf{Q}$. Specifically, we have $C_{\mathbf{P}}(\{x_1, x_2\}) = \{x_1\} \neq \{x_1, x_2\} = C_{\mathbf{Q}}(\{x_1, x_2\})$, while $x_2 \notin \emptyset = C_{\mathbf{P}}(\{x_1, x_2\}) - C_{\mathbf{Q}}(\{x_1, x_2\})$.

It is also easy to check the COM does not respect the above unambiguous improvement: Suppose further that $\succ_D = (\succ_{d_1}, \succ_{d_2})$ is such that $x_1 \succ_{d_1} \varnothing$ and $x_2 \succ_{d_2} \varnothing$. Then, the COM outputs $C_{\mathbf{P}}(\{x_1, x_2\}) = \{x_1\}$ at $(\succ_D, C_{\mathbf{P}})$ and $C_{\mathbf{Q}}(\{x_1, x_2\}) = \{x_1, x_2\}$ at $(\succ_D, C_{\mathbf{Q}})$. Apparently, d_2 is clearly worse off at the former than at the latter, even though $\mathbf{P}$ is an unambiguous improvement over $\mathbf{Q}$ for d_2, according to the original definition. □

H.2 Afacan's (2017) Improvements

Afacan (2017) provides the following definition of an improvement and shows that the COM respects this class of improvements under the assumptions of unilateral substitutability and size-monotonicity.

Definition 22. A profile C_H of choice functions is called an *Afacan (2017) improvement* over another profile C'_H for agent d, if for any $h \in H$ and $X \subseteq X^G$, the following hold:

- if there is x such that $\mathsf{d}(x) = d$ and $x \in C'_h(X)$, then there exists y such that $\mathsf{d}(y) = d$ and $y \in C_h(X)$; and
- if $z \notin C_h(X)$ for all z such that $\mathsf{d}(z) = d$, then $C_h(X) = C'_h(X)$. □

An Afacan improvement for agent d would appear similar to our weak Y-improvement with $Y = \{y \in X^G : \mathsf{d}(y) = d\}$. The key (and only) difference is whether or not they allow the choice to vary keeping a same contract for d. That is, an Afacan improvement for $d = \mathsf{d}(y)$ allows $C_h(X) \neq C'_h(X)$ and $y \in C_h(X) \cap C'_h(X)$, whereas our weak improvement does not. While it might appear a minor difference, it actually is not: Other contracts chosen along with y affects what offers will be made along the COP, which in turn, can alter the final outcome for $\mathsf{d}(y)$. As a consequence, Afacan improvements are too broad to be respected by the COM when size-monotonicity is not satisfied. As the following example shows, the COM can fail to respect Afacan improvements even when it is strategy-proof and the choice functions are fully substitutable:

Example 6. Suppose that $D = \{d_1, d_2\}$, $H = \{h\}$, and $X^G = \{x_1, x_2, y_2\}$, where x_1 is a contract between d_1 and h, while x_2 and y_2 represent two distinct ones between d_2 and h. Consider two preferences for h over $\mathscr{A}$, given by

$$\succ_h: \quad \{y_2\} \succ_h \{x_1\} \succ_h \{x_2\} \succ_h \emptyset \succ_h \{x_1, x_2\}, \quad \text{and}$$
$$\succ'_h: \quad \{y_2\} \succ'_h \{x_1, x_2\} \succ'_h \{x_1\} \succ'_h \{x_2\} \succ'_h \emptyset,$$

and let C_h and C'_h be the choice functions induced by $\succ_h$ and $\succ'_h$, respectively. Notice that C_h and C'_h are Afacan improvements over each other, both for d_1. For a mechanism to respect Afacan improvements, the contract assigned to d_1 cannot vary across C_h and C'_h.

