UNIVERSITY *of* York

Discussion Papers in Economics

No. 23/05

Stable Matching Mechanisms under Distributional Constraints

Yao Cheng and Zaifu Yang

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Stable Matching Mechanisms under Distributional Constraints[1]

Yao Cheng[2] and Zaifu Yang[3]

Abstract: Motivated by the Chinese doctors, hospitals, and regions matching problem, we study a general matching model with complex distributional constraints. Doctors have preferences over hospitals. Hospitals have preferences over sets of doctors. Every hospital faces its floor and ceiling constraints on the number of doctors, and every region which has several hospitals also faces its floor and ceiling constraints on the number of doctors. We examine how to assign doctors to hospitals and regions in efficient, fair, stable, and strategy-proof ways. We propose two mechanisms for finding such solutions and explore their properties.

Keywords: Matching, distributional constraints, efficiency, fairness, incentive, stability.

JEL Classification: C78, D47.

---

[1]Revised on 26 November 2023. We wish to thank Yan Chen, Karl Claxton, Bingling Gong, Peng Liu, Luigi Siciliani, Jie Zheng, Xiaohan Zhong and other colleagues for their helpful comments on various occasions including 2019 SWUFE Microeconomic Theory Symposium, 2020 Conference on Mechanism and Institution Design (Klagenfurt), 2022 Conference on Mechanism and Institution Design (Singapore), East China Normal University, 2023 Conference on Economic Design (Girona), 2023 Tsinghua Conference on Behavioral, Experimental and Theoretical Economics, and 2023 Beijing International Workshop on Microeconomics: Empirics, Experiments and Theory. The first author is supported by the grant #72203176 of the National Natural Science Foundation of China.

[2]School of Economics and China Center for Behavioral Economics and Finance, Southwestern University of Finance and Economics, Chengdu, 611130, China. Email: chengyao@swufe.edu.cn

[3]Department of Economics and Related Studies, University of York, York YO10 5DD, UK. E-mail: zaifu.yang@york.ac.uk.

1 Introduction

Real-life matching problems always face a variety of distributional constraints. One of the most common and natural constraints is the capacity or ceiling constraint, which places the maximum number of such as positions or people that can be accepted. Essentially, every organization encounters this constraint. The other most common and natural constraint is the floor constraint, which imposes a minimum number on the distribution of resources or rights to ensure certain fairness and maintain sustainability. For instance, a hospital needs to have a certain number of doctors and nurses in order to operate properly.

This paper aims to study a general and practical matching market consisting of a finite number of doctors, hospitals, and regions. Hospitals are located in different districts or regions. Doctors have preferences over hospitals. Hospitals have preferences over sets of doctors. Both hospitals and regions must satisfy certain distributional constraints. Every hospital (region) faces a ceiling constraint and a floor constraint, which give the maximum and minimum number of recruited doctors in the hospital (the region), respectively. The floor constraint on the number of doctors of each hospital is to ensure that the hospital can properly function, and the floor constraint on the number of doctors of each region is to guarantee that the region can provide at least a certain level of medical services for its residents. In China, such floor constraints are very important as the country is huge and varies greatly from one region to another. Without such constraints, less developed regions could have a severe shortage of doctors; see for instance Li and Zhong (2006); Yu and Yang (2008); Zheng (2019); Report (2020); Xia et al. (2021); Report (2021) who pointed out uneven and inefficient distribution of medical resources and called for improvement in China. We believe such problems can also occur in other nations with many regions of unequal development and similar problems arise from civil service sectors and other sectors.[4] We investigate how to assign doctors to hospitals and regions under their distributional con-

[4]Every level of government in every region and its departments need to provide services to its residents and also face their respective floor constraint and ceiling constraint.

straints in terms of (Pareto) efficiency, fairness, stability, and incentive-compatibility.

To deal with this market with complex distributional constraints, we propose two new solutions that exhibit several desirable and important properties. The first solution satisfies efficiency, fairness, and strategy-proofness. The second solution satisfies group stability and fairness within the same quota type.

Stability is the most fundamental solution concept in two-sided matching markets; see Gale and Shapley (1962) and Roth and Sotomayor (1990). It is well-known that markets without distributional constraints possess at least one stable matching. For such markets, stability is defined via blocking pair, which can make both members of the pair better off and at least one member strictly better off via themselves. However, when there are distributional constraints, a feasible matching has to satisfy all the constraints. Not every pair can form a blocking pair against a feasible matching, because matching the blocking pair together may violate some of the constraints. It seems natural to require that a pair can block a feasible matching if matching the blocking pair together does not violate any constraint. Unfortunately, stable matchings via such blocking pairs are too strong to exist.

To break the impasse, we introduce a novel concept of blocking pair called *admissible blocking pair.* A doctor-hospital pair (d, h) forms an admissible blocking pair against a feasible matching μ if either matching the pair together results in a feasible matching or the hospital's hiring the doctor with the intention of replacing another doctor $d' \in \mu(h)$ later and keeping all its hired doctors $\mu(h)$ satisfies all the constraints except the hospital's capacity constraint. This temporary violation of the hospital's capacity constraint was inspired by a common practice that many organizations do give their dismissed employees a grace period of time to leave. Clearly, after the dismissed doctor d' leaves the hospital, this will result in a feasible matching. A feasible matching is (pairwise) stable if there does not exist any admissible blocking pair. We can analogously define admissible blocking coalitions and group stability. Group stability implies efficiency and pairwise stability.

Fairness is another fundamental concept concerning equitable distribution of resources and means that agents at their assignments do not envy each other.[5] We find, however, that fairness is not compatible with pairwise stability. To maintain and be compatible with stability, we relax fairness to fairness within the same quota type. In our model, doctors who can be assigned to a hospital are constrained by the quotas that the hospital and its region face. Fairness within the same quota type means that doctors who share the same quota type do not envy each other and their associated hospitals do not envy each other. Nonwastefulness concerns a different kind of efficiency.[6] In the current context, we say that a feasible matching is nonwasteful if there do not exist any doctor and any hospital who are unmatched such that the doctor prefers the hospital to her assignment and the hospital has unfilled positions and wants to hire the doctor and doing so does not violate any constraint. We show that in our model, pairwise stability implies nonwastefulness, but (Pareto) efficiency and nonwastefulness do not imply each other.

To achieve the two proposed solutions, we develop two new mechanisms. The first one is called the deferred acceptance algorithm with distributional constraints (the DA-D algorithm). This mechanism can be seen as a significant generalization of the celebrated deferred acceptance algorithm of Gale and Shapley (1962). It is shown that the outcome generated by the DA-D algorithm is fair and efficient, and the algorithm is *strategy-proof* for all doctors. Strategy-proofness is a fundamental and appealing property for a mechanism to be successful and guarantees that it is in the best interest of all doctors to reveal their preferences truthfully. This property has been widely discussed in the literature.[7] In the DA-D algorithm, because of regional floor and ceiling constraints, we have to use a picking order rule to adjust and balance

[5]See Foley (1967) and Varian (1974) for earlier studies and Abdulkadiroglu and Sönmez (2003); Sun and Yang (2003); Andersson et al. (2010); Kesten and Yazici (2012); Sönmez and Switzer (2013); Fragiadakis and Troyan (2017) for recent studies of various matching models.

[6]This concept is studied by Balinski and Sönmez (1999); Ehlers et al. (2014); Kamada and Kojima (2017).

[7]See e.g., Hurwicz (1973); Dubins and Freedman (1981); Roth (1982); Roth and Sotomayor (1990); Abdulkadiroglu and Sönmez (2003); Ausubel (2006); Hatfield and Milgrom (2005); Sönmez and Switzer (2013); Sun and Yang (2014); Kamada and Kojima (2015, 2018); Fragiadakis and Troyan (2017).

the number of doctors among all hospitals and all regions. Kamada and Kojima (2015) initialized the use of such a rule for their model with ceiling constraints on regions. Our second mechanism is called the sequential deferred acceptance algorithm with distributional constraints (the SDA-D algorithm). This mechanism makes use of the DA-D algorithm repeatedly by adjusting doctors in hospitals and regions according to their demands. We prove that the matching found by the SDA-D algorithm is group stable and fair within the same quota type. We also find an impossibility result that regardless of whether a market faces ceiling constraints or floor constraints or both, there is no strategy-proof and nonwasteful mechanism that doctor-dominates the DA-D or SDA-D algorithm. Related results are due to Abdulkadiroğlu et al. (2009) and Kesten (2010). Furthermore, we compare our stability solution with two closely related concepts of Kamada and Kojima (2015) and Akin (2021).

1.1 Related Literature

We briefly review a number of previous studies on matching markets under distributional constraints. Ehlers et al. (2014) considered a school choice matching model in which schools set their lower bound and an upper bound for the number of students of each type. They showed the existence of nonwasteful and fair assignments of students to schools. Kamada and Kojima (2015) studied the Japanese medical residency matching market in which each region/hospital faces a ceiling constraint.[8] Their major result shows that their flexible deferred acceptance algorithm finds a stable matching and is strategy-proof for doctors. Kamada and Kojima (2017) proposed two notions of stability for a similar model with more general ceiling constraints. Kamada and Kojima (2018) identified necessary and sufficient conditions on the structure of ceiling constraints for matching models under which strategy-proof and stable mechanisms exist. Our model is close to these three models but differs from them in one crucial aspect that our model also permits each hospital/region to accommodate

[8] Regional ceiling or floor constraints in matching models make these models crucially different from those with ceiling or floor constraints only on hospitals or schools.

a floor constraint besides a ceiling constraint. Our solutions are considerably different from theirs. Akin (2021) investigated a doctor-hospital matching model with regional floor constraints where each hospital has a ceiling constraint but has no floor constraint. She introduced the floor respecting stability concept and proposed a modified deferred acceptance algorithm for finding her solution. Our model differs from this model in one crucial aspect that our model allows every hospital/region to have both a floor constraint and a ceiling constraint. Our solutions differ from hers substantially.

Motivated by the college admission in Hungary, Biró et al. (2010) studied a matching model in which each college has a floor constraint and a ceiling constraint on the number of students. They showed that checking whether a stable matching exists or not is NP-complete in the theory of computing complexity but found a condition under which a solution exists. See also Huang (2010) for a similar problem. Goto et al. (2015) presented an algorithm for improving fairness in nonwasteful matching for a market with regional minimum quotas; see also Goto et al. (2014). Fragiadakis et al. (2016) proposed strategy-proof mechanisms for a matching model in which each school has a floor constraint and a ceiling constraint. For a similar model, Fragiadakis and Troyan (2017) introduced a mechanism that possesses efficiency, fairness and incentive properties. Tomoeda (2018) presented a sufficient condition for the existence of stable matching in the sense of Ehlers et al. (2014).

This paper is organized as follows. Section 2 presents the model and its solutions. Section 3 contains the two mechanisms and their key properties. Section 4 discusses other properties of our mechanisms and compare our stability with others. Section 5 concludes. Most of the proofs are given in the appendix.

2 Model

2.1 Preliminaries

There are a finite set D of doctors with $n = |D|$, a finite set H of hospitals, and a finite set R of regions. Let h_0 stand for every doctor's outside option with an unlimited capacity. Each doctor d has a complete, strict, and transitive preference $\succ_d$ over $H \cup \{h_0\}$ with $h \succ_d h_0$ for all $h \in H$. For any $h, h' \in H$, $h \succeq_d h'$ if and only if $h \succ_d h'$ or $h = h'$. Let $\succ_D = (\succ_d)_{d \in D}$ be the preference profile of doctors. Each hospital h has a strict preference $\succ_h$ over $D \cup \{\emptyset\}$ with $\{d\} \succ_h \emptyset$ for all $d \in D$ and faces a ceiling quota $q(h)$ and a floor quota $l(h)$. We assume that the preference relation $\succ_h$ is *responsive* for every hospital h.[9] Let $\succ_H = (\succ_h)_{\in H}$ be the preference profile of hospitals. Let Ω denote the set of all agents' preference relations. Hospitals are located at different regions. Let $r(h) \in R$ be the region where hospital $h \in H$ is. Let H_r be the set of hospitals in region $r \in R$. That is, $H_r \cap H_{r'} = \emptyset$ if $r \neq r'$ and $H = \cup_{r \in R} H_r$.

A matching μ is a mapping that satisfies (i) $\mu(d) \in H \cup \{h_0\}$ for all $d \in D$, (ii) $\mu(h) \subseteq D$ for all $h \in H$, and (iii) for any $d \in D$ and $h \in H$, $\mu(d) = h$ if and only if $d \in \mu(h)$. Let $\mathcal{A}$ be the set of all matchings.

Hospitals and regions face distributional constraints. Let $L = (l(i))_{i \in H \cup R}$ and $Q = (q(i))_{i \in H \cup R}$ be the vector of floors and ceilings on hospitals and regions, respectively. These floor and ceiling constraints are hard constraints. Without loss of generality, we assume that $\sum_{h \in H_r} l(h) \leq l(r) \leq q(r)$ for all $r \in R$ and that the ceiling constraint of every region $r \in R$ has either $q(r) < \sum_{h \in H_r} q(h)$ or $q(r) \gg \sum_{h \in H_r} q(h)$.[10] Let R^b be the set of regions with ceiling constraints satisfying $q(r) < \sum_{h \in H_r} q(h)$ and R^u the set

[9] Hospital h's preference $\succ_h$ is responsive if, for any $S \subseteq D$, $d \in D \setminus S$ and $d' \in S$, we have $(S \cup \{d\}) \setminus \{d'\} \succeq_h S$ if and only if we have $\{d\} \succeq_h \{d'\}$. See Roth (1985); Roth and Sotomayor (1990); Kamada and Kojima (2015) for this widely-used condition. Note that here the definition of responsiveness is not related to the hospital's floor or capacity constraint.

[10] Note that in the latter case, if one can fill the ceiling capacity $q(h)$ of every hospital $h \in H_r$ in region r, we must have $q(r) \geq \sum_{h \in H_r} q(h)$ so it is harmless to set $q(r) \gg \sum_{h \in H_r} q(h)$.

of regions with ceiling constraints satisfying $q(r) \gg \sum_{h \in H_r} q(h)$.

We represent the described market by a tuple $\mathcal{M} = (D, H, \succ_D, \succ_H, L, Q)$ or simply by $\mathcal{M}$. Let Ω be the family of all such markets $\mathcal{M}$. Observe that the distributional constraints are not binding if (i) $l(h) = 0$ for every hospital $h \in H$; and (ii) $l(r) = 0$ and $\sum_{h \in r} q(h) \ll q(r)$ for every region $r \in R$. When distributional constraints are not binding, the market $(D, H, \succ_D, \succ_H, L, Q)$ reduces to a standard one with no distributional constraints.

A matching μ is **feasible** if $l(h) \leq |\mu(h)| \leq q(h)$ for all $h \in H$ and $l(r) \leq |\mu(r)| \leq q(r)$ for all $r \in R$, where $\mu(r) = \bigcup_{h \in H_r} \mu(h)$. Let $\mathcal{A}^f$ be the set of feasible matchings. We assume that the number of doctors is greater than or equal to the sum of every region's minimum quota, i.e., $n \geq \sum_{r \in R} l(r)$. So we have at least one feasible matching. A matching μ is **individually rational** if (i) $\mu(d) \succeq_d h_0$ for all $d \in D$ and (ii) $\{d\} \succeq_h \emptyset$ for all $d \in \mu(h)$. Note that every feasible matching is individually rational. A feasible matching μ is **(Pareto) efficient** if there does not exist a feasible matching ν such that $\nu(i) \succeq \mu(i)$ for all $i \in D \cup H$ and $\nu(i) \succ \mu(i)$ for some $i \in D \cup H$. A related efficiency concept is non-wastefulness. A feasible matching μ is **wasteful** if there is a doctor-hospital pair (d, h) with $h \succ_d \mu(d)$ and $|\mu(h)| < q(h)$ and if one can make a new feasible matching ν with $\nu(d) = h$ and $\nu(k) = \mu(k)$ for all $k \in D \setminus \{d\}$. This implies that it is possible to match the wasteful pair and improve their welfare without violating any constraint. A feasible matching is **non-wasteful** if there does not exist any wasteful doctor-hospital pair.

A **mechanism** φ is a procedure that generates a matching $\varphi(\succ_D, \succ_H) \in \mathcal{A}$ for every instance $(\succ_D, \succ_H) \in \Omega$. A mechanism φ is feasible if $\varphi(\succ_D, \succ_H) \in \mathcal{A}^f$ for every $(\succ_D, \succ_H) \in \Omega$. A mechanism φ is **strategy-proof for doctors** if $\varphi_i(\succ_D, \succ_H) \succeq_i \varphi_i(\succ'_i, \succ_{-i}; \succ_H)$ for all $i \in D$. It is well-known that no mechanism can be both strategy-proof and stable for both sides of a market and all possible preference profiles even for the basic Gale-Shapley marriage matching model; see e.g., Roth and Sotomayor (1990). It is common and natural to focus on the doctor side's incentive.