In this market, the COM is well-defined, as both C_h and C'_h meet the substitutability condition. Furthermore, the COM is strategy-proof both at C_h and C'_h: Since d_1 has only one non-null contract, she has no room for profitable manipulation. The other agent, d_2, has no incentive to misreport, either, because she can always secure y_2 even if she offers x_2 first.

However, the COM does not respect Afacan improvements. Suppose that $x_1 \succ_{d_1} \varnothing$ and $x_2 \succ_{d_2} y_2 \succ_{d_2} \varnothing$. The outcome of the COM is $\{y_2\}$ at C_h and $\{x_1, x_2\}$ at C'_h. That is, d_1 gets strictly worse off at C_h than at C'_h, while the former is an improvement over the latter for d_1. □

I More Examples

I.1 Non-COM Stable Mechanisms may Respect Improvements

Example 7. Let $D = \{d_1, d_2\}$, $H = \{h\}$, and $X^G = \{x_i, y_i\}_{i \in \{1,2\}}$, where for each $i \in \{1, 2\}$, x_i and y_i are two possible contracts between d_i and h. Let $\succ_h$ be a preference relation over $\mathscr{A}$ such that

$$\{x_1\} \succ_h \{y_1, x_2\} \succ_h \{x_2\} \succ_h \{y_2\} \succ_h \{y_1\} \succ_h \emptyset,$$

where all the subsets of X^G unspecified above are unacceptable. Then, the choice function C_h induced by $\succ_h$ is substitutably completable. To see the point, let $\succ_h^+$ be a preference relation over the subsets of X^G such that

$$\{x_1\} \succ_h^+ \{x_2, y_2\} \succ_h^+ \{y_1, x_2\} \succ_h^+ \{x_2\} \succ_h^+ \{y_2\} \succ_h^+ \{y_1\} \succ_h^+ \emptyset,$$

where all the subsets of X^G unspecified above are unacceptable. Define $C_h^+ : 2^{X^G} \to 2^{X^G}$ so that for each $X \subseteq X^G$, $C_h^+(X)$ is the best subset of X according to $\succ_h^+$. Then, it is easy to check that this C_h^+ is a substitutable completion of C_h. Therefore, C_h is substitutably completable. By Proposition 4 and Fact 4, $\mathscr{C}_H := \{C_h^{-Y} : Y \subseteq X^G\}$ is a strongly OS domain.

In what follows, we show that over this $\mathscr{C}_H$, the COM does not respect strong improvements while another stable mechanism respects weak improvements. To see the first claim, let $\rhd_D$ be the preference profile such that

$$y_1 \rhd_{d_1} x_1 \rhd_{d_1} \varnothing \qquad \text{and} \qquad y_2 \rhd_{d_2} x_2 \rhd_{d_2} \varnothing.$$

Then, it is immediate to check $F^{\star}(\rhd_D, C_h) = \{x_1\}$ and $F^{\star}\left(\rhd_D, C_h^{-y_2}\right) = \{y_1, x_2\}$. Note that agent d_2 signs a non-null contract at $C_h^{-y_2}$ but not at C_h, even though the latter is a strong improvement over the former for her. That is, the COM does not respect strong improvements.

In contrast, the following mechanism F is stable and respects improvements over $\mathscr{P}_D \times \mathscr{C}_H$: For each $(\succ'_D, C'_h) \in \mathscr{P}_D \times \mathscr{C}_H$, define

$$F(\succ'_D, C'_h) := \begin{cases} \{y_1, x_2\} & \text{if } (\succ'_D, C'_h) = (\rhd_D, C_h) \\ F^{\star}(\succ'_D, C'_h) & \text{otherwise,} \end{cases}$$

where $\rhd_D$ is the one defined above. This F is stable because $\{y_1, x_2\}$ is stable at $(\rhd_D, C_h)$. To see it respects improvements, assume for a contradiction that for some $\succ'_D \in \mathscr{P}_D$, (possibly empty) $Y \subseteq X^G$, and $z \in X^G$,