In order to design a mechanism for finding a desired matching between doctors, hospitals, and regions that satisfies the distributional constraints, we need to introduce a recruitment quota system consisting of hospital rigid quota τ_{rh} for $h \in H$, region rigid quota τ_{rr} for $r \in R$, and region elastic quota τ_{er} for $r \in R$. These quota types are represented by $\mathcal{T}_H = \{\tau_{rh}\}_{h\in H}$, $\mathcal{T}_R = \{\tau_{rr}\}_{r\in R}$, and $\mathcal{T}_{RE} = \{\tau_{er}\}_{r\in R}$, respectively. Let τ_o stand for the null quota type which has a sufficiently large capacity, say, at least the number n of all doctors. Let $\mathcal{T} = \mathcal{T}_H \cup \mathcal{T}_R \cup \mathcal{T}_{RE}$ be the set of all quota types. The quota of every $\tau \in \mathcal{T}$ will be denoted by $\kappa(\tau)$. For each hospital h, we define its rigid quota by $\kappa(\tau_{rh}) = l(h)$ the floor bound of doctors for the hospital that must be occupied. For each region r, we define its rigid quota by $\kappa(\tau_{rr}) = l(r) - \sum_{h\in H_r} l(h)$ the difference between the floor number of the region and the sum of the floor number of every hospital in the region. This number $\kappa(\tau_{rr})$ of doctors can be distributed to any hospital in the region and must be fulfilled. For each region r, we define its elastic quota by $\kappa(\tau_{er}) = q(r) - l(r)$ the gap between its ceiling number and its floor number. This elastic quota of doctors has some flexibility so need not be distributed.

2.2 Stability

Stability is a key solution concept used in matching models. The standard stability concept of Gale and Shapley (1962) has been extensively studied in the literature. Recall from Roth and Sotomayor (1990) that given a matching μ, a doctor-hospital pair (d, h) is a blocking pair if (i) $h \succ_d \mu(d)$; (ii) either $\{d\} \succ_h \{d'\}$ for some $d' \in \mu(h)$ or $|\mu(h)| < q(h)$. A matching is (pairwise) stable if it is individually rational and there is no blocking pair. Note that with an abuse of notation, we write a doctor-hospital pair as (d, h) meaning the set $\{d, h\}$.

An intuitive and straightforward way of adapting the above definition of blocking pair to the current model with distributional constraints is to ensure that when a feasible matching is blocked by a doctor-hospital pair, we should match the pair together and make the new matching respect all the constraints. Formally,

Definition 1 *A feasible matching μ is* ***perfectly blocked by a doctor-hospital pair*** (d, h) *if (i) $h \succ_d \mu(d)$; (ii) either $|\mu(h)| < q(h)$ or $\{d\} \succ_h \{d'\}$ for some $d' \in \mu(h)$; and (iii) there* *is a feasible matching ν that satisfies either $\nu(d) = h$ and $\nu(k) = \mu(k)$ for all $k \in D \setminus \{d\}$, or $\nu(d) = h$, $\nu(d') = h_0$ and $\nu(k) = \mu(k)$ for all $k \in D \setminus \{d, d'\}$. The pair (d, h) is called a perfect blocking pair.*

A feasible matching is *perfectly stable* if it is not perfectly blocked by any doctor-hospital pair. Appealing as this concept may appear, the following example shows that it fails to exist.

Example 1 *There are two doctors $D = \{d_1, d_2\}$ and two hospitals $H = \{h_1, h_2\}$ in one region with $q(h_1) = q(h_2) = 1$, $l(h_1) = l(h_2) = 0$, and $q(r) = 1$ and $l(r) = 0$. The preferences of doctors and hospitals are given by*

$$\succ_{d_1}: h_2, h_1 \qquad \succ_{d_2}: h_1, h_2$$

$$\succ_{h_1}: d_1, d_2 \qquad \succ_{h_2}: d_2, d_1$$

We will see that this example has four feasible matchings given in Table 1 but has no perfectly stable matching.

Table 1: Feasible matchings in Example 1.

Matching	h_1	h_2	h_0
μ_1	d_1	$\emptyset$	d_2
μ_2	$\emptyset$	d_1	d_2
μ_3	d_2	$\emptyset$	d_1
μ_4	$\emptyset$	d_2	d_1

Let us show that every feasible matching is perfectly blocked by a doctor-hospital pair. Indeed, μ_1 is perfectly blocked by (d_1, h_2), as $h_2 \succ_{d_1} \mu_1(d_1) = h_1$, $|\mu_1(h_2)| < q(h_2)$, matching d_1 and h_2 and keeping the rest intact makes a new feasible matching. Similarly, μ_4 is perfectly blocked by (d_2, h_1), as $h_1 \succ_{d_2} \mu_4(d_2) = h_2$, $|\mu_4(h_1)| < q(h_1)$, matching d_2 and h_1 and keeping the rest unchanged makes a new feasible matching

ν. μ_2 is perfectly blocked by (d_2, h_2), as $h_2 \succ_{d_2} \mu_2(d_2) = h_0$, $d_2 \succ_{h_2} d_1$, and matching h_2 and d_2 by rejecting d_1 makes a new feasible matching ν. Similarly, μ_3 is perfectly blocked by (d_1, h_1), as $h_1 \succ_{d_1} \mu_3(d_1) = h_0$, $d_1 \succ_{h_1} d_2$, and matching h_1 and d_1 by rejecting d_2 makes also a new feasible matching. This demonstrates that there does not exist any perfectly stable matching.

The above example suggests that the concept of doctor-hospital blocking pair should be modified in a more sensible way. To achieve this, we first introduce the following two new concepts.

Definition 2 *Given a feasible matching μ, we say a new matching ν is **permissible at** μ if there is a nonempty group D' of doctors such that $\nu(d) \neq \mu(d)$ for all $d \in D'$, $\nu(d') = \mu(d')$ for all $d' \in D \setminus D'$, $|\nu(h)| \geq l(h)$ for all $h \in H$, and $l(r) \leq |\nu(r)| \leq q(r)$ for all $r \in R$.*

By definition, a new matching ν is permissible at μ if the reassignment of a group D' of doctors without altering any other doctor's assignment violates neither floor constraints on hospitals nor distributional constraints on regions. It is possible that the number of doctors in some hospital may exceed its ceiling constraint. As both ceiling and floor constraints on every region are satisfied, it is always possible to make every permissible matching ν into a new feasible matching when matching μ is feasible.

Definition 3 *A feasible matching μ is **admissibly blocked by a doctor-hospital pair** (d, h) if (i) $h \succ_d \mu(d)$; (ii) either $|\mu(h)| < q(h)$ or $\{d\} \succ_h \{d'\}$ for some $d' \in \mu(h)$; and (iii) there is a permissible matching ν that satisfies $\nu(d) = h$ and $\nu(k) = \mu(k)$ for all $k \in D \setminus \{d\}$. The pair (d, h) is called an admissible blocking pair of μ.*

In the current model, all distributional constraints are hard constraints so must be respected. The definition says that a pair of doctor d and hospital h can possibly block a feasible matching μ if they both can cooperate to make themselves better off by breaking almost none of the hard constraints and requiring no aid of any other member.

By definition, if the number of doctors in hospital h at μ is less than its capacity $q(h)$, then hospital h hires doctor d and the new matching ν is a feasible matching. However, if the number of doctors in hospital h at μ is equal to its capacity $q(h)$, then hospital h hires doctor d but still keeps doctor d' and the matching ν satisfies all the distributional constraints except the capacity constraint of hospital h. Counter-intuitive as it may appear, this temporary violation of the capacity constraint is actually very common and mild, because in practice many firms give their dismissed employees a grace period of time to leave. In fact, many organisations have an official and open policy that if they decide to terminate an employee's job, they will inform the concerned employee of the decision several weeks or months in advance. In the definition, as ν satisfies all other constraints, it is possible to fire d' or assign her to work at some hospital so that all the distributional constraints will be satisfied. Namely, there always exists a feasible matching in which h accepts d and rejects d'.

Observe that in the admissible blocking pair (d, h) of μ, doctor d either has no position (i.e $\mu(d) = h_0$) or works at a hospital $\mu(d) \neq h_0$ whose number of hired doctors at μ exceeds its floor constraint $l(\mu(d))$. Therefore d can leave hospital $\mu(d)$ without violating the constraint. There are two possible cases: Case 1. The pair (d, h) are in the same region. In this case, all doctors and hospitals in the other regions will remain unchanged as in μ and hiring d to work at h will not change the number of doctors in the region. Case 2. The pair (d, h) are in different regions. In this case, the number of hired doctors at μ in the hospital h's region is strictly below its regional capacity. Therefore, d can leave his hospital $\mu(d)$ and move out of the region to work at hospital h without violating any floor constraints on hospitals or any distributional constraints on any region.

From the above discussion it is easy to see that every admissible blocking pair must be a perfect blocking pair. We will show by Example 1 that the reverse is not true. Clearly, perfect blocking pairs (d_1, h_2) and (d_2, h_1) are also admissible blocking pairs. However, the perfect blocking pair (d_2, h_2) of μ_2 is not an admissible blocking pair, because h_2's accepting d_2 without rejecting d_1 will violate the region's ceiling constraint. Similarly, the perfect blocking pair (d_1, h_1) of μ_3 is not an admissible blocking pair, because h_1's accepting d_1 without rejecting d_2 will also violate the region's ceiling constraint. This shows that perfect blocking pairs are strictly more general than admissible blocking pairs so are hard for perfectly stable matchings to exist.

In the following, we refer *distributional constraints to the floor constraints on hospitals and both floor and ceiling constraints on regions* in order to ease description.

Definition 4 *A feasible matching is **(pairwise) stable** if it is not admissibly blocked by any doctor-hospital pair.*

It is easy to check that in Example 1, matchings μ_2 and μ_3 are stable but matchings μ_1 and μ_4 are not. To facilitate a better understanding of the above concepts, let us consider the following more general example.

Example 2 *There are three doctors $D = \{d_1, d_2, d_3\}$ and four hospitals $H = \{h_1, h_2, h_3, h_4\}$ located in two regions $r_1 = \{h_1, h_2\}$ and $r_2 = \{h_3, h_4\}$ with $q(h_1) = q(h_2) = q(h_3) = q(h_4) = 1$, $l(h_1) = l(h_2) = 1$, $l(h_3) = l(h_4) = 0$, $q(r_1) = 10$, $q(r_2) = 1$, $l(r_1) = 2$, and $l(r_2) = 0$. The preferences of doctors and hospitals are given by*

$$\succ_{d_1}: h_4, h_3, h_1, h_2 \qquad \succ_{d_2}: h_2, h_3, h_4, h_1 \qquad \succ_{d_3}: h_3, h_4, h_1, h_2$$

$$\succ_{h_1}: d_1, d_3, d_2 \qquad \succ_{h_2}: d_2, d_3, d_1 \qquad \succ_{h_3}: d_2, d_1, d_3 \qquad \succ_{h_4}: d_3, d_2, d_1$$

Table 2 lists all feasible matchings. We need to consider the following two cases.

Case (1): $|\mu(h)| < q(h)$. For instance, (d_3, h_3) is an admissible blocking pair of μ_2 as $h_3 \succ_{d_3} h_4$, $|\mu(h_3)| < q(h_3)$, and moving d_3 to h_3 is feasible. However, the blocking pair (d_1, h_4) of matching μ_1 is not admissible, because moving d_1 to h_4 will violate the floor constraint of h_1.

Case (2): $\{d\} \succ_h \{d'\}$ **for some** $d' \in \mu(h)$. In this case, when there are only ceiling constraints, matching d and h by rejecting d' is feasible. However, it may not be feasible when there are floor constraints. Consider matching μ_9. There are two possible blocking pairs (d_1, h_1) and (d_2, h_2) against μ_9. Matching d_2 and h_2 and keeping d_1 at h_2 makes an admissible blocking pair. In this case, d_2 leaves h_3 and region r_2 to join h_2 in region r_1. As h_2 still keeps d_1, it overuses its capacity but all other constraints are satisfied. Observe that the pair (d_1, h_1) is not admissible because d_1's leaving h_2 violates the floor constraint of h_2.

In this example, μ_1, μ_5, μ_8, and μ_{12} are stable, as none of them has any admissible blocking pair against it.

Table 2: Feasible matchings in Example 2.

$\mathcal{A}^f$	r_1		r_2		Admissible blocking pair	Admissible blocking coalition
	h_1	h_2	h_3	h_4	(pairwise stability)	(group stability) when $\|S\| > 2$
μ_1	d_1	d_2	d_3	$\emptyset$	$\emptyset$	$\varnothing$
μ_2	d_1	d_2	$\emptyset$	d_3	(d_3, h_3)	$\varnothing$
μ_3	d_1	d_3	d_2	$\emptyset$	(d_2, h_2)	$\varnothing$
μ_4	d_1	d_3	$\emptyset$	d_2	(d_2, h_3), (d_2, h_2)	$\{d_2, h_2, d_3, h_3\}$, $\{d_2, h_2, d_3, h_4\}$
μ_5	d_2	d_1	d_3	$\emptyset$	$\emptyset$	$\{d_2, h_2, d_1, h_1\}$
μ_6	d_2	d_1	$\emptyset$	d_3	(d_3, h_3)	$\{d_2, h_2, d_1, h_1\}$, $\{d_2, h_2, d_1, h_1.d_3, h_3\}$
μ_7	d_2	d_3	d_1	$\emptyset$		$\{d_2, h_2, d_3, h_1\}$
μ_8	d_2	d_3	$\emptyset$	d_1	$\emptyset$	$\{d_2, h_2, d_3, h_1\}$
μ_9	d_3	d_1	d_2	$\emptyset$	(d_2, h_2)	$\{d_1, h_4, d_2, h_2\}$, $\{d_3, h_4, d_1, h_1, d_2, h_2\}$
μ_{10}	d_3	d_1	$\emptyset$	d_2	(d_2, h_2), (d_2, h_3)	$\{d_1, h_3, d_2, h_2\}$, $\{d_3, h_4, d_1, h_1, d_2, h_2\}$ $\{d_3, h_3, d_1, h_1, d_2, h_2\}$
μ_{11}	d_3	d_2	d_1	$\emptyset$	(d_1, h_4)	$\varnothing$
μ_{12}	d_3	d_2	$\varnothing$	d_1	$\emptyset$	$\varnothing$

In this example, admissible blocking coalitions S for each matching include admissible blocking pairs and admissible blocking coalitions of $|S| > 2$.

The following observation follows immediately from the definition of pairwise stability.

Lemma 1 *A pairwise stable matching must be nonwasteful.*

Note that in the previous example, (pairwise) stability may not be strong enough to exclude some undesirable outcomes. For instance, the stable matching μ_5 is inefficient as it is Pareto dominated by μ_1 or blocked by the coalition $\{d_1, h_1, d_2, h_2\}$. The stable matching μ_8 is also inefficient as it is Pareto dominated by μ_{12} or blocked by the coalition $\{d_2, h_2, d_3, h_1\}$. This shows that inefficiency can be caused by the limited scope of blocking pairs. It is therefore necessary and important to consider larger coalitions. We introduce a new notion of group stability.

Definition 5 *A feasible matching μ is* ***admissibly blocked by a coalition*** *$S \subseteq D \cup H$ consisting a group of doctors and hospitals, if there exists a redistribution ν^S among the agents in the coalition such that*

(i) $\nu^S(d) \in S$ for every $d \in S \cap D$;

(ii) for every $h \in S \cap H$, $d \in \nu^S(h)$ implies $d \in S \cup \mu(h)$;

(iii) $\nu^S(i) \succeq_i \mu(i)$ for all $i \in S$ and $\nu^S(i) \succ_i \mu(i)$ for some $i \in S$;

(iv) matching ν given by $\nu(i) = \nu^S(i)$ for all $i \in S \cap D$ and $\nu(i) = \mu(i)$ for all $i \in D \setminus S$ is permissible at μ.

A matching μ is blocked by a coalition if there exits a redistribution among coalition members such that each coalition member becomes weakly better off and some coalition member becomes strictly better off. Similar to pairwise stability, a coalition S is admissible to block a matching μ if assigning every coalition member $i \in S$ with the redistribution $\nu^S(i)$ and maintaining every other doctor's (outside the coalition) assignment as the same as in μ will not violate any distributional constraint.