$$F^{\star}\left(\succ'_D, C_h^{-Y\cup\{z\}}\right) \equiv F\left(\succ'_D, C_h^{-Y\cup\{z\}}\right) \succ_{\mathsf{d}(z)} F\left(\succ'_D, C_h^{-Y}\right), \tag{30}$$

where the identity is by the definition of F. To begin, suppose further that $\left(\succ'_D, C_h^{-Y}\right) = (\rhd_D, C_h)$. It is easy to check that $F^{\star}\left(\rhd_D, C_h^{-z}\right)$ is equal to $\{y_2\}$ if $z = x_1$ and to $\{x_1\}$ otherwise. Thus, (30) cannot hold true for any z. Next, consider the case of $\left(\succ'_D, C_h^{-Y}\right) \neq (\rhd_D, C_h)$. Then, by Fact 2 and the definition of F, the assumption of (30) is equivalent to

$$F^{\star}\left(\succ_D, C_h^{-z}\right) \succ_{\mathsf{d}(z)} F^{\star}(\succ_D, C_h), \tag{31}$$

where $\succ_D := \left(\succ'_D\right)^{-Y}$. Note that (31) cannot hold with $\succ_D = \rhd_D$: If it does, $Y = \emptyset$ follows from the definition of $\succ_D$, and hence, $\left(\succ'_D, C_h^{-Y}\right) = (\rhd_D, C_h)$; this would contradict the assumption of $\left(\succ'_D, C_h^{-Y}\right) \neq (\rhd_D, C_h)$. To conclude F respects strong improvements, thus, it suffices to confirm that (31) cannot hold with $\succ_D \neq \rhd_D$, either.

Towards a contradiction, suppose that (31) holds with some $\succ_D \neq \rhd_D$. Then, we must have $\succ_{d_1} = \rhd_{d_1}$ for the following reasons:

- First, suppose $x_1 \notin \mathrm{Ac}(\succ_{d_1})$. Then, (31) cannot hold with $\mathsf{d}(z) = d_2$, because $F^{\star}(\succ_D, C_h)$ should contain the best contract for $\succ_{d_2}$. Moreover, it cannot hold with $\mathsf{d}(z) = d_1$, either: If $z = x_1$, the two sides of (31) must coincide. If $z = y_1$, agent d_1 should sign the null contract at $F^{\star}\left(\succ_D, C_h^{-z}\right)$. Therefore, $x_1 \in \mathrm{Ac}(\succ_{d_1})$ is necessary for (31).

- Second, suppose that x_1 is the best acceptable contract for $\succ_{d_1}$. Then, $F^\star\left(\succ_D, C_h^{-z}\right) = F^\star\left(\succ_D, C_h\right)$ holds unless $z = x_1$. Even if $z = x_1$, (31) cannot hold true because the right-hand side should be $\{x_1\}$ and $\mathsf{d}(z) = d_1$. Therefore, x_1 must be the *second* acceptable contract for $\succ_{d_1}$; that is, $\succ_{d_1} = \rhd_{d_1}$.

Given $\succ_{d_1} = \rhd_{d_1}$, however, (31) cannot hold unless $\succ_{d_2} = \rhd_{d_2}$:

- First, suppose $x_2 \notin \mathrm{Ac}(\succ_{d_2})$. Then, $F^\star\left(\succ_D, C_h\right)$ must contain y_1, which is the best contract for $\succ_{d_1} = \rhd_{d_1}$. Hence, (31) cannot hold with $\mathsf{d}(z) = d_1$. It cannot hold with $\mathsf{d}(z) = d_2$, either, because by the assumption of $y_2 \notin \mathrm{Ac}(\succ_{d_2})$, $F^\star\left(\succ_D, C_h^{-x_2}\right)$ contains no non-null contract fo d_2 and $F^\star\left(\succ_D, C_h^{-y_2}\right)$ is equal to $F^\star\left(\succ_D, C_h\right)$.
- Second, suppose that x_2 is the best acceptable contract for $\succ_{d_2}$. Then, $F^\star\left(\succ_D, C_h\right) = \{y_1, x_2\}$. Since y_1 and x_2 are the best contract for $\succ_{d_1} = \rhd_{d_1}$ and for $\succ_{d_2}$, respectively, equation (31) cannot hold no matter what z is. Therefore, x_2 must be the *second* acceptable contract for $\succ_{d_2}$; that is, $\succ_{d_2} = \rhd_{d_2}$.