Definition 6 *A feasible matching is* ***group stable*** *if it cannot be admissibly blocked by any coalition.*

Let us look at Example 2 again. In this example, μ_1 and μ_{12} are group stable as neither of them can be admissibly blocked by any coalition

Lemma 2 *Every group stable matching is (pairwise) stable, efficient, and non-wasteful.*

Proof: Let μ be a group stable matching. By definition, μ is feasible and therefore individually rational. Suppose μ is not pairwise stable. By definition, there exists a doctor-hospital pair (d, h) blocking μ admissibly. That is, $\{d, h\}$ blocks μ, contradicting the assumption. Suppose matching μ is not efficient. There exists another feasible matching ν such that $\nu(i) \succeq_i \mu(i)$ for every $i \in D \cup H$ and $\nu(i) \succ_i \mu(i)$ for some $i \in D \cup H$. By definition, the grand coalition $S = D \cup H$ blocks the matching μ, contradicting the assumption. □

Lemma 3 *Efficiency and nonwastefulness do not imply each other.*

Proof: We prove by example. Suppose that there are two doctors $D = \{d_1, d_2\}$ and two hospitals $H = \{h_1, h_2\}$ located in one region r. Their quotas are given by $q(h_1) = 1, q(h_2) = 2$, $l(h_1) = l(h_2) = 0$, $q(r) = 20$, and $l(r) = 0$. The preferences of doctors and hospitals are given by

$$\succ_{d_1}: h_2, h_1 \qquad \succ_{d_2}: h_1, h_2$$

$$\succ_{h_1}: d_2, d_1 \qquad \succ_{h_2}: d_1, d_2$$

Consider the matching μ satisfying $\mu(d_1) = \mu(d_2) = h_2$ and $\mu(h_2) = \{d_1, d_2\}$. Clearly, we have $|\mu(h_2)| = q(h_2)$. μ is efficient, because any other feasible matching ν must satisfy $|\nu(h_2)| < q(h_2)$. Such matching ν makes h_2 worse off than at μ. However, μ is not non-wasteful, because the pair (d_2, h_1) have $h_1 \succ_{d_2} h_2$, $|\mu(h_1)| < q(h_1)$ and we can make a new feasible matching ν satisfying $\nu(d_2) = h_1$, $\nu(h_1) = d_2$, $\nu(d_1) = h_2$ and $\nu(h_2) = d_1$. This shows that an efficient matching is not necessarily nonwasteful.

We now prove that a nonwasteful matching need not be efficient. We use the same example by only modifying the capacity of h_2 from 2 to 1. Consider a matching ρ satisfying $\rho(d_1) = h_1$, $\rho(h_1) = d_1$, $\rho(d_2) = h_2$ and $\rho(h_2) = d_2$. This matching ρ is non-wasteful, because $|\rho(h)| = q(h)$ for all $h \in \{h_1, h_2\}$. However, ρ is not efficient, because another feasible matching ρ' satisfying $\rho'(d_1) = h_2$, $\rho(h_2) = d_1$, $\rho'(d_2) = h_1$ and $\rho'(h_1) = d_2$ makes all doctors and hospitals better off. □

2.3 Fairness

Fairness is a fundamental concept concerning equitable distribution of resources and has long and widely been explored in the literature. We will adapt this concept to our current model. Given a matching μ, we say that doctor d and hospital h who are unmatched at μ form **an envious pair** if $h \succ_d \mu(d) = h'$ and $d \succ_h d'$ for some $d' \in \mu(h)$. Roughly speaking, the condition means that doctor d envies every doctor hired by hospital h and hospital h envies hospital h' which hires doctor d.

Definition 7 *A matching is **fair** if there does not exist any envious doctor-hospital pair.*

The following result shows incompatibility between stability and fairness.

Lemma 4 *In general, there does not exist a matching that is both stable and fair.*

Proof: We prove this by the following example. There are two doctors d_1, d_2, three hospitals h_1, h_2, h_3, and two regions r_1, r_2 with $r_1 = \{h_1\}$ and $r_2 = \{h_2, h_3\}$. The maximum and minimum quotas for hospitals and regions are given by $q(h_1) = q(h_2) = q(h_3) = 1$, $l(h_1) = l(h_2) = l(h_3) = 0$, $q(r_1) = q(r_2) = 20$, $l(r_1) = 1$, and $l(r_2) = 0$. The preferences of doctors and hospitals are given by

$$\succ_{d_1}: \; h_2, \; h_3, \; h_1 \qquad \succ_{d_2}: \; h_3, \; h_2, \; h_1$$
$$\succ_{h_1}: \; d_2, \; d_1 \qquad \succ_{h_2}: \; d_2, \; d_1 \qquad \succ_{h_3}: \; d_1, \; d_2$$

All feasible matching are shown in Table 3. Matchings μ_2 and μ_4 are (pairwise) stable but not fair. While matchings μ_1 and μ_3 are fair but not stable. □

Table 3: Feasible Matchings

Matching	h_1	h_2	h_3
μ_1	d_1	d_2	$\emptyset$
μ_2	d_1	$\emptyset$	d_2
μ_3	d_2	$\emptyset$	d_1
μ_4	d_2	d_1	$\emptyset$

The above lemma indicates that fairness is too strong for stable matchings. To maintain stability, we need to relax the notion of fairness. Recall that τ_0 stands for the null quota type which can have any large positive integer quota, say no less than the number n of all doctors. A hospital can hire a doctor beyond its rigid quota only if a quota from the region rigid quota and elastic quota is distributed to the hospital. In other words, assigning a doctor to a hospital is possible if the hospital is given a quota from the hospital's rigid quota, its region rigid or elastic quota. A mapping $\sigma : D \to \mathcal{T} \cup \{\tau_0\}$ is a **quota mapping** if it satisfies $|\sigma^{-1}(\tau)| \leq \kappa(\tau)$ for all $\tau \in \mathcal{T}$. The condition says that for any quota type τ, the number of doctors who are assigned the quota type should be no greater than the quota $\kappa(\tau)$.

We can also use a pair (μ, σ) of matching μ and quota mapping σ to represent an outcome if (i) $\sigma(d) \in \{\tau_{rh}, \tau_{rr(h)}, \tau_{er(h)}\}$ for all $d \in \mu(h)$ and $h \in H$; (ii) $\sigma(d) = \tau_0$ for all $d \in \mu(h_0)$; and (iii) $|\sigma^{-1}(\tau)| \leq \kappa(\tau)$ for all $\tau \in \mathcal{T}_{RE}$, and $|\sigma^{-1}(\tau)| = \kappa(\tau)$ for all $\tau \in \mathcal{T}_H \cup \mathcal{T}_R$. Condition (i) says that every doctor hired by a hospital has either the hospital's rigid quota type or its regional rigid or elastic quota type. Condition (ii) states that any unemployed doctor must be assigned the null quota type. Condition (iii) says that the number of doctors who are assigned any regional elastic quota type is not more than the elastic quota, and that the number of doctors who are assigned any rigid quota type must equal the rigid quota. We can now introduce the following concept of fairness with respect to every outcome.

Definition 8 *An outcome (μ, σ) is **fair within the same quota type** if there does not exist any doctor-hospital pair (d, h) such that $d, d' \in \sigma^{-1}(\tau)$, $\mu(d') = h \succ_d \mu(d) = h'$ and $\{d\} \succ_h \{d'\}$ for some type $\tau \in \mathcal{T}_R \cup \mathcal{T}_{RE}$.*

An outcome (μ, σ) is fair within the same quota type if there does not exist any pair (d, h) of doctor d and hospital h unmatched at μ such that doctor d prefers h to his current match $\mu(d)$ and hospital h prefers d to any doctor d' hired by the hospital at μ who shares the same quota type τ with doctor d. In other words, doctors who share the same quota type will not envy each other and to some extent the associated hospitals will not envy others. We use the following example to illustrate this concept.

Example 3 *There are three doctors* d_1, d_2*, three hospitals* h_1, h_2, h_3*, and two regions* r_1, r_2 *with* $r_1 = \{h_1\}$ *and* $r_2 = \{h_2, h_3\}$*. The maximum and minimum quotas for hospitals and regions are given by* $q(h_1) = q(h_2) = q(h_3) = 1$*,* $l(h_1) = l(h_2) = l(h_3) = 0$*,* $q(r_1) = 8$*,* $q(r_2) = 8$*,* $l(r_1) = 1$*, and* $l(r_2) = 0$*. The preferences of doctors and hospitals are given by*

$$\succ_{d_1}: \ h_3, \ h_2, \ h_1 \qquad \succ_{d_2}: \ h_3, \ h_2, \ h_1 \qquad \succ_{d_3}: \ h_3, \ h_2, \ h_1$$
$$\succ_{h_1}: \ d_2, \ d_1, \ d_3 \qquad \succ_{h_2}: \ d_2, \ d_3, \ d_1 \qquad \succ_{h_3}: \ d_3, \ d_1, \ d_2$$

In this example, there are no rigid hospital quota, one region rigid quota τ_{rr_1} of r_1, seven region elastic quota τ_{er_1} of r_1, and eight region elastic quotas τ_{er_2} of r_2. Any feasible matching μ with $\mu(h_2) = \emptyset$ or $\mu(h_3) = \emptyset$ is neither efficient nor pairwise stable. Therefore, we will only consider efficient matchings shown in Table 4.

In the outcome (μ_2, σ_2), doctor d_1 obtains the region rigid quota τ_{rr_1} of r_1; d_3, d_2 obtain the region elastic quotas τ_{er_2} of r_2. We can see that this outcome is not fair within the same quota type because there exists a doctor-hospital pair (d_3, h_3) such that $d_3, d_2 \in \sigma^{-1}(\tau_{er_2})$, $\mu(d_2) = h_3 \succ_{d_3} h_2$ and $d_3 \succ_{h_3} d_2$. Similarly, we can show that matchings μ_3 and μ_5 are not fair within the same quota type. Matchings μ_1, μ_4, and μ_6 are all group stable and fair within the same quota type. In fact, μ_1 is fair but μ_4 and μ_6 are not, because at μ_4 doctor d_2 envies d_1 and at μ_6 doctor d_3 envies d_1.

Table 4: Efficient matchings in which all doctors are recruited.

Matching	h_1	h_2	h_3	Quota Matching	τ_{rr_1}	τ_{er_2}	τ_{er_2}
μ_1	d_1	d_2	d_3	σ_1	d_1	d_2	d_3
μ_2	d_1	d_3	d_2	σ_2	d_1	d_3	d_2
μ_3	d_2	d_3	d_1	σ_3	d_2	d_3	d_1
μ_4	d_2	d_1	d_3	σ_4	d_2	d_1	d_3
μ_5	d_3	d_1	d_2	σ_5	d_3	d_1	d_2
μ_6	d_3	d_2	d_1	σ_6	d_3	d_2	d_1

3 Main Results

In this section we propose two mechanisms. We will show that the first mechanism finds a fair and efficient matching and is strategy-proof and the second mechanism finds a group stable and fair matching within the same quota type. The second mechanism is built upon the first one. The first mechanism can be seen as significant generalization of the deferred acceptance (DA) algorithm of Gale and Shapley (1962) from the setting without distributional constraints to the setting with complex distribution constraints. In each step of our mechanisms, every hospital can hire doctors by using either its rigid quota, region rigid quota or elastic quota. In the mechanisms a picking order rule will be used in order to balance region rigid quotas and elastic quotas among hospitals and regions. Kamada and Kojima (2015) introduced a first picking order rule for their model with regional ceiling constraints.

Let $rank_{D'}(d,h)$ be the ranking of every doctor d in a given set D' on the preference list of any given hospital h. For instance, if d is hospital h's first choice among all doctors in D', then $rank_{D'}(d,h) = 1$.

Let $B_h \subseteq D$ be a set of doctors who apply for a position at hospital h. Let $B = \bigcup_{h \in H} B_h$ be the set of all such doctors. Let $\succ^t$ be a tie-breaking order of hospitals, for instance, by using a lottery. This determines an order by which doctors will be assigned to hospitals when there is a tie between hospitals. A picking order $\succ_B$ is constructed such that for any $d \in B_h$ and $d' \in B_{h'}$, we have $d \succ_B d'$ if either $rank_{B_h}(d,h) < rank_{B_{h'}}(d',h')$ or $rank_{B_h}(d,h) = rank_{B_{h'}}(d',h')$ and $h \succ^t h'$ for $h \neq h'$. This picking order stipulates that first every hospital h calculates its own ranking $rank_{B_h}(d,h)$ for every doctor $d \in B_h$ and then all doctors in B will be ranked such that a doctor with a higher ranking should be ranked higher than any lower ranked doctor and a doctor with a higher ranked hospital should be ranked higher than any doctor of the same ranking with a lower ranked hospital.

Given a tie-breaking order $\succ^t$, we can now present our first mechanism, which will be the key building block of our second mechanism.

Deferred Acceptance Algorithm with Distributional Constraints (The DA-D Algorithm)

Step 1. Every doctor who has not yet made any proposal or has been rejected previously proposes to her favorite hospital that she has not previously proposed to. Otherwise, she will not make any proposal but match with h_0. Go to Step 2.

Step 2. (Distribute hospital rigid quota) Every hospital $h \in H$ selects its favorite doctors up to its capacity $q(h)$ from those who have just proposed to the hospital or have been tentatively accepted by the hospital, but rejects all other doctors who have also proposed to it. Let M_h be the set of doctors who are currently selected by the hospital. If there exists any rejection, go back to Step 1. Otherwise, every hospital h tentatively accepts the first $\min\{\kappa(\tau_{rh}), |M_h|\}$ top doctors from M_h by using its rigid quota. Let $\bar{D}_h$ denote this set of tentatively accepted top doctors. If every doctor is tentatively accepted by a hospital, go to Step 4. Otherwise, go to Step 3.

Step 3. Let $B_h = M_h \setminus \bar{D}_h$ for every hospital h and $B = \cup_{h \in H} B_h$. Based on the given tie-breaking order of hospitals $\succ^t$, construct a picking order $\succ_B$. Do the following operation:

Process 1. (Distribute region rigid quota) Pick doctors from B one by one according to the picking order $\succ_B$ and assign each of them to the hospital which has selected her, until there is no region rigid quota. Let $\bar{B}$ be the set of doctors who are assigned to hospitals in this step. Any doctor who is assigned to a hospital in this step is tentatively accepted by the hospital. If $\bar{B} = B$, go to Step 4. Otherwise, go to Process 2.

Process 2. (Distribute regional elastic quota) Pick at most the number of

$$\min\{n - \sum_{r \in R} l(r), \sum_{r \in R} \min\{q(r) - l(r), \sum_{h \in H_r} q(h) - l(r)\}\} \quad (1)$$

doctors one by one from the remaining ones in $B \setminus \bar{B}$ according to the picking order $\succ_B$ and assign each of them to the hospital which has selected her, until there is no regional elastic quota. Any doctor who is assigned to a hospital in this step is tentatively accepted by the hospital. Every doctor in $B \setminus \bar{B}$ who is not tentatively accepted will be rejected by her proposed hospital. If there is no rejection, go to Step 4. Otherwise, go to Step 1.

Step 4. In the end, any doctor who has been tentatively accepted by a hospital will match with the hospital. Any doctor who has not been accepted by any hospital will have no position. Any hospital which has not accepted any doctor will have no doctor. The algorithm stops.

Note that in formula (1) the first term $n - \sum_{r \in R} l(r)$ is the number of doctors still available after they fill every region's minimum quota, and the second term reflects that the total number of doctors can be possibly filled after doctors fill every hospital's minimum quota and every region's minimum quota. The minimum of the two terms is the actual number of doctors that will be distributed after they fill every hospital's minimum quota and every region's minimum quota. We call the number *the total distributable elastic (TDE) quota*. We can rewrite formula (1) as follows:

$\min\{n - \sum_{r \in R} l(r), \sum_{r \in R} \min\{q(r) - l(r), \sum_{h \in H_r} q(h) - l(r)\}\}$
$= \min\{n, \sum_{r \in R} \min\{q(r), \sum_{h \in H_r} q(h)\}\} - \sum_{r \in R} l(r)$.

The following result implies that the DA-D algorithm will generate a feasible matching.

Proposition 1 *In each round when the DA-D algorithm processes to Step 3, we have*

$$|B \setminus \bar{B}| \geq \min\{n, \sum_{r \in R} \min\{q(r), \sum_{h \in H_r} q(h)\}\} - \sum_{r \in R} l(r).$$

The algorithm distributes rigid and elastic quotas among hospitals and regions and checks whether any distributional constraint is violated or not. We have the following result for the algorithm.

Theorem 1 *A matching μ produced by the DA-D algorithm is feasible, fair and efficient.*

Theorem 2 *The DA-D algorithm is strategy-proof for all doctors.*

Now we give an example to illustrate the process of the DA-D algorithm.

Example 4 *There are six doctors* $D = \{d_1, d_2, d_3, d_4, d_5, d_6\}$ *and five hospitals* $H = \{h_1, h_2, h_3, h_4, h_5\}$ *located in two regions* $R = \{r_1, r_2\}$ *with* $r_1 = \{h_1, h_2\}$ *and* $r_2 = \{h_3, h_4, h_5\}$. *The maximum and minimum quotas for hospitals and regions are given by* $q(h_1) = q(h_2) = q(h_5) = 2$, $q(h_3) = 3$, $q(h_4) = 1$, $l(h_1) = l(h_2) = l(h_3) = 0$, $l(h_4) = l(h_5) = 1$, $q(r_1) = 2$, $q(r_2) = 20$, $l(r_1) = 2$, *and* $l(r_2) = 2$. *The preferences of doctors and hospitals are given by*

$\succ_{d_1}$	$\succ_{d_2}$	$\succ_{d_3}$	$\succ_{d_4}$	$\succ_{d_5}$	$\succ_{d_6}$	$\succ_{h_1}$	$\succ_{h_2}$	$\succ_{h_3}$	$\succ_{h_4}$	$\succ_{h_5}$
h_1	h_1	h_2	h_2	h_3	h_3	d_3	d_1	d_4	d_5	d_3
h_2	h_2	h_1	h_1	h_4	$\vdots$	d_4	d_2	d_6	d_2	d_1
$\vdots$	h_3	$\vdots$	h_4	$\vdots$		d_2	d_4	d_5	d_4	d_2
	h_4		h_3			d_1	d_3	$.d_2$	$\vdots$	$\vdots$
	h_5					$\vdots$	$\vdots$	$\vdots$		

According to the distributional constraints, we can calculate the number of quotas with each type shown in Table 5.

Table 5: Quotas for Example 4.

Region	r_1		r_2		
Hospital	h_1	h_2	h_3	h_4	h_5
region elastic quota	$\kappa(\tau_{er_1}) = q(r_1) - l(r_1) = 0$		$\kappa(\tau_{er_2}) = 18$		
region rigid quota	$\kappa(\tau_{rr_1}) = l(r_1) - \sum_{h \in H_{r_1}} l(h) = 2$		$\kappa(\tau_{rr_2}) = 0$		
hospital rigid quota	$\kappa(\tau_{rh_1}) = 0$	$\kappa(\tau_{rh_2}) = 0$	$\kappa(\tau_{rh_3}) = 0$	$\kappa(\tau_{rh_4}) = 1$	$\kappa(\tau_{rh_5}) = 1$

The process of the DA-D algorithm is shown in Table 6. In the table, we record each round of the algorithm and count the process of the algorithm as a new round as soon as it goes back to Step 1. In each round, at most

$$\min\{n - \sum_{r \in R} l(r), \sum_{r \in R} \min\{\kappa(\tau_{er}), \sum_{h \in H_r} q(h) - l(r)\}\} = 2$$

the total distributable elastic (TDE) quota can be used in each region. The tie-breaking order for hospitals is given by $P = (h_1, h_2, h_3, h_4, h_5)$. Doctors underlined in the last column are selected according to the picking order. Doctors in boxes are temporarily accepted by hospitals in each round when hospital rigid quota, region rigid quota, regional elastic quota, and TDE quota are used.

Table 6: Illustration of the DA-D algorithm for Example 4.

Region	r_1		r_2			Picking Order
Hospital	h_1	h_2	h_3	h_4	h_5	$P = (h_1, h_2, h_3, h_4, h_5)$
Round 1	$d_1, \boxed{d_2}$	$d_3, \boxed{d_4}$	$\boxed{d_5}, \boxed{d_6}$			$\succ_B$: $\underline{d_2}, \underline{d_4}, \underline{d_6}, d_1, d_3, \underline{d_5}$
Round 2	$d_2, \boxed{d_3}$	$d_4, \boxed{d_1}$	$\boxed{d_5}, \boxed{d_6}$			$\succ_B$: $\underline{d_3}, \underline{d_1}, \underline{d_6}, d_2, d_4, \underline{d_5}$
Round 3	$d_4, \boxed{d_3}$	$d_2, \boxed{d_1}$	$\boxed{d_5}, \boxed{d_6}$			$\succ_B$: $\underline{d_3}, \underline{d_1}, \underline{d_6}, d_4, d_2, \underline{d_5}$
Round 4	$\boxed{d_3}$	$\boxed{d_1}$	$d_2, \boxed{d_5}, \boxed{d_6}$	$\boxed{d_4}$		$\succ_B$: $\underline{d_3}, \underline{d_1}, \underline{d_6}, \underline{d_5}, d_2$
Round 5	$\boxed{d_3}$	$\boxed{d_1}$	$\boxed{d_5}, \boxed{d_6}$	$d_4, \boxed{d_2}$		
Round 6	$\boxed{d_3}$	$\boxed{d_1}$	$d_5, \boxed{d_4}, \boxed{d_6}$	$\boxed{d_2}$		$\succ_B$: $\underline{d_3}, \underline{d_1}, \underline{d_4}, \underline{d_6}, d_5$
Round 7	$\boxed{d_3}$	$\boxed{d_1}$	$\boxed{d_4}, \boxed{d_6}$	$d_2, \boxed{d_5}$		
Round 8	$\boxed{d_3}$	$\boxed{d_1}$	$\boxed{d_4}, \boxed{d_6}$	$\boxed{d_5}$	$\boxed{d_2}$	$\succ_B$: $\underline{d_3}, \underline{d_1}, \underline{d_4}, \underline{d_6}$

To facilitate a better understanding of the algorithm, we explain several rounds in detail. In Round 1, d_1 and d_2 propose to h_1, d_3 and d_4 propose to h_2, and d_5 and d_6 propose to h_3 in Step 1. We have $M_{h_1} = \{d_1, d_2\}$, $M_{h_2} = \{d_3, d_4\}$, $M_{h_3} = \{d_5, d_6\}$, $\bar{D}_{h_1} = \bar{D}_{h_2} = \bar{D}_{h_3} = \emptyset$ in Step 2. At Step 3, we have $B_{h_1} = \{d_1, d_2\}$, $B_{h_2} = \{d_3, d_4\}$, and $B_{h_3} = \{d_5, d_6\}$, and the picking order $\succ_B$: $d_2, d_4, d_6, d_1, d_3, d_5$. In Process 1, d_2 is temporarily accepted by h_1 and d_4 by h_2. In Process 2, d_5 and d_6 are temporarily accepted by h_3, and then d_1 and d_3 are rejected by h_1 and h_2, respectively.

In Round 2, d_3 proposes to h_1 and d_1 to h_2 in Step 1. We have $M_{h_1} = \{d_2, d_3\}$, $M_{h_2} = \{d_1, d_4\}$, $M_{h_3} = \{d_5, d_6\}$, $\bar{D}_{h_1} = \bar{D}_{h_2} = \bar{D}_{h_3} = \emptyset$ in Step 2. At Step 3, we have $B_{h_1} = \{d_2, d_3\}$, $B_{h_2} = \{d_1, d_4\}$, and $B_{h_3} = \{d_5, d_6\}$, and the picking order $\succ_B$: $d_3, d_1, d_6, d_2, d_4, d_5$. In Process 1, d_3 is temporarily accepted by h_1 and d_1 by h_2. In

Process 2, d_5 and d_6 are temporarily accepted by h_3, and then d_2 and d_4 are rejected by h_1 and h_2, respectively. Round 3 is omitted.

In Round 4, d_2 proposes to h_3 and d_4 to h_4 in Step 1. We have $M_{h_1} = \{d_3\}$, $M_{h_2} = \{d_1\}$, $M_{h_3} = \{d_2, d_5, d_6\}$, $M_{h_4} = \{d_4\}$, $\bar{D}_{h_1} = \bar{D}_{h_2} = \bar{D}_{h_3} = \emptyset$ and $\bar{D}_{h_4} = \{d_4\}$ in Step 2. At Step 3, we have $B_{h_1} = \{d_3\}$, $B_{h_2} = \{d_1\}$, and $B_{h_3} = \{d_2, d_5, d_6\}$, and the picking order $\succ_B$: d_3, d_1, d_6, d_5, d_2. In Process 1, d_3 is temporarily accepted by h_1 and d_1 by h_2. In Process 2, d_5 and d_6 are temporarily accepted by h_3, and then d_2 is rejected by h_3.

In Round 5, d_2 proposes to h_4 in Step 1. h_4 takes d_2 in by rejecting d_4 in Step 2. We have $M_{h_1} = \{d_3\}$, $M_{h_2} = \{d_1\}$, $M_{h_3} = \{d_5, d_6\}$, and $M_{h_4} = \{d_2\}$ in Step 2. Then the algorithm goes back to Step 1. Rounds 6 and 7 are omitted.

In Round 8, d_2 proposes to h_5 in Step 1. We have $M_{h_1} = \{d_3\}$, $M_{h_2} = \{d_1\}$, $M_{h_3} = \{d_4, d_6\}$, $M_{h_4} = \{d_5\}$, $M_{h_5} = \{d_2\}$, $\bar{D}_{h_1} = \bar{D}_{h_2} = \bar{D}_{h_3} = \emptyset$, $\bar{D}_{h_4} = \{d_5\}$, and $\bar{D}_{h_5} = \{d_2\}$ in Step 2. At Step 3, we have $B_{h_1} = \{d_3\}$, $B_{h_2} = \{d_1\}$, and $B_{h_3} = \{d_4, d_6\}$, and the picking order $\succ_B$: d_3, d_1, d_4, d_6. In Process 1, d_3 is temporarily accepted by h_1 and d_1 by h_2. In Process 2, d_4 and d_6 are temporarily accepted by h_3. There is no rejection so the DA-D algorithm stops with the following matching:

$$\mu = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 & h_5 \\ d_3 & d_1 & \{d_4, d_6\} & d_5 & d_2 \end{pmatrix}$$

Observe that this matching μ is efficient and fair but not (pairwise) stable as it can be admissibly blocked by either (d_3, h_2) or (d_1, h_1).

Given a feasible matching μ, let $M_\mu(h) = \{d \in D \mid h \succ_d \mu(d)\}$ be the set of doctors d who prefer hospital h to her assignment $\mu(d)$. For any doctor $d \in D$ and any set H' of hospitals, let $\mathcal{H}_\mu^{H'}(d) = \{h \in H' \mid h \succ_d \mu(d)\}$ be the set of hospitals h that doctor d prefers to her assignment $\mu(d)$, i.e. $h \succ_d \mu(d)$. Let $N_\mu(r) = \{h \in H_r \mid |\mu(h)| < q(h)\}$ be the set of hospitals h located in region r which still have vacant positions under matching μ.

A doctor d is **underdemanded** if she has no position in any hospital, i.e. $\mu(d) = h_0$. A hospital h is **underdemanded** if $M_\mu(h) = \emptyset$ and $|\mu(h)| = l(h)$. A hospital h is underdemanded at matching μ if the number of doctors hired by the hospital at μ equals its floor quota $l(h)$ and no doctor who is not matched with the hospital prefers the hospital to her assignment. A region r is **underdemanded** if $|\mu(r)| = l(r)$ and either $N_\mu(r) = \emptyset$ or $\bigcup_{h\in N_\mu(r)} M(h) = \emptyset$ when $N_\mu(r) \neq \emptyset$. A region r is underdemanded at matching μ if the number of doctors hired by the region r equals its floor quota $l(r)$ and either the number of doctors hired by every hospital $h \in H_r$ in the region equals its ceiling quota $q(h)$ or there exists some hospital in the region which has vacant positions, but no doctor whose assignment is not the hospital prefers the hospital to her assignment. A region r is **nonwasteful** if $|\mu(r)| > l(r)$ and either $N_\mu(r) = \emptyset$ or $\bigcup_{h\in N_\mu(r)} M(h) = \emptyset$ when $N_\mu(r) \neq \emptyset$.

Proposition 2 *Let μ be the matching produced by the DA-D algorithm. There exists either*

(a) a doctor d which has $\mu(d) = h_0$, or

(b) a hospital h which has $M_\mu(h) = \emptyset$ and $|\mu(h)| = l(h)$, or

(c) a region r which has $N_\mu(r) \neq \emptyset$ and $\bigcup_{h\in N(r)} M(h) = \emptyset$, or

(d) a region r which has $N_\mu(r) = \emptyset$.

We have the following Corollary.

Corollary 1 *Let μ be the matching produced by the DA-D algorithm. There exists at least one underdemanded doctor, one underdemanded hospital, one underdemanded region, or one nonwasteful region.*

Based on the DA-D algorithm, we propose the following mechanism for finding a feasible and group stable matching.

Sequential Deferred Acceptance Algorithm with Distributional Constraints (The SDA-D Algorithm)

Step 0. Start with a given market $\mathcal{M}$. Set $\mathcal{M}^1 = \mathcal{M}$ and $k = 1$. Go to Step k.

Step k. Implement the DA-D algorithm for the market $\mathcal{M}^k$ which yields matching μ^k. Remove every underdemanded doctor d from $\mathcal{M}^k$ by setting $\mu^*(d) = h_0$. Do the following:

– For every removed doctor d with $\mu^*(d) = h_0$, remove every remaining hospital $h \in \bigcup_{r \in R^u} \mathcal{H}_{\mu^k}^{H_r}(d)$ from every remaining doctor $\hat{d}$'s preference list, where $\hat{d}$ is given by $d \succ_h \hat{d}$. Let $k = k+1$ and go back to Step k for the remaining market $\mathcal{M}^k$.

If μ^k does not have any underdemanded doctor, remove all underdemanded hospitals h and their hired doctors from $\mathcal{M}^k$ by setting $\mu^*(h) = \mu^k(h)$. Let $k = k+1$ and go to Step k for the remaining market $\mathcal{M}^k$.

If μ^k does not have any underdemanded hospital, remove all underdemanded regions and their hospitals h and their hired doctors from $\mathcal{M}^k$ by setting $\mu^*(h) = \mu^k(h)$. Let $k = k+1$ and go to Step k for the remaining market $\mathcal{M}^k$.

If μ^k does not have any underdemanded region, remove all nonwasteful regions and their hospitals h and their hired doctors from $\mathcal{M}^k$ by setting $\mu^*(h) = \mu^k(h)$. Let $k = k+1$ and go to Step k for the remaining market $\mathcal{M}^k$.

The algorithm stops with the matching μ^* when no doctor is left.

We say that a market is *sparse* if the number of available positions is larger than the number of available doctors in the market, i.e. $n < \sum_{h \in H} q(h)$. Otherwise, the market is *thick*. The SDA-D algorithm is a general mechanism that can deal with both sparse and thick markets. When the market is sparse, a sensible quota system should be given such that every doctor will find a position, i.e. $n < \sum_{r \in R} \min\{q(r), \sum_{h \in r} q(h)\}$. In this case, every doctor can find a position so there will be no underdemanded doctor.

Again we use Example 4 to illustrate the SDA-D algorithm. Note that the market in this example is sparse. We use the same tie breaking order $P = (h_1, h_2, h_3, h_4, h_5)$.

Table 7 shows the process of the SDA-D algorithm for Example 4. In each step of the table, we indicate those underlined doctors who have gotten their final assignments. In Step 1, doctor d_3 is assigned to h_1, d_1 to h_2, both d_4 and d_6 to h_3, d_5 to h_4, and d_2 to h_5. h_5 is underdemanded so removed together with its hired d_2 from the market. In Step 2, doctor d_3 is assigned to h_1, d_1 to h_2, both d_5 and d_6 to h_3, and d_4

to h_4. h_4 is underdemanded so removed together with its hired d_4. In Step 3, doctor d_1 is assigned to h_1, d_3 to h_2, and both d_5 and d_6 to h_3. Both regions r_1 and r_2 are nonwasteful so removed together with the three hospitals and the hired doctors. The algorithm now stops and gives the following matching:

$$\mu^* = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 & h_5 \\ d_1 & d_3 & \{d_5, d_6\} & d_4 & d_2 \end{pmatrix}$$

The outcome (μ^*, σ) is efficient, nonwasteful, group stable, and fair within the same quota type, where

$$\sigma = \begin{pmatrix} d_1 & d_2 & d_3 & d_4 & d_5 & d_6 \\ \tau_{rr_1} & \tau_{rh_5} & \tau_{rr_1} & \tau_{rh_4} & \tau_{er_2} & \tau_{er_2} \end{pmatrix}.$$

Table 7: Illustration of the SDA-D algorithm for Example 4.

Step 1	Region	r_1		r_2			Removed agents
	Hospital	h_1	h_2	h_3	h_4	h_5	underdemanded hospital h_5
	μ^1	d_3	d_1	d_4, d_6	d_5	$\underline{d_2}$	with $\mu^*(h_5) = \{d_2\}$
Step 2	Region	r_1		r_2			Removed agents
	Hospital	h_1	h_2	h_3	h_4		underdemanded hospital h_4
	μ^2	d_3	d_1	d_5, d_6	$\underline{d_4}$		with $\mu^*(h_4) = \{d_4\}$
Step 3	Region	r_1		r_2			Removed agents
	Hospital	h_1	h_2	h_3			nonwasteful regions r_1 and r_2 with
	μ^3	$\underline{d_1}$	$\underline{d_3}$	$\underline{d_5, d_6}$			$\mu^*(h_1) = \{d_1\}$, $\mu^*(h_2) = \{d_3\}$, $\mu^*(h_3) = \{d_5, d_6\}$

We have the following two results for the SDA-D algorithm. Let D^k and H^k stand for the set of doctors and hospitals in the market $\mathcal{M}^k$, respectively.