In summary, (31) cannot hold for any $\succ_D$, and as a consequence, F respects strong improvements. Since any weak improvements are also strong improvements, it also respects weak improvements. ■

I.2 Strong OS without the AOSM and Rulral Hospital Theorem

Example 8. Let $D = \{d_1, d_2\}$, $H = \{h\}$, and $X^G = \{x_i, y_i\}_{i\in\{1,2\}}$, where for each $i \in \{1, 2\}$, x_i and y_i are two possible contracts between d_i and h. Let $\succ_h$ be a preference relation over $\mathscr{A}$ such that

$$\{x_1, x_2\} \succ_h \{x_1, y_2\} \succ_h \{x_1\} \succ_h \{x_2\} \succ_h \{y_1\} \succ_h \{y_2\} \succ \emptyset,$$

and all the subsets of X^G unspecified above are unacceptable. Let C_h be the choice function induced by $\succ_h$. Since C_h is induced by a (strict) preference relation, it must satisfy the IRC condition. In what follows, we establish that C_h is size-monotonic and strongly OS, and thus, the COM with C_h is strategy-proof by Theorem 7. At the same time, this market fails to maintain some key structures; specifically, we will observe below that the agent-optimal stable matching (for short, AOSM) fails to exist and the "rural hospital" theorem fails to hold in this market.

$\mathbb{X}(\mathbf{w}^2)$	$R_h(\mathbb{X}(\mathbf{w}^2))$	$\mathbb{X}(\mathbf{w}^3)$	$R_h(\mathbb{X}(\mathbf{w}^3))$	$\mathbb{X}(\mathbf{w}^4)$	$R_h(\mathbb{X}(\mathbf{w}^4))$
$\{x_1, x_2\}$	$\emptyset$				
$\{x_1, y_2\}$	$\emptyset$				
$\{y_1, x_2\}$	$\{y_1\}$	$\{x_1, y_1, x_2\}$	$\{y_1\}$		
$\{y_1, y_2\}$	$\{y_2\}$	$\{y_1, x_2, y_2\}$	$\{y_1, y_2\}$	$\{x_1, x_2, y_1, y_2\}$	$\{y_1, y_2\}$

Table 5: Observable offer processes in Example 8.

Size-Monotonicity and Strong OS. First observe that $Z \succ_h Z' \succ_h \emptyset$ implies $|Z| \geq |Z'|$ for any $Z, Z' \subseteq X^G$; therefore, C_h is size-monotonic. To check the strong OS of C_h, let $\mathbf{w}^t = (w_1, \ldots, w_t)$ denote a generic observable offer process at C_h. Table 5 lists all paths along which observable processes can evolve in this market. With this table, it is easy to confirm that C_h is strongly OS.

The AOSM and "rural hospital" theorem. Let $\succ_D$ be such that $y_1 \succ_{d_1} x_1 \succ_{d_1} \varnothing$ and $y_2 \succ_{d_2} \varnothing \succ_{d_2} x_2$. At $(\succ_D, C_h)$, there are two stable allocations: $\{y_1\}$ and $\{x_1, y_2\}$. However, neither is the AOSM, since d_1 prefers $\{y_1\}$ while d_2 prefers $\{x_1, x_2\}$. Further, the "rural hospital" theorem fails to hold at this preference, as d_2 is assigned the null-contract at $\{y_1\}$ but not at $\{x_1, x_2\}$. □