Proposition 3 *Given any given problem $\mathcal{M}$, let K be the last round in the processes of the SDA-D algorithm. Let μ^k be the matching produced by the algorithm for the remaining market $\mathcal{M}^k$ in round k ($k < K$). Then every doctor $d \in D^k$ has $\mu^{k+1}(d) \succeq_d \mu^k(d)$.*

Theorem 3 *The matching found by the SDA-D algorithm is feasible, group stable and fair within the same quota type.*

Recall that a group stable matching must be efficient, pairwise stable and non-wasteful.

4 Comparison and Discussion

In this section, we discuss compatibility between several solution concepts used in this paper and compare our stability with two closely related stability concepts.

4.1 An Impossibility Result

We first discuss compatibility among several properties including non-wastefulness, fairness, and strategy-proofness. Let $\mathcal{P}$ denote the class of problems under consideration and let φ and ψ be two different mechanisms designed for the problems. We say that a mechanism φ doctor-dominates another mechanism ψ if, for every problem, every doctor weakly prefers the matching produced by φ to that by ψ and if, for some problem, every doctor weakly prefers the matching produced by φ to that by ψ and some doctor strictly prefers the matching produced by φ to that by ψ. Formally, we have

Definition 9 *A mechanism φ **doctor-dominates** ψ if*

(i) for all $P \in \mathcal{P}$, $\varphi_i(P) \succeq_i \psi_i(P)$ for all $i \in D$; and

(ii) for some $P \in \mathcal{P}$, $\varphi_i(P) \succeq_i \psi_i(P)$ for all $i \in D$ and $\varphi_i(P) \succ_i \psi_i(P)$ for some $i \in D$.

Let us reexamine the outcomes generated by the SDA-D algorithm and the DA-D algorithms for Example 4. We can see that doctors prefer the matching produced by the SDA-D algorithm to that by the DA-D algorithm. This is because the SDA-D algorithm is based on the DA-D algorithm and there exists a dominance relationship between the two algorithms. Let φ^{SDA} be the SDA-D algorithm and φ^{DA} the DA-D algorithm. By Proposition 3, we immediately have the following observation.

Proposition 4 *The SDA-D algorithm φ^{SDA} doctor-dominates the DA-D algorithm φ^{DA}.*

Let φ^{NW} represent a mechanism that yields a non-wasteful matching and is strategy-proof for all doctors. Let φ^{DA} denote the DA-D mechanism. Then we have the following general result.

Proposition 5 *Regardless of whether a market faces ceiling constraints or floor constraints or both, there exists no strategy-proof and non-wasteful mechanism φ^{NW} that doctor-dominates the DA-D mechanism φ^{DA}.*

Abdulkadiroğlu et al. (2009) found an impossibility result for a school choice model without distribution constraints saying that there is no strategy-proof mechanism that dominates the deferred acceptance mechanism. Kesten (2010) discovered another impossibility saying that there is no efficient and strategy-proof mechanism selecting a fair and Pareto-efficient matching.

We have the following immediate Corollary.

Corollary 2 *There exists no strategy-proof and group stable mechanism that doctor-dominates the DA-D mechanism φ^{DA}.*

Obviously, by this result and Proposition 4, there is no strategy-proof and group stable mechanism that doctor-dominates the SDA-D mechanism.

4.2 Stability of Kamada and Kojima (2017)

Kamada and Kojima (2017) studied a market with general ceiling constraints by extending Kamada and Kojima (2015) . They propose a feasibility constraint $f : Z_+^{|H|} \to \{0,1\}$ such that $x \leq y$ implies $f(x) \geq f(y)$ and $f(0) = 1$, where 0 means the zero vector and Z_+ is the set of nonnegative integers. Each coordinate x_h in the vector x indicates the number of doctors assigned to hospital $h \in H$. x is feasible if and only if $f(x) = 1$. Otherwise x is infeasible. If x is feasible and $0 \leq y \leq x$, y is also feasible. A matching is **feasible** if and only if $f(w(\mu)) = 1$, where $w(\mu) := (|\mu(h)|)_{h \in H}$.

Their weak stability is defined as follows. A doctor-hospital pair (d, h) block a feasible matching μ if $h \succ_d \mu(d)$ and either $|\mu(h)| < q(h)$ and $\{d\} \succ_h \emptyset$ or $\{d\} \succ_h \{b\}$ for some $b \in \mu(d)$. Given a feasible constraint f, a feasible matching μ is *weakly stable* if it is individually rational, and whenever there is a blocking pair (d, h), we must have (1) $f(w(\mu) + e(h)) = 0$ and (2) $\{b\} \succ_h \{d\}$ for every $b \in \mu(h)$, where $e(h)$ is the h-th unit vector in $R^{|H|}$. Condition (1) says that adding one doctor to hospital h without changing anything else violates feasibility. Condition (2) means that hospital h prefers all its doctors in $\mu(h)$ to doctor d but likes to hire doctor d to fill its vacant position. So, this weak stability does not eliminate certain blocking pairs that still satisfy the feasibility constraint. They also introduced strong stability. They proved the existence of weakly stable matching but disproved the existence of strongly stable matching in their model.

We use the following example to show that their weak stability is not suitable for a market with floor constraints. Clearly, their strong stability cannot be applied either.

Example 5 *There are one doctor $D = \{d\}$ and two hospitals $H = \{h_1, h_2\}$ located in one region r with maximum and minimum quotas $q(h_1) = q(h_2) = 1$, $l(h_1) = 1$, $l(h_2) = 0$, $q(r) = 2$, and $l(r) = 1$. The preferences of doctors and hospitals are given by $\succ_d$: h_2, h_1, $\succ_{h_1}$: d and $\succ_{h_2}$: d.*

It is easy to see that this problem has a unique matching μ with $\mu(h_1) = d$, $\mu(d) = h_1$, and $\mu(h_2) = \emptyset$. By definition, (d, h_2) is a blocking pair of μ. Hypothetically matching the blocking pair (d, h_2) together without removing d from h_1 would satisfy all the constraints but violate the above condition (1). Hence, μ is not weakly stable. We can also look at this issue from a different perspective. Because of the floor constraint on h_1, there exists no feasible matching that can assign d to h_2, as μ is the only feasible matching.

In summary, without floor constraints, removing d from h_1 would not cause any problem with feasibility. However, the presence of floor constraint makes the movement of doctor d from hospital h_1 impossible.

By our definition in Section 2, μ is stable, because (d, h_2) is not an admissible blocking pair.

Note that for any given matching model with unbinding floor constraints, i.e., $l(i) = 0$ for all $i \in H \cup R$, weak stability in Kamada and Kojima (2017) and our concept of pairwise stability do not superset each other, because every weakly stable matching is fair but every pairwise stable matching is non-wasteful.

4.3 Stability of Akin (2021)

We now discuss the notion of floor respecting stability of Akin (2021) for a doctor-hospital matching model with floor constraints. Following Akin (2021), a feasible matching μ is blocked by coalition $A \subseteq D \cup H$ if there is another feasible matching μ' such that (i) $\mu'(d) \in A$ and $\mu'(d) \succeq_d \mu(d)$ for all doctors $d \in A \cap D$; (ii) $\mu'(h) \subseteq A$ and $\mu'(h) \succeq_h \mu(h)$ for all $h \in A \cap H$; (iii) $\mu'(i) \succ_i \mu(i)$ for some doctor or hospital $i \in A$; and (iv) if doctor $d \notin A$, then $\mu'(d) = \mu(d)$ or $\mu'(d) = \emptyset$. A feasible matching μ is *floor respecting stable* if it is not blocked by any coalition. (iv) is a key condition, saying that a coalition can block a feasible matching μ if there exists another feasible matching μ' when in the new matching μ' any doctor not in the blocking coalition can either keep her previous position in μ or has no position at all.

Although Akin (2021) did not discuss blocking pairs, we can immediately use her definition to define a doctor-hospital blocking pair. We say that a feasible matching is floor respecting pairwise stable if it is not blocked by any doctor-hospital pair. We compare these two concepts with our pairwise and group stability concepts.

Let FRS be the set of floor-respecting stable matchings, and $PWFRS$ the set of floor-respecting pairwise stable matchings.

Proposition 6 *For any given matching model with unbinding-disjoint regions, i.e.,* $q(r) \gg \sum_{h \in H_r} q(h)$ *for all* $r \in R$ *and* $H_r \cap H_{r'} = \emptyset$ *for all* $r, r' \in R$*, we have the relationship* $PWS \subseteq PWFRS$*.*

The following example shows that a floor-respecting pairwise stable matching need not be pairwise stable.

Example 6 *There are two doctors $D = \{d_1, d_2\}$ and two hospitals $H = \{h_1, h_2\}$ located in one region r with quotas $q(h_1) = q(h_2) = 1$, $l(h_1) = l(h_2) = 0$, $q(r) = 20$, and $l(r) = 2$. The preferences of doctors and hospitals are given by $\succ_{d_1}$: h_1, h_2 and $\succ_{d_2}$: h_1, h_2, $\succ_{h_1}$: d_2, d_1, and $\succ_{h_2}$: d_2, d_1.*

Table 8: Feasible Matchings in Example 6.

Matching	h_1	h_2
μ_1	d_1	d_2
μ_2	d_2	d_1

There are two feasible matchings μ_1 and μ_2 as shown in the table. Both μ_1 and μ_2 are floor-respecting stable matchings. But only μ_2 is pairwise stable, as it is not admissibly blocked by any doctor-hospital pair and μ_1 is admissibly blocked by the pair (d_2, h_1).

Similarly, we have the following result.

Proposition 7 *For any given matching model with unbinding-disjoint regions, i.e., $q(r) \gg \sum_{h \in H_r} q(h)$ for all $r \in R$ and $H_r \cap H_{r'} = \emptyset$ for all $r, r' \in R$, we have the relationship $GS \subseteq FRS$.*

The above example shows that μ_1 and μ_2 are floor-respecting stable but only μ_2 is group stable.

5 Conclusion

Many real-life markets face a variety of distributional constraints. This paper is motivated by the Chinese doctors, hospitals, and regions matching problem. In the problem, hospitals are located in different regions of the country. Because of the heterogeneity of the development in hospitals and regions, some hospitals or regions are much more popular than others. To ensure certain fairness and a proper level of services at each hospital and each region, floor constraints on the number of doctors are imposed upon every hospital and every region, besides their ceiling constraints.

To deal with this matching market problem with complex distributional constraints, we have introduced two solutions. The first solution is efficient, fair, and strategy-proof, while the second one is group stable and fair within the same quota type. Unlike the conventional two-sided matching problem without distributional constraints where every coalition can be formed as a blocking coalition, in the current market some coalition cannot be formed because it may violate the distributional constraints. We have proposed an intuitive, novel, and practical concept that a blocking coalition is admissible if a new matching resulted by the coalition satisfies all the constraints except some hospital's ceiling constraint but can be made feasible after a grace period of time for dismissed doctors. A feasible matching is (pairwise) group stable if it cannot be blocked by any admissible coalition (any admissible pair). Pairwise stability implies nonwastefulness but is not compatible with fairness. We have weakened fairness to fairness within the same quota type. The latter means that doctors sharing the same quota type do not envy each other and their associated hospitals do not envy each other. Group stability implies efficiency and pairwise stability.

We have proposed new mechanisms for finding the two solutions. They are called the deferred acceptance algorithm with distributional constraints (the DA-D algorithm) and the sequential deferred acceptance algorithm with distributional constraints (the SDA-D algorithm). We have proved that the outcome generated by the DA-D algorithm is fair and efficient, and the algorithm is incentive-compatible for doctors, and the outcome by the SDA-D algorithm is group stable and fair within the same quota type. The DA-D algorithm can be viewed as a significant generalization of the deferred acceptance algorithm and the SDA-D algorithm is built upon the DA-D algorithm. We found that efficiency and nonwasteful do not imply each other. We showed that there exists no strategy-proof and nonwasteful mechanism that doctor-dominates any fair mechanism. We compared our stability concepts with two closely related concepts by Kamada and Kojima (2017); Akin (2021).

In summary, we have introduced a general matching problem with practical and complex distributional constraints by taking doctors, hospitals and regions matching as premier examples. Similar problems arise from civil service sectors and others. We hope this study has given fresh insights into this important class of resource allocation problems.

A Appendix: Proofs

Proof of Proposition 1: In each round when the DA-D processes to Step 3, $|(M_h)_{h\in H}| \geq \sum_{r\in R} l(r) \geq \sum_{h\in H} l(h)$ holds. Otherwise, there does not exist a feasible matching because each hospital is acceptable for each doctor and vice versa. Moeover, at Step 2 of each round of the DA-D procedure, every hospital h tentatively accepts at most $l(h)$ number of doctors by using its rigid quota. So $|\bar{D}_h| \leq l(h)$ holds for each $h \in H$.

Now we consider any round when the DA-D processes to Step 3. Clearly, $|B| = |(M_h \setminus \bar{D}_h)_{h\in H}| \geq |(M_h)_{h\in H}| - \sum_{h\in H} l(h)$ holds true. At Process 1 of Step 3 of the round, hospitals in every region r can tentatively accept in total at most $l(r) - \sum_{h\in H_r} l(h)$ doctors by using their region rigid quota. So $|\bar{B}| \leq \sum_{r\in R}(l(r) - \sum_{h\in H_r} l(h)$ holds and hence $|B \setminus \bar{B}| \geq |(M_h)_{h\in H}| - \sum_{h\in H} l(h) - (\sum_{r\in R}(l(r) - \sum_{h\in H_r} l(h))) \Rightarrow |B \setminus \bar{B}| \geq |(M_h)_{h\in H}| - \sum_{r\in R} l(r)$ holds true. We need to consider the following two cases.

Case 1. $\min\{n, \sum_{r\in R} \min\{q(r), \sum_{h\in H_r} q(h)\}\} = n$. In this case, the total distributable ealstic (TDE) quota is $n - \sum_{r\in R} l(r)$. The total number of hospital rigid quota, region rigid quota and region elastic quota that can be distributed is n. In the DA-D procedure, a doctor will match with h_0 when she has proposed but been rejected by all hospitals; every hospital h will reject a doctor from M_h when all hospital rigid quota, region rigid quota and region elastic quota have been completely used. It implies that $|(M_h)_{h\in H}| = n$ holds true. Hence, $|B \setminus \bar{B}| \geq |(M_h)_{h\in H}| - \sum_{r\in R} l(r) \Rightarrow |B \setminus \bar{B}| \geq n - \sum_{r\in R} l(r) \Rightarrow |B \setminus \bar{B}| \geq \min\{n, \sum_{r\in R} \min\{q(r), \sum_{h\in H_r} q(h)\}\} - \sum_{r\in R} l(r)$ holds true.

Case 2. $\min\{n, \sum_{r\in R}\min\{q(r), \sum_{h\in H_r} q(h)\}\} = \sum_{r\in R}\min\{q(r), \sum_{h\in H_r} q(h)\}$. In this case, the total distributable ealstic (TDE) quota is $\sum_{r\in R}\min\{q(r), \sum_{h\in H_r} q(h)\} - \sum_{r\in R} l(r)$. The total number of hospital rigid quota, region rigid quota and region elastic quota that can be distributed is $\sum_{r\in R}\min\{q(r), \sum_{h\in H_r} q(h)\}$. In the DA-D procedure, a doctor will match with h_0 when she has proposed but been rejected by all hospitals; every doctor in M_h will be rejected by h when all hospital rigid quota, region rigid quota and region elastic quota have been completely used. It implies that $|(M_h)_{h\in H}| \geq \sum_{r\in R}\min\{q(r), \sum_{h\in H_r} q(h)\}$ holds true. Hence, $|B \setminus \bar{B}| \geq |(M_h)_{h\in H}| - \sum_{r\in R} l(r) \geq \sum_{r\in R}\min\{q(r), \sum_{h\in H_r} q(h)\} - \sum_{r\in R} l(r)$ holds.

□

To prove Theorems 1, 2, and 3, we need to introduce several new definitions and lemmas. Let μ be the matching produced by the DA-D algorithm and K be the last round. We first show that as long as hospital h rejects some doctor at Step 3 in some round k, the number of region rigid quota and region elastic quota distributed to h in any later round will not exceed the number of the quotas obtained by h in round k. Let $\bar{D}^k(h)$ and $\hat{D}^k(h)$ be the set of doctors who are temporarily accepted at Step 2 and at Step 3 in round k, respectively.

Lemma 5 *If hospital h rejects some doctor at Step 3 of the DA-D algorithm in round k, then for all $k' \geq k$, we have $|\bar{D}^{k'}(h)| = |\bar{D}^k(h)| = l(h)$ and $|\hat{D}^{k'}(h)| \leq |\hat{D}^k(h)|$.*

Proof: Suppose hospital h rejects some doctor at Step 3 in round k.

We first show that $|\bar{D}^{k'}(h)| = |\bar{D}^k(h)| = l(h)$ for all $k' \geq k$. Hospital h rejects some doctor at Step 3 in round k. It indicates that hospital h has completely used its rigid quota to recruit doctors in round k, i.e. $|\bar{D}^k(h)| = l(h)$. In the DA-D procedure, a hospital rejects a doctor who has been tentatively accepted by using the hospital's rigid quota only when the hospital receives a more preferred doctor's proposal. So $|\bar{D}^{k'}(h)| = l(h)$ for all $k' \geq k$ holds true.

We further prove the case of $|\hat{D}^{k'}(h)| \leq |\hat{D}^k(h)|$ for all $k' \geq k$ by induction. When $k' = k$, it is obvious that $|\hat{D}^k(h)| \leq |\hat{D}^k(h)|$ holds. We will show that if $|\hat{D}^m(h)| \leq$

$|\hat{D}^k(h)|$ holds true for all rounds m where $k \leq m < K$, $|\hat{D}^{m+1}(h)| \leq |\hat{D}^k(h)|$ will hold true in round $m+1$. Suppose the relation does not hold such that (i) $|\hat{D}^{m+1}(h)| > |\hat{D}^m(h)|$ in round $m+1$. Hospital h rejecting some doctor at Step 3 in round k of the DA-D procedure indicates that all region rigid quotas $\tau_{rr(h)}$ and the number

$$\min\{n - \sum_{r\in R} l(r), \sum_{r\in R} \min\{\kappa(\tau_{er}), \sum_{h\in H_r} q(h) - l(r)\}\}$$

of region elastic quotas has been distributed. Therefore, h obtaining additional region rigid or elastic quotas means that either (1) hospital h' in region $r(h)$ loses at least one region rigid or elastic quota in round $m+1$ or (2) hospital h' in other region $r \neq r(h)$ loses at least one region elastic quota in round $m+1$. So the hospital h' has (ii) $|\hat{D}^m(h')| > |\hat{D}^{m+1}(h')|$.

In the DA-D procedure, region rigid quotas and region elastic quotas are distributed one by one to hospitals to hire doctors according to the picking order given the tie-breaking order $\succ^t$. According to the rule of the picking order, we have (iii) $|\hat{D}^m(h')| \leq |\hat{D}^k(h)| + 1$ because h rejects some doctor at Step 3 in round k; and (iv) $|\hat{D}^{m+1}(h)| \leq |\hat{D}^{m+1}(h')| + 1$ because h' rejects some doctor at Step 3 in round $m+1$. According to relations (i),(ii),(iii) and (iv), we have the follwing relation

$$|\hat{D}^{m+1}(h)| \leq |\hat{D}^{m+1}(h')| + 1 \leq |\hat{D}^m(h')| \leq |\hat{D}^k(h)| + 1 \leq |\hat{D}^{m+1}(h)|,$$

so that $|\hat{D}^{m+1}(h)| = |\hat{D}^{m+1}(h')| + 1 = |\hat{D}^m(h')| = |\hat{D}^k(h)| + 1$ holds true. According to the rule of the picking order, we have $h' \succ^t h$ in round m while $h \succ^t h'$ in round $m+1$, which yields a contradiction.

□

Proof of Theorem 1: Let μ be the matching produced by the DA-D algorithm and K be the last round in the procedure.

(i) Feasiblity: First, matching μ does not violate ceilling constraints on hospitals and regions. Step 2 of the DA-D algorithm shows that every hospital only accepts a set of doctors up to its capacity. Step 3 of the algorithm shows that the number of doctors accepted by hospitals in any region will not exceed the ceiling of the region in every round.

Second, matching μ does not violate floor constraints on hospitals and regions. Suppose that some floor constraint is violated in matching μ. There are two cases.

Case (i). There exists some doctor d with $\mu(d) = h_0$. It means that doctor d has been rejected by all hospitals. In the algorithm, no hospital rejects any proposal from acceptable doctors as long as it has undistributed hospital rigid quotas and its region rigid quota. Hence, all hospital rigid quotas and region rigid quotas are filled in this case, which contradicts the assumption.

Case (ii). There exists no doctor d with $\mu(d) = h_0$ so that $|\mu| = n$. It indicates that $n \le \sum_{r\in R} \min\{q(r), \sum_{h\in H_r} q(h)\}$. According to the quota system, the TDE quota is

$$\min\{n - \sum_{r\in R} l(r), \sum_{r\in R} \min\{q(r) - l(r), \sum_{h\in H_r} q(h) - l(r)\}\} = n - \sum_{r\in R} l(r).$$

So the total number of distributable quota in the market is n. Every doctor is matched with a hospital associated a quota from either hospital rigid quotas, region rigid quotas or region elastic quotas. If there exists some hospital h with $|\mu(h)| < l(h)$ or some region r with $|\mu(r)| < l(r)$, then at least one distributable quota is undistributed, i.e. $|\mu| < n$, which contradicts the assumption.

(ii) Fairness: In each round of the algorithm, every hospital selects its most preferred set of doctors according to the currently available proposals. In each round, $d \succ_h d'$ implies $(d, h) \succ_B (d', h)$ for all $h \in H$. By Lemma 5, if $h \succ_d \mu(d)$, then $d' \succ_h d$ for all $d' \in \mu(h)$.

(iii) Efficiency: Suppose that μ is not efficient. There exists another feasible matching ν such that $\nu(i) \succeq_i \mu(i)$ for all $i \in D \cup H$ and $\nu(i) \succ_i \mu(i)$ for some $i \in D \cup H$. Under strict preference, there exists some hospital $h \in H$ such that $\nu(h) \succ_h \mu(h)$. Otherwise, matching μ is not individually rational. When $\nu(h) \succ_h \mu(h)$, we need to consider the following cases:

Case 1. Hospital h has $|\nu(h)| \le |\mu(h)|$. Due to the responsiveness of the preference of h, there exists some doctor $d \in \nu(h) \setminus \mu(h)$ such that $\{d\} \succ_h \{d'\}$ for some $d' \in \mu(h)$. Under strict preference, we have $h \succ_d \mu(d)$, which contradicts fairness.

Case 2. Hospital h has $|\nu(h)| > |\mu(h)|$. Based on the algorithm, we have $|\mu| = \min\{n, \sum_{r\in R} \min\{\kappa(\tau_{er}), \sum_{h\in H_r} q(h)\}\}$ and hence $|\mu| = |\nu|$. There exists some hospital

h' with $|\nu(h')| < |\mu(h')|$. Due to the responsiveness of the preference of h, there exists some doctor $d \in \nu(h') \setminus \mu(h')$ such that $\{d\} \succ_{h'} d'$ for some $d' \in \mu(h')$. Under strict preference, we have $h' \succ_d \mu(d)$, contradicting fairness. □

The following proof of Theorem 2 generalizes the well-known argument given by Dubins and Freedman (1981) for the Gale-Shapley marriage matching model to our current more sophisticated model in which both hospitals and regions face their own floor and ceiling constraints on the number of doctors.

Proof of Theorem 2: Let $\succ_D$ be the set of true preference profiles for doctors and $\mu = \varphi(\succ_D)$ be the matching produced by the DA-D algorithm under $\succ_D$.

Consider any preference profile $\succ_D^{'}$. Let $\nu = \varphi(\succ_D^{'})$ be the matching produced by the DA-D algorithm under $\succ_D'$. We show that we only need to consider a specific kind of misrepresentation. Consider a manipulator d_m who makes a simple equivalent misrepresentation $\succ_D^{''} = (\succ_{d_m}^{'}, \succ_{-d_m}^{'})$ such that the manipulator d_m simply puts his match in $\nu(d_m)$ to the top of his list. That is, d_m prefers $\nu(d_m)$ to all other hospitals. We have the following Lemma 6.

Lemma 6 *Let $\nu = \varphi(\succ_D^{'})$ and $\nu^{'} = \varphi(\succ_D^{''})$. Then, the set $M = \{d \in D \mid \nu(d) \succ_d^{''} \nu^{'}(d)\}$ is empty.*

Proof: By Theorem 1, matching ν and $\nu^{'}$ are fair under preference $\succ_D^{'}$ and $\succ_D^{''}$, respectively. Since the only change from $\succ_D^{'}$ to $\succ_D^{''}$ is d_m who moves $\nu(d_m)$ to the top of his preference list, ν is also fair under preference $\succ_D^{''}$. Let $N = (\nu(d))_{d \in M}$. If the manipulator d_m receives different assignments at ν' and ν, i.e. $\nu^{'}(d_m) \neq \nu(d_m)$, then $\nu(d_m) \succ_{d_m}^{''} \nu^{'}(d_m)$, and hence $d_m \in M$.

Suppose $M \neq \emptyset$. We have the following claim .

<u>Claim (1)</u>. Either $\nu(d) = h$ or $\nu(d) \succ_d^{'} \nu'(d) = h$ for every doctor $d \in \nu^{'}(h)$ with $h \in N$. Proof: Supoose there exists $\nu'(d) = h \succ_d^{'} \nu(d)$ for some $d \in \nu^{'}(h)$ with $h \in N$. It indicates that the dotor is not the manipulator, i.e. $d \neq d_m$ and hence $\succ_d^{'} = \succ_d^{''}$. Since matching ν is fair under $\succ'$ and $h \succ_d^{'} \nu(d)$, then $d' \succ_h^{'} d$ ($d' \succ_h^{''} d$) for each doctor

$d' \in \nu(h)$. Since $h \in N$, there exists some doctor $d' \in M$ who has $\nu(d') = h \succ''_{d'} \nu'(d)$ so that (d', h) forms an envious pair, which contradicts fairness of matching ν'. □

Let k be the earliest round in which some doctor $d \in M$ is rejected by $\nu(d)$ in the procedure $\phi(\succ''_D)$. By Claim (1), every doctor d who has been assigned to hospital $\nu(d)$ at ν' is assigned to either $\nu(d)$ or a better choice at ν. No doctor d' who has $\nu(d') \succ''_{d'} \nu(d)$ applies to $\nu(d)$ in round k. (Otherwise, we have $d' \in M$, and hence k is not the earliest round.) Therefore, each doctor who makes a proposal to ν in round k of the procedure $\phi(\succ''_D)$ is assigned to either $\nu(d)$ or a worse choice at ν. According to the fairness of the DA-D procedure, $d \succ'_{\nu(d)} d'$ holds true for each doctor d' who is assigned a worse choice than $\nu(d)$ at ν, i.e. $\nu(d) \succ'_{d'} \nu(d')$. So, no doctor d' who has $\nu(d) \succ'_{d'} \nu(d')$ is tentatively accepted by $\nu(d)$ in round k of the procedure $\phi(\succ''_D)$ because of the rejection of d. This implies that the number of doctors tentatively accepted by $\nu(d)$ in round k of the procedure $\phi(\succ''_D)$ is smaller than $|\nu(\nu(d))|$. Hence, the rejection of doctor d is not because of $\nu(d)$'s capacity, which means that doctor d is rejected by $\nu(d)$ in Step 3 of round k; and by Lemma 5, $|\nu^{'}(\nu(d))| < |\nu(\nu(d))|$ hold.s true, which means that $\nu(d)$ loses at least one region rigid quota or region elastic quota (in round k) at ν' than that at ν.

Because doctor d has been rejected in Step 3 of round k of the procedure $\phi(\succ''_D)$, all region rigid quota of region $r(\nu(d))$ and all distributable region elastic quota have been distributed in round k. So there must exist some hospital h' (1) (tentatively) obtaining the "losing" quota so that the number of doctors tentatively accepted by h' is more than $|\nu(h')|$ in round k; and (2) assigning to some doctor d' who has a higher priority (of obtaining the "losing" quota) than d at the picking order $\succ_B$ in round k of the procedure $\phi(\succ''_D)$. By Lemma 5, h' has not rejected any doctor in step 3 of procedure $\varphi(\succ'')$.

Given the set of proposals from doctors, the picking order constructed in each round of the DA-D algorithm depends on the number of doctors tentatively accepted by each hospital and the picking ordering among hospitals. Therefore, hospital h' is underdemanded at ν such that no doctor $d'' \notin \nu(h')$ prefers h' to $\nu(d'')$. Otherwise, doctor d'' who has $\nu(d'') \succ'_{d''} h'$ (so that $d'' \neq d_m$ and $\succ'_{d''}=\succ''_{d''}$) has higher priority (of obtaining the "losing" quota) than d, which causes a contradiction to the fact that $\nu(d) = h$. Therefore, doctor d' has $\nu(d') \succ''_{d'} h'$. That is, $d' \in M \setminus \{d\}$ and d' is rejected by $\nu(d')$ before round k, which contradicts the assumption.

Consequently, $M = \emptyset$. Therefore, we have $\nu(d_m) = \nu'(d_m)$.

□

According to Lemma 6, compared to $\succ'_D$, no doctor becomes worse off under the simple misrepresentation $\succ''_D$. Since $\nu(d_m)$ is the top choice of d_m under $\succ''_D$, we have $\nu(d_m) = \nu'(d_m)$. Therefore, we only need to consider the simple misrepresentation of the manipulator d_m.

Let $\mu = \varphi(\succ_D)$ be the outcome produced by the DA-D algorithm under true preference profile and $\nu = \varphi(\succ'_D)$ be the outcome produced by the algorithm under preference profile, where doctor d_m makes a simple misrepresentation $\succ'_{d_m}$ and other doctors present their true preference. That is $\succ'_D = (\succ'_{d_m}, \succ_{-d_m})$. It is clear that matching μ is fair under the simple misrepresentation $\succ'_D$. We prove that d_m's simple misrepresentation cannot be successful. By Lemma 6, we have the following Corollary.

Corollary 3 *If $\nu(d_m) \succ_{d_m} \mu(d_m)$ or $\nu(d_m) = \mu(d_m)$, then $\nu(d) \succ_d \mu(d)$ or $\nu(d) = \mu(d)$ for every $d \in D$.*

We continue the proof of Theorem 2. Suppose the manipulator d_m has $\nu(d_m) \succ_{d_m} \mu(d_m)$ or $\nu(d_m) = \mu(d_m)$. Let k be the round when doctor d_m proposes to $\mu(d_m)$ in the process of $\varphi(\succ_D)$. We will show that every doctor $d \in D$ who proposes to $\mu(d)$ in round k or later has $\nu(d) = \mu(d)$. There are three possibilities:

<u>Case 1</u>. If $\mu(d) = \emptyset$, then $\nu(d) = \emptyset$. Otherwise, there exists a doctor d with $\mu(d) = \emptyset$ having $\nu(d) \in H$. According to the DA-D procedure, the total number of quota that can be distributed to hospitals in the process of $\phi(\succ_D)$ is the same as that in the process of $\phi(\succ'_D)$. So there exists some doctor $d' \in D \setminus \{d_m, d\}$ with $\mu(d') \in H$ having

$\nu(d') = \emptyset$, i.e. $\mu(d') \succ'_{d'} \nu(d')$, which contradicts Lemma 6 that $\{d \in D | \mu(d) \succ'_d \nu(d)\} = \emptyset$. Therefore, if $\mu(d_m) = \emptyset$, then $\nu(d_m) = \emptyset$ holds.

Case 2. If d_m proposes to $\mu(d_m) \in H$ in the last round K of the process $\varphi(\succ_D)$, then every doctor d including d_m, who proposes $\mu(d) \in H$ in the last round K of the process $\varphi(\succ_D)$, has $\nu(d) = \mu(d)$.

Let M^K be the set of those doctors who have proposed in the last round K. Suppose that there exists a doctor $d \in M^K$ such that $\nu(d) \neq \mu(d)$. By Corollary 3, we have $\nu(d) \succ_d \mu(d)$; By Lemma 5, $\mu(d)$ rejects no doctor in the process of $\varphi(\succ_D)$. (Otherwise, round K is not the last round.) So $|\nu(\mu(d))| < |\mu(\mu(d))|$ holds true by Corollary 3. This implies that (i) at matching μ, $\mu(d)$ consumes at least one region rigid quota or region elastic quota, and at least one region rigid quota or region elastic quota is available in last round K; (ii) at matching ν, compared to the quotas obtained by $\mu(d)$ at μ, $\mu(d)$ loses at least one region rigid quota or region elastic quota. From (i), any rejection by hospitals which can potentially obtain the "losing"quota is caused by capacity in the process of $\varphi(\succ_D)$ so that those hospitals have no capacity to accept the additional "losing"quota. From (ii), there exists some hospital h which obtains the "losing"quota at ν but rejects no doctor in the process $\varphi(\succ_D)$. By Corollary 3, doctor $d' \in \nu(h) \setminus \mu(h)$ has $h \succ_{d'} \mu(d')$, contradicting the fact that h rejects no doctor.