I.3 Lemma 5 may fail w/o strong OS even if the COM is Strategy-Proof

Example 9. Let $D = \{d_1, d_2, d_3\}$, $H = \{h\}$, and $X^G = \{x_1, y_1, x_2, y_2, x_3\}$; x_1 and y_1 (resp. x_2 and y_2) are two different contracts between d_1 and h (resp. d_2 and h), while x_3 is the unique contract between d_3 and h. Let C_h be a choice function that is induced by a preference $\succ_h$ over $\mathscr{A}$ such that

$$\{x_1, x_2, x_3\} \succ_h \{x_1, x_3\} \succ_h \{y_2, x_3\} \succ_h \{y_1, x_2\}$$
$$\succ_h \text{ [all the other doubleton allocations] } \succ_h \text{ [all singletons] } \succ_h \emptyset,$$

$\mathbb{X}(\mathbf{w}^3)$	$R_h\left(\mathbb{X}\left(\mathbf{w}^3\right)\right)$	$\mathbb{X}(\mathbf{w}^4)$	$R_h\left(\mathbb{X}\left(\mathbf{w}^4\right)\right)$	$\mathbb{X}(\mathbf{w}^5)$	$R_h\left(\mathbb{X}\left(\mathbf{w}^5\right)\right)$
$\{x_1, x_2, x_3\}$	$\emptyset$				
$\{x_1, y_2, x_3\}$	$\{y_2\}$	$\{x_1, x_2, y_2, x_3\}$	$\{y_2\}$		
$\{y_1, x_2, x_3\}$	$\{x_3\}$				
$\{y_1, y_2, x_3\}$	$\{y_1\}$	$\{x_1, y_1, y_2, x_3\}$	$\{y_1, y_2\}$	X^G	$\{y_1, y_2\}$

Table 6: Observable offer processes in Example 9.

where all the tripletons but $\{x_1, x_2, x_3\}$ are unacceptable. Since C_h is induced by a (strict) preference, it satisfies the IRC. In what follows, we confirm that C_h meets OS, the COM with C_h is strategy-proof, and yet that the conclusion of Lemma 5 fails to hold in this market.

Observable Substitutability of C_h. As in the previous examples, let $\mathbf{w}^t = (w_1, \ldots, w_t)$ denote a generic offer process. Table 6 lists all possible paths along which observable offer processes evolve in this market. With this table, it is easy to confirm that C_h is OS. Note, however, that it is not strongly OS because $x_3 \in R_h(\{y_1, x_2, x_3\})$ but $x_3 \notin R_h(X^G)$.

Strategy-Proofness of the COM. Referring back to Table 6, it is easy to confirm that x_1 and x_2 are never rejected along any observable offer process; thus, neither d_1 nor d_2 has an incentive to misreport. Moreover, d_3 cannot manipulate the COM, either, because she has only one relevant contract and has a chance to obtain it only if she reports it as acceptable.

Violation of Lemma 5. Let $\succ_D$ be such that

$$y_1 \succ_{d_1} x_1 \succ_{d_1} \varnothing,$$
$$y_2 \succ_{d_2} x_2 \succ_{d_2} \varnothing, \text{ and}$$
$$x_3 \succ_{d_3} \varnothing.$$

A complete offer process at this $\succ_D$ is $(y_1, y_2, x_3, x_1, x_2)$ and the outcome of the COM is $f^\star(\succ_D) = \{x_1, x_2, x_3\}$. At $\succ_D^{-y_2}$, in contrast, a complete offer process is (y_1, x_2, x_3) and the outcome of the COM is $f^{\star}\left(\succ_D^{-y_2}\right) = \{y_1, x_2\}$. Note that d_3 is strictly worse off at $f^{\star}\left(\succ_D^{-y_2}\right)$ than at $f^{\star}(\succ_D)$, even though $y_2 \notin f^{\star}(\succ_D)$. □