Case 3. If d_m proposes $\mu(d_m)$ in round $k < K$ of the procedure $\phi(\succ_D)$, then every doctor d including d_m, who proposes $\mu(d)$ in round k' where $k \leq k' < K$, has $\nu(d) = \mu(d)$. We prove this by induction. Let M^i be the set of doctors who have proposed in round i of matching μ. From Case 2, we have shown that $\nu(d) = \mu(d)$ for every doctor $d \in M^K$. The inductive part is that if $\nu(d) = \mu(d)$ for every $d \in M^{k'+1}$, then $\nu(d) = \mu(d)$ for every $d \in M^{k'}$. Suppose that there exists some doctor $d \in M^{k'}$ having $\nu(d) \neq \mu(d)$. By Corollary 3, we have $\nu(d) \succ_d \mu(d)$. Let $M'_h = \{d' \in D | h \succ_{d'} \mu(d')\}$. There are the following cases:

Case 3.a If $|\nu(\mu(d))| < |\mu(\mu(d))|$. This implies that (i) at matching μ, $\mu(d)$ is distribued at least one region rigid quota or region elastic quota; (ii) at matching ν, compared to the quotas obtained by $\mu(d)$ at μ, $\mu(d)$ loses at least one region rigid quota or region elastic quota. According to the DA-D algorithm, there exists some hospital h' with $|\mu(h')| < |\nu(h')|$ which obtains the "losing" quota.

We first show that all region rigid quota of $r(h')$ and distributable region elastic quota have been distributed before round k' of the process $\varphi(\succ_D)$. Let $d^* \in M'_{h'}$ be the most preferred doctor by h' among doctors from $M'_{h'}$. Then, we show that d^* is rejected by h' at Step 3 of some round $s < k'$ of the process $\varphi(\succ_D)$. Since $|\mu(h')| < |\nu(h')|$, h' rejects d^* at Step 3. Suppose that d^* is rejected by h' in round k' or later. By the inductive assumption, $\nu(d^*) = \mu(d^*)$. According to Corollary 3 and $|\mu(h')| < |\nu(h')|$, there must exist some doctor $d^{'}$ who is rejected by h' before round k' in the process of $\varphi(\succ_D)$ and has $\nu(d^{'}) = h'$. Then, $h' \succ_{d^*} \nu(d^*)$ and $d^* \succ_{h'} d'$ holds true, which contradicts fairness of matching ν.

Recall that $\mu(d)$ obtains at least one region rigid or region elastic quota in the procedure $\varphi(\succ_D)$ but loses one such quota in the procedure $\varphi(\succ'_D)$. $\mu(d)$ accepting d in round k' of $\varphi(\succ_D)$ implies that there must exist some hospital $h^{''}$ losing the "losing" quota and rejecting some doctor $d^{''} \neq d_m$ in or after round k'. By inductive assumption, we have $\nu(d'') = \mu(d'')$. Since h' rejects d^* at Step 3 before round k' and h'' loses the "losing" quota until round k' or after round k', d'' should have priority over d^* according to the picking order. In the process of $\varphi(\succ'_D)$, h' obtains the "losing" quota while h'' does not ($d'' \notin \nu(h'')$), which contradicts the fact that d'' has priority over d^*.

Case 3.b If $|\nu(\mu(d))| \geq |\mu(\mu(d))|$. By Lemma 6, $M'_{\mu(d)} \neq \emptyset$ holds true. Let d^* be the doctor who is most preferred by $\mu(d)$ in $M'_{\mu(d)}$. Recall that d makes his match $\mu(d)$ in round k' of the procedure $\varphi(\succ_D)$. This means that d^* is rejected by $\mu(d)$ in or later than round k' so that d^* proposes his match $\mu(d^*)$ at μ after round k'. Hence, $d^* \neq d_m$ and $\nu(d^*) = \mu(d^*)$ holds by the inductive assumption. In the procedure $\varphi(\succ'_D)$, there exists some doctor $d^{'} \in M'_{\mu(d)} \setminus \{d^*\}$ with $\nu(d^{'}) = \mu(d)$. So $d^* \succ_{\mu(d)} d'$ and $\mu(d) \succ_{d^*} \nu(d^*)$ hold true, which contradicts fairness of ν.

We have now proved that every doctor $d \in D$, who proposes to $\mu(d)$ in round k or later, has $\nu(d) = \mu(d)$. So d_m's mispresentation cannot be successful. □

Proof of Proposition 2: Let μ be the matching produced by the DA-D algorithm and

K be the last round. Suppose that at matching μ, there does not exist (a) a doctor d who has $\mu(d) = h_0$; or (b) a hospital h which has $M_\mu(h) \neq \emptyset$ and $|\mu(h)| = l(h)$; or (c) a region r which has $N_\mu(r) \neq \emptyset$ and $\bigcup_{h \in N_\mu(r)} M_\mu(h) = \emptyset$; and (d) a region r which has $N_\mu(r) \neq \emptyset$. In other words, at matching μ, the following cases hold

Case 1. Every doctor d has $\mu(d) \neq h_0$.

Case 2. Every hospital h has either $M_\mu(h) \neq \emptyset$ or $|\mu(h)| > l(h)$.

Case 3. Every region r has $N_\mu(r) \neq \emptyset$ and $\bigcup_{h \in N_\mu(r)} M_\mu(h) \neq \emptyset$.

Let Z^K be the set of doctors who propose to their match at μ in round K. Consider any doctor $d \in Z^K$. According to Case 1, d has $\mu(d) \in H$. According to Lemma 5 and Case 2, $\mu(d)$ has $M_\mu(\mu(d)) = \emptyset$ and $|\mu(\mu(d))| > l(\mu(d))$. So there remains (at least) either one undistributed region rigid quota, or one undistributed region elastic quota such that distributable elastic quota are not fully distributed in the last round K. Otherwise, K is not the last round. According to Case 3, there exists some hospital $h' \neq \mu(d)$ in region $r(\mu(d))$ with $|\mu(h')| < q(h')$ who rejects some doctor at Step 3 of some round before round K. According to the DA-D algorithm, the rejection indicates that one of the following situations holds: (1) all region rigid quota of $r(\mu(d))$ have been distributed before round K; (2) all distributable elastic quota have been distributed before round K. Either situation yields a contradiction. □

Before we prove the Theorem 3, we first prove Propositions 3 and 8.

Proof of Proposition 3: Let $\mathcal{M}^k$ be the remaining market in round k ($k < K$), and μ^k be the matching produced in round k. According to the SDA-D algorithm, $D^{k+1} \subseteq D^k$ holds true for every round $1 \leq k < K$; and every doctor who is removed in round k, i.e. $d \in D^k \setminus D^{k+1}$ has $\mu^k(d) = \mu^{k+1}(d)$.

Now, we show that every doctor $d \in D^{k+1}$ has $\mu^{k+1}(d) \succeq_d \mu^k(d)$. Let $Z^{k+1} \subseteq D^{k+1}$ be the set of doctors who have $\mu^k(d) \succ_d \mu^{k+1}(d)$. Suppose to the contrary that $Z^{k+1} \neq \emptyset$. Let ν be a matching that $\nu(d) = \mu^k(d)$ for all $d \in D^{k+1}$. The following two claims hold true.

Claim (1). For every doctor $d \in Z^{k+1}$, $\mu^k(d) \neq h_0$. According to the SDA-D algorithm, if $\mu^k(d) = h_0$, then $\mu^{k+1}(d) = h_0$ because doctor d is underdemanded and removed in round k.

Claim (2). Matching ν is feasible for the market $\mathcal{M}^{k+1}$, and if doctor $d \in D^{k+1}$ prefers a hospital $h \in H^{k+1}$ to $\nu(d)$, then $d' \succ_h d$ for all $d' \in \nu(h)$. Since μ^k is feasible for $\mathcal{M}^k$, clearly ν is feasible for $\mathcal{M}^{k+1}$. By the DA-D algorithm, μ^k is a fair matching in $\mathcal{M}^k$. So if there exists some doctor d who prefers $h \in H^k$ to $\mu^k(d)$, then $d' \succ_h d$ holds for all $d' \in \mu^k(h)$ Since $H^{k+1} \subseteq H^k$ and $D^{k+1} \subseteq D^k$, Claim (2) holds true.

Let doctor $d \in Z^{k+1}$ be the first doctor among doctors from Z^{k+1} rejected by their match at μ^k in the SDA-D procedure for the market $\mathcal{M}^{k+1}$. Let s be the step of round $k+1$ of the SDA-D procedure for $\mathcal{M}^{k+1}$ in which d is rejected by $h = \mu^k(d)$. According to Claim (2), each doctor $d' \in \mu^{k+1}(h) \setminus \mu^k(h)$ has $\nu(d') \succ_{d'} h$ and hence $d' \in Z^{k+1}$. Since d is the first doctor rejected by his match at μ^k in some step s of round $k+1$ of the SDA-D procedure, every other doctor from Z^{k+1} has not been rejected by their match at μ^k in step s. Since μ^{k+1} is fair in $\mathcal{M}^{k+1}$ (Theorem 1), any doctor who is less preferred than d by h is also rejected by h in or earlier than step s. So h rejects d in step s of round $k+1$ because all region rigid quotas and all distributable elastic quota have been distributed; and the number of doctors who are tentatively accepted by h at step s is smaller than $|\mu^k(h)|$. According to Lemma 5, $|\mu^k(h)| > |\mu^{k+1}(h)| \geq l(h)$ holds.

So compared to $\mu^k(h)$ produced in round k of the SDA-D procedure for $\mathcal{M}^k$, h loses at least one quota from region rigid quotas or region elastic quotas in step s of round $k+1$. According to the SDA-D algorithm, all region rigid quotas and region elastic quotas obtained by each $h \in H^{k+1}$ will remain in round $k+1$. There exists some hospital $h' \in H^{k+1} \setminus \{h\}$ (temporarily) obtaining the "losing" quota in step s of round $k+1$. So that the number of doctors tentatively accepted by h' in step s is larger than $|\mu^k(h')|$. Since d is the "first" doctor, some doctor $d' \in H^{k+1}$ who has $h' \succ_{d'} \mu^k(d')$ is tentatively accepted by h' in step s. This implies that d' has a higher priority (of obtaining the "losing" quota) than d in round $k+1$. In the round k of the SDA-D procedure for $\mathcal{M}^k$, h obtains the "losing" quota while h' does not. So d must have a higher priority (of obtaining the "losing" quota) than d' in round k, which yields a contradiction, because all markets $\mathcal{M}^k$ use the same picking order rule and the tie-breaking order of hospitals.

□

We now introduce the following result.

Proposition 8 *Let μ be the matching produced by the DA-D algorithm for problem P. If there does not exist any underdemanded doctor, underdemanded hospital, or underdemanded region, then any hospital h which makes a rejection has either $|\mu(h)| = q(h)$ or $|\mu(r(h))| = q(r(h))$.*

Proof: Let μ be the matching produced by the DA-D algorithm and K be the last round. According to Proposition 2, there must exist at least one nonwasteful region r with $|\mu(r)| > l(r)$ such that either (i) $N_\mu(r) \neq \emptyset$ and $\bigcup_{h \in N_\mu(r)} M_\mu(h) = \emptyset$, or (ii) $N_\mu(r) = \emptyset$.

In case (ii), we have $|\nu(h)| = q(h)$ for all $h \in r(h)$ so the Proposition holds immediately. Now suppose that the Proposition is false. Then there must exist a region with $|\mu(r)| > l(r)$ such that $N_\mu(r) \neq \emptyset$ and $\bigcup_{h \in N_\mu(r)} M_\mu(h) \neq \emptyset$. Then there exists some hospital $h \in r$ which makes a rejection, having $|\mu(h)| < q(h)$. So h must have rejected some doctor at Step 3 in some round $k < K$. This means that all region rigid quotas of region r and the distributable elastic quota have been used in round k. According to the DA-D algorithm, we have the following fact that (i) any hospital in region r which is assigned an extra quota from region rigid quotas and region elastic quotas of r in any round later than k will cause a rejection; and (ii) any hospital in region $r' \neq r$ which is assigned an extra quota from region elastic quota of r' in any round later than k will cause a rejection. Therefore, in the last round K, there exists at least one underdemanded doctor or one underdemanded hospital or an underdemanded region, which contradicts the assumption. □

Proof of Theorem 3: Let μ be the matching produced by the SDA-D algorithm and K be the last round. Let μ^k be the matching produced in the round k for subproblem P^k. Let $k(i)$ denote the round when doctor or hospital i finalizes their match $\mu(i)$.

Feasibility: Since the DA-D algorithm is feasible, the SDA-D algorithm produces a feasible allocation μ^k in each round k for problem P^k and therefore matching μ is feasible.

Group stability: Suppose that matching μ is not group stable. There exists a coalition $S \subseteq D \cup H$ which can admissibly block μ. That is, there exists a redistribution ν^S among S such that

(i) $\nu^S(d) \in S$ for every $d \in S \cap D$;

(ii) for every $h \in S \cap H$, $d \in \nu^S(h)$ implies $d \in S \cup \mu(h)$;

(iii) $\nu^S(i) \succeq_i \mu(i)$ for all $i \in S$ and $\nu^S(i) \succ_i \mu(i)$ for some $i \in S$;

(iv) matching ν given by $\nu(i) = \nu^S(i)$ for all $i \in S \cap D$ and $\nu(i) = \mu(i)$ for all $i \in D \setminus S$ is permissible at μ.

First, since any feasible matching is individually rational, coalition S will not be a single doctor.

Second, we show that any underdemanded doctor will not make any improvement in the coalition S. According to the SDA-D algorithm, any underdemanded doctor will be removed in the first round. If there exist underdemanded doctors, then the TDE quota is $\sum_{r \in R} \min\{q(r) - l(r), \sum_{h \in H_r} q(h) - l(r)\}$ and all distributable elastic quotas are distributed. So (i) the ceiling constraint of each region $r \in R^b$ is filled; and (ii) the capacity of each hospital in each region $r \in R^u$ is filled. Suppose that there exists some underdemanded doctor d who makes improvement in the coalition S, i.e. $\nu(d) \in H$. By definition, matching ν given that $\nu(i) = \nu^S(i)$ for all $i \in S \cap D$ and $\nu(i) = \mu(i)$ for all $i \in D \setminus S$ is permissible at μ. By definition, every doctor $d' \in S$ has $\nu(d') \succeq_{d'} \mu(d')$ so that each doctor $d' \in S$ with $\mu(d') \in H$ has $\nu(d') \in H$, and every doctor $d' \in D \setminus S$ has $\nu(d') = \mu(d')$. So $|\nu| > |\mu|$. According to the DA-D algorithm, all hospital rigid quotas, region elastic quotas and distributable elastic quotas are distributed at μ. It indicates that matching ν is infeasible. Therefore, a coalition S does not necessarily include any underdemanded doctor.

Let $X = \{d \in S | \nu(d) \succ_d \mu(d)\}$ be the set of doctors in the coalition S who make improvements under the redistribution ν. Suppose doctor $d \in X$ is the earliest doctor who finalizes his match $\mu(d)$ among doctors from X. Let $H(k)$ be the set of hospitals removed in round k. We will prove that any hospital $h \in \bigcup_{k=1}^{k(d)} H(k)$ will not make improvement at μ so that d will not be matched to any hospital $h \in \bigcup_{k=1}^{k(d)} H(k)$ at ν under strict prferences of hospitals. Since d is the "earliest" doctor, $k(d) \leq k(i)$ holds for every doctor $i \in X$, and hence $i \in D^{k(h)}$ holds for every $h \in \bigcup_{k=1}^{k(d)} H(k)$ and every doctor $i \in X$. According to Proposition 2, each hospital $h \in \bigcup_{k=1}^{k(d)} H(k)$ has either $M_{\mu^{k(h)}}(h) = \emptyset$ or $|\mu(h)| = q(h)$ in round $k(h)$. We will prove it by showing the following two claims hold.

Claim (1) No hospital $h \in \bigcup_{k=1}^{k(d)} H(k)$ with $M_{\mu^{k(h)}}(h) = \emptyset$ will make improvement at ν. According to Proposition 3, every doctor $i \in S$ prefers her match $\mu(i)$ to h because $i \in D^{k(h)}$ holds.

Claim (2) No hospital $h \in \bigcup_{k=1}^{k(d)} H(k)$ with $|\mu(h)| = q(h)$ will make improvement at ν. Suppose h is in the coalition S such that $\nu(h) \succ_h \mu(h)$. There must exist some doctor $d' \in X$ who has $\nu(d') = h \succ_{d'} \mu(d')$ and $d' \succ_h d''$ for some $d'' \in \mu(h)$ because of the responsive preferences of hospitals. Since $\mu^{k(h)}$ is fair for problem $P^{k(h)}$ (Theorem 1) and $d' \in D^{k(h)}$, $d'' \succ_h d'$ holds for all $d'' \in \mu(h)$, which yields a contradiction.

According to Claims (1) and (2), hospital $\mu(d) \in \bigcup_{k=1}^{k(d)} H(k)$ will make no improvement at ν. So $\mu(d)$ is not in the coalition S because d is not assigned to h at ν. Since matching ν is permissible at μ, hospital $\mu(d)$ is not underdemanded, i.e. $|\mu(\mu(d))| > |\nu(\mu(d))| \geq l(\mu(d))$. According to the SDA-D procedure and $\mu(d)$ is removed in round $k(\mu(d)) = k(d)$, region $r(\mu(d))$ is either underdemanded or non-wasteful at matching $\mu^{k(\mu(d))}$ for subproblem $P^{k(\mu(d))}$.

If region $r(\mu(d))$ is underdemanded at matching $\mu^{k(h)}$ for problem $P^{k(h)}$, then $|\mu(r(\mu(d)))| = l(r(\mu(d)))$ holds. According to Claims (1) and (2), no hospital in $r(\mu(d))$ makes improvement at ν so that $\nu(d') \notin H_{r(\mu(d))}$ holds for each doctor $d' \in X$. This implies that the number of doctors matched to every hospital in $r(\mu(d))$ at ν is no larger than that at μ. The "earliest" doctor d becomes strictly better at ν so that $\nu(d) \notin r(\mu(d))$ and hence $|\nu(r(\mu(d)))| < l(r(\mu(d)))$, which contradicts the fact that ν is permissible at μ.

So region $r(\mu(d))$ is nonwasteful i.e., $|\mu(r(\mu(d)))| > l(r(\mu(d)))$ at matching $\mu^{k(h)}$ for problem $P^{k(h)}$. We will complete the proof work by proving that d will not be matched to any hospital $h \in \bigcup_{j=k(d)+1}^{K} H^j$ at ν. Suppose $\nu(d) \in \bigcup_{j=k(d)+1}^{K} H^j$ in contradiction. According to the SDA-D procedure, there does not exist any underdemanded doctor, underdemanded hospital or underdemanded region in round $k(d)$ because the removed region $r(\mu(d))$ is nonwasteful. So in round $k(d)$, all nonwasteful regions are removed for problem $P^{k(h)}$. According to Proposition 8, every remaining region $r \in R^{k(d)+1}$ has $|\mu^{k(d)}(r)| = q(r)$. Otherwise, region r has $|\mu^{k(d)}(r)| = \sum_{h \in r} q(h)$ so that the region is nonwasteful in round $k(d)$, which should have been removed. We have shown that no doctor $i \in X$ is matched to hospital $h \in \bigcup_{k=1}^{k(d)} H(k)$ at ν. If d is matched to some hospital $h \in \bigcup_{j=k(d)+1}^{K} H^j$ at ν, then some region $r \in R^{k(d)+1}$ has $|\nu(r)| > \sum_{h \in r} q(h)$, which contradicts the fact that ν is permissible at μ.

Fairness within the same quota type: In the SDA-D algorithm, every doctor who is accepted by a hospital occupies a quota from hospital rigid quotas, or region rigid quotas, or region elastic quotas. Let (μ^k, σ^k) be the outcome produced by the SDA-D algorithm in round k. Suppose that (μ, σ) is not fair within the same type quota. That is, there exists a pair (d, h) such that $d, d' \in \sigma^{-1}(\tau_i)$, $\mu(d') = h \succ_d \mu(d)$ and $d \succ_h d'$, where $\tau_i \in \mathcal{T}_r \cup \mathcal{T}_e$. This implies that $r(\mu(d)) = r(\mu(d'))$. Otherwise, d and d' will not obtain the same type quota.

Let k be the round when region $r = r(h)$ is removed by the SDA-D algorithm. Hospital h receives at least one quota from region rigid quotas or region elastic quotas of $r(h)$. So hospital h is removed in round k because any hospital $h' \in r$ removed before round k is underdemanded so that $\sigma^{-1}(\tau_i) \cap \mu(h) = \emptyset$ ($\tau_i \in \{\tau_{rr}, \tau_{er}\}$). The assignment of positions with region r's rigid quota or elastic quota is determined in round k. Let H' be the set of underdemanded hospitals in region $r(h)$. Clearly, $\mu(d), \mu(d') \notin H'$ holds because an underdemanded hospital does not obtain a quota from region rigid quotas or region elastic quotas. So $\mu(d), \mu(d') \in r(h) \setminus H'$. Since μ^k is fair (Theorem 1) and $d, d' \in D^k$, then $d' \succ_h d$ holds, which yields a contradiction.

□

Proof of Proposition 5: We prove via an example that acting truthfully is not optimal for doctors in a non-wasteful mechanism that doctor-dominates the DA-D mechanism. Suppose that there are two doctors $D = \{d_1, d_2\}$ and three hospitals $H = \{h_1, h_2, h_3\}$ in two regions with $r_1 = \{h_1, h_2\}$ and $r_2 = \{h_3\}$. Consider the following market distributional constraints:

- Binding ceiling constraints and unbinding floor constraints for the two regions:

  $q(h_1) = q(h_2) = q(h_3) = 1, l(h_1) = l(h_2) = l(h_3) = 0$ and $q(r_1) = q(r_2) = 1$ and $l(r_1) = l(r_2) = 0$.

- Unbinding ceiling constraints and binding floor constraints for the two regions:

  $q(h_1) = q(h_2) = q(h_3) = 1, l(h_1) = l(h_2) = l(h_3) = 0$ and $q(r_1) = q(r_2) = \infty$ and $l(r_1) = 1$ and $l(r_2) = 1$.

- Binding ceiling constraint for the region 1 and binding floor constraint for the region 2:

  $q(h_1) = q(h_2) = q(h_3) = 1, l(h_1) = l(h_2) = l(h_3) = 0, q(r_1) = 1, q(r_2) = \infty$ and $l(r_1) = 0, l(r_2) = 1$.

Let P be the true preference profile in which the preferences of doctors and hospitals are given by

$$\succ_{d_1}: h_1, h_3, h_2 \qquad \succ_{d_2}: h_2, h_3, h_1$$
$$\succ_{h_1}: d_2, d_1 \qquad \succ_{h_2}: d_1, d_2 \qquad \succ_{h_3}: d_1, d_2$$

Under any of the three given distributional constraints, there are only four feasible matchings as given in the Table 9. For the given preferences, fair matchings are μ_1, μ_2, and μ_3, and non-wasteful matchings are μ_2 and μ_3.

Consider the DA-D mechanism φ^{DA} and any non-wasteful mechanism φ^{NW} that doctor-dominates φ^{DA}. Under the preference profile P, according to the precess of the DA-D mechanism φ^{DA}, it produces one matching from μ_2, and μ_3. We prove the proposition by discussing the following two cases.

Table 9: Four feasible matchings.

Matching	h_1	h_2	h_3
μ_1	d_2	$\emptyset$	d_1
μ_2	$\emptyset$	d_2	d_1
μ_3	d_1	$\emptyset$	d_2
μ_4	$\emptyset$	d_1	d_2

Case (1). Under the preference profile P, the DA-D mechanism produces $\varphi^{DA}(P) = \mu_2$. Then the mechanism $\varphi^{NW}(P) = \mu_2$ because there is only one non-wasteful matching μ_2 that doctor-dominates μ_2 under P.

We show that doctor d_1 has an incentive to misreport his/her preference. Consider the following d_1's preference:

$$\succ'_{d_1}: h_1, h_2, h_3$$

Let P^* be the corresponding preference profile, in which doctor d_1 reports $\succ'_{d_1}$ and other doctors and hospitals report their true preferences. Under the preference profile P^*, fair matchings are μ_1, μ_3, μ_4 and non-wasteful matchings are μ_2, μ_3. There are two sub-cases. Case (1.a). When the tie between hospitals $\succ^t$ has $h_2 \succ^t h_1$, the DA-D mechanism produces $\varphi^{DA}(P) = \mu_4$. Case (1.b). When the tie between hospitals $\succ^t$ has $h_1 \succ^t h_2$, the DA-D mechanism produces $\varphi^{DA}(P) = \mu_3$. In both cases, the mechanism $\varphi^{NW}(P^*) = \mu_3$ because non-wasteful matching μ_3 doctor-dominates μ_4 under P^* and no other non-wasteful matching doctor-dominates μ_3 and μ_4.. In this case, $\varphi^{NW}_{d_1}(P^*) \succ_{d_1} \varphi^{NW}_{d_1}(P)$. Clearly, doctor d_1 has incentive to misreport his/her preference.

Case (2). Under the preference profile P, the DA-D mechanism produces $\varphi^{DA}(P) = \mu_3$. Then the mechanism $\varphi^{NW}(P) = \mu_3$ because non-wasteful matching μ_3 doctor-dominates μ_3 under P and no other non-wasteful matching doctor-dominates μ_3.

We show that doctor d_2 has an incentive to misreport his/her preference. Consider the following d_2's preference:

$$\succ'_{d_2}: h_2, h_1, h_3$$

Let P' be the corresponding preference profile in which doctor d_2 reports $\succ'_{d_2}$ and other doctors and hospitals report their true preferences. Under the preference profile P', fair matchings are μ_1, μ_2, μ_4 and non-wasteful matchings are μ_2, μ_3. There are two sub-cases. Case (1.a). When the tie between hospitals $\succ^t$ has $h_2 \succ^t h_1$, the DA-D mechanism produces $\varphi^{DA}(P) = \mu_2$. Case (1.b). When the tie between hospitals $\succ^t$ has $h_1 \succ^t h_2$, the DA-D mechanism produces $\varphi^{DA}(P) = \mu_1$. In both cases, the mechanism $\varphi^{NW}(P') = \mu_2$ because non-wasteful matching μ_2 doctor-dominates μ_1 under P' and no other non-wasteful matching doctor-dominates μ_1 and μ_2. In this case, $\varphi^{NW}_{d_2}(P') \succ_{d_2} \varphi^{NW}_{d_2}(P)$. Clearly, doctor d_2 has incentive to misreport his/her preference.

In summary the mechanism φ^{NW} is not strategy-proof.

□

Proof of Proposition 6: Let μ be a feasible matching. Suppose μ is not floor-respecting pairwise stable. There exists a doctor-hospital pair (d, h) that can block μ such that one of the following cases happens:

(i) if $|\mu(h)| < q(h)$, then matching μ' satisfying $\mu'(d) = h$ and $\mu'(d') = \mu(d')$ for all $d' \in D \setminus \{d\}$ is feasible;

(ii) if $|\mu(h)| = q(h)$, then matching μ' satisfying $\mu'(d) = h$, $\mu'(d') = h_0$ for some $d' \in \mu(h)$, and $\mu'(d'') = \mu(d'')$ for all $d'' \in D \setminus \{d, d'\}$ is feasible.

We will show that μ is not pairwise stable either. In both cases (i) and (ii), that μ' is feasible indicates that a new matching ν satisfying $\nu(d) = \mu'(d)$ and $\nu(d') = \mu(d')$ for all $d' \in D \setminus \{d\}$ is permissible at μ because all floor constraints are satisfied and all regions have unbinding ceiling constraint. Therefore, (d, h) is an admissible blocking pair of μ and μ is not pairwise stable. □

Proof of Proposition 7: Let μ be a feasible matching. Suppose μ is not floor-respecting stable. There exists a coalition S that blocks μ via a feasible matching μ' such that

(i) $\mu'(d) \in S$ and $\mu'(d) \succeq_d \mu(d)$ for all doctors $d \in S \cap D$

(ii) $\mu'(h) \subseteq S$ and $\mu'(h) \succeq_h \mu(h)$ for all $h \in A \cap H$

(iii) $\mu'(i) \succ_i \mu(i)$ for some doctor or hospital $i \in S$

(iv) if doctor $d \notin S$, then $\mu'(d) = \mu(d)$ or $\mu'(d) = \emptyset$.

We show that μ is not group stable. Since μ' is feasible, then a new matching ν satisfying $\nu(d) = \mu'(d)$ for all $d \in S$ and $\nu(d) = \mu(d)$ for all $d \in D \setminus S$ is permissible at μ because all floor constraints are satisfied and all regions have unbinding ceiling constraints. This means that S admissibly blocks μ and hence μ cannot be group stable. □

References

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

Abdulkadiroglu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. American Economic Review, 93(3):729–747.

Akin, S. (2021). Matching with floor constraints. Theoretical Economics, 16(3):911–942.

Andersson, T., Sevensson, L. G., and Yang, Z. (2010). Constrained fair job assignments under minimum wages. Games and Economic Behavior, 68(2):428–442.

Ausubel, L. (2006). An efficient dynamic auction for heterogeneous commodities. American Economic Review, 96:602–629.

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: Student placement. Journal of Economic theory, 84(1):73–94.

Biró, P., Fleiner, T., Irving, R. W., and Manlove, D. F. (2010). The college admissions problem with lower and common quotas. Theoretical Computer Science, 411(34-36):3136–3153.

Dubins, L. and Freedman, D. (1981). Maciavelli and the Gale-Shapley algorithm. American Mathematical Monthly, 88:485–494.

Ehlers, L., Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. (2014). School choice with controlled choice constraints: Hard bounds versus soft bounds. Journal of Economic Theory, 153:648–683.

Foley, D. (1967). Resource allocation and the public sector. Yale Economic Essays, 7:45–98.

Fragiadakis, D., Iwasaki, A., Troyan, P., Ueda, S., and Yokoo, M. (2016). Strategyproof matching with minimum quotas. ACM Transactions on Economics and Computation, 4(1):6.

Fragiadakis, D. and Troyan, P. (2017). Improving matching under hard distributional constraints. Theoretical Economics, 12(2):863–908.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. American Mathematical Monthly, 69(1):9–15.

Goto, M., Iwasaki, A., Kawasaki, Y., Yasuda, Y., and Yokoo, M. (2014). Improving fairness and efficiency in matching with distributional constraints: An alternative solution for the Japanese medical residency match.

Goto, M., Kurata, R., Hamada, N., Iwasaki, A., and Yokoo, M. (2015). Improving fairness in nonwasteful matching with hierarchical regional minimum quotas. In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, pages 1887–1888. International Foundation for Autonomous Agents and Multiagent Systems.

Hatfield, J. and Milgrom, P. (2005). Matching with contracts. American Economic Review, 95:913–935.

Huang, C.-C. (2010). Classified stable matching. In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, pages 1235–1245. Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms.

Hurwicz, L. (1973). The design of mechanisms for resource allocation. American Economic Review, 63:1–39.

Kamada, Y. and Kojima, F. (2015). Efficient matching under distributional constraints: Theory and applications. American Economic Review, 105(1):67–99.

Kamada, Y. and Kojima, F. (2017). Stability concepts in matching under distributional constraints. Journal of Economic theory, 168:107–142.

Kamada, Y. and Kojima, F. (2018). Stability and strategy-proofness for matching with constraints: A necessary and sufficient condition. Theoretical Economics, 13(2):761–793.

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

Kesten, O. and Yazici, A. (2012). The Pareto-dominant strategy-proof and fair rule for problems with indivisible goods. Economic Theory, 50(2):463–488.

Li, S. and Zhong, W. (2006). An empirical study on fairness and efficiency of medical services in china (in Chinese). Management World, 5:146–147.

Report (2020). High-quality medical resources are unevenly distributed? A pilot project for the construction of regional medical centers (in Chinese). Chinese National Development and Reform Commission Report. See http://health.china.com.cn/2020-07/29/content_41236451.htm.

Report (2021). To promote the balanced distribution of high-quality medical resources and improve the hierarchical diagnosis and treatment system (in Chinese). Chinese Government Report. See https://www.gov.cn/xinwen/2021-07/08/content_5623639.htm.

Roth, A. and Sotomayor, M. (1990). Two-sided Matching. Cambridge University Press.

Roth, A. E. (1982). Incentive compatibility in a market with indivisible goods. Economics Letters, 9:127–132.

Roth, A. E. (1985). The college admissions problem is not equivalent to the marriage problem. Journal of economic Theory, 36(2):277–288.

Sönmez, T. and Switzer, T. B. (2013). Matching with (branch-of-choice) contracts at the United States Military Academy. Econometrica, 81(2):451–488.

Sun, N. and Yang, Z. (2003). A general strategy-proof fair allocation mechanism. Economics Letters, 81(1):73–79.

Sun, N. and Yang, Z. (2014). An efficient and incentive-compatible dynamic auction for multiple complements. Journal of Political Economics, 122:422–465.

Tomoeda, K. (2018). Finding a stable matching under type-specific minimum quotas. Journal of Economic Theory, 176:81–117.

Varian, H. (1974). Equity, envy, and efficiency. Journal of Economic Theory, 9:63–91.

Xia, W., Ke, P., Di, H., Wang, C., and Lu, Z. (2021). Evaluation of national medical resource allocation efficiency based on data envelopment analyses (in Chinese). Chinese Journal of Health Policy, 14(7):68–73.

Yu, Y. and Yang, D. (2008). A theoretical and empirical study on the fairness of medical resource allocation in china (in Chinese). Reform of Economic System, 6:160–163.

Zheng, J.-C. (2019). Study on the equilibrium of health resource allocation in China (in Chinese). Chinese Health Resources, 22(5):362–366.