EFFICIENT MATCHING UNDER DISTRIBUTIONAL CONSTRAINTS: THEORY AND APPLICATIONS

YUICHIRO KAMADA AND FUHITO KOJIMA

Cowles Foundation for Research in Economics, Yale University, New Haven, CT 06511,
and Haas School of Business, University of California, Berkeley, Berkeley, CA 94720
y.cam.24@gmail.com

Department of Economics, Stanford University, Stanford, CA 94305
fkojima@stanford.edu

ABSTRACT. Many real matching markets are subject to distributional constraints. These constraints often take the form of restrictions on the numbers of agents on one side of the market matched to certain subsets of the other side. Real-life examples include restrictions imposed on regions in medical residency matching, academic master's programs in graduate school admission, and state-financed seats for college admission. Motivated by these markets, we study the design of matching mechanism under distributional constraints. We show that the existing matching mechanisms around the world may result in avoidable inefficiency and instability, and propose a better mechanism that has desirable properties in terms of efficiency, stability, and incentives while respecting the distributional constraints.

JEL Classification Numbers: C70, D61, D63.

Keywords: medical residency matching, college and graduate school admission, distributional constraints, efficiency, stability, strategy-proofness, matching with contracts, the rural hospital theorem

1. INTRODUCTION

Many real matching markets are subject to distributional constraints. In health care, there is often a concern that certain medical specialties attract too many doctors while others suffer from shortage, and regulations on the number of doctors in certain specialties are imposed or proposed.[1] In some public school districts, multiple school programs often share one school building, so there is a bound on the total number of students in these programs in addition to each program's capacity because of the building's physical size.[2] Achieving demographic balance of the incoming class is often one of the most important goals in school and college admission, which may lead to distributional constraints on the matching outcome.

An interesting example of a concrete matching market with such distributional constraints is the one for Japanese medical residency in which around 8,000 doctors (mostly consisting of graduating medical students) are matched to about 1,500 residency programs each year. In 2008, the Japanese government introduced a "regional cap" which, for each of the 47 prefectures that partition the country, restricts the total number of residents

---

[1]In the United States, the plaintiff of *Jung v. Association of American Medical Colleges* alleged that the association called Accreditation Council for Graduate Medical Education regulates the number of medical residents in each specialty, although the allegation was dismissed. Nicholson (2003) finds some suggesting evidence that Residency Review Committees, which works closely with ACGME, essentially has complete control over the number of residents who train in each specialty.

[2]See Abdulkadiroğlu and Sönmez (2003) for the introduction to school choice problems.

matched within the prefecture. This measure was taken to regulate the geographical distribution of doctors, which would otherwise be concentrated too much in urban areas at the expense of rural areas.[3] Since the introduction of the regional caps, they have used a mechanism that respects the caps, which is a modification of the standard deferred acceptance mechanism that they used before 2008. Specifically, in the modified mechanism, which we call the Japan Residency Matching Program (JRMP) mechanism, if the sum of the hospital capacities in a region exceeds its regional cap, then the capacity of each hospital is reduced to equalize the total capacity with the regional cap. Then the deferred acceptance algorithm is implemented under the reduced capacities.[4]

A similar policy is taken in the context of graduate school admission in China, which has placed more than 400,000 students every year since 2009, where master's programs are categorized as either academic or professional. To enlarge the labor force with professional master's degrees, in 2010 the Chinese government started restricting the number of admissions to academic master's programs. More specifically, the government decided to reduce the number of available seats of each academic master's program by about 25 percent by 2015, just as in Japan where a rigid restriction is imposed on each hospital. A clearinghouse mechanism is run given the reduced capacities.

In the context of college admission, the Ukrainian government provides a limited number of "state-financed" seats in public universities, for which tuition is free for the students. The number of state-financed seats and that of privately financed seats in each college are determined first, and then a matching procedure assigns students to these college-specific state-financed or privately financed seats. The situation resembles the market for

---

[3]Regional imbalance of doctors is a serious concern in many other parts of the world as well. For instance, a Washington Post article entitled "Shortage of Doctors Affects Rural U.S." describes a dire situation in the United States (Talbott, 2007) as follows: "The government estimates that more than 35 million Americans live in underserved areas, and it would take 16,000 doctors to immediately fill that need, according to the American Medical Association." Similar problems are present around the world. For example, one can easily find reports of doctor shortages in rural areas in the United Kingdom, India, Australia, and Thailand (see Shallcross (2005), Alcoba (2009), Nambiar and Bavas (2010), and Wongruang (2010)).

[4] The capacity of a hospital is reduced proportionately to its original capacity in principle (subject to integrality constraints) although there are a number of fine adjustments and exceptions. This rule might suggest that hospitals have incentives to misreport their true capacities, but in Japan, the government regulates how many positions each hospital can offer so that the capacity can be considered exogenous. More specifically, the government decides the physical capacity of a hospital based on verifiable information such as the number of beds in it.

Japanese medical residents if the total number of state-financed seats is interpreted as a regional cap.

A different type of mechanism is used in the U.K. medical match in which about 7,000 new doctors participate. Their matching procedure has two rounds, in which a doctor is assigned to one of the country's 25 regions first, and then to a program within their assigned region. A similar two-round procedure is employed in the matching of new teachers in Scotland as well.

An interesting observation is that distributional constraints appear in many different contexts (we discuss more examples in the Appendix), and many different policies have been tried to accommodate these constraints, but it has been unclear which mechanism in practice, if any, achieves appealing properties such as efficiency, stability, and incentive compatibility under such constraints.

Motivated by these real-life policies, we study the design of matching markets under constraints on the doctor distribution. This paper shows that each of those existing mechanisms may result in avoidable instability and inefficiency. We further find that a mechanism that appears to be intuitively appealing (and is similar to a mechanism that policy makers often informally suggest to us) suffers from incentive problems. We propose an alternative mechanism that overcomes these shortcomings while respecting the distributional goals. More specifically, we first introduce concepts of stability and (constrained) efficiency that take distributional constraints into account. We point out that none of the aforementioned existing mechanisms always produces a stable or efficient matching, and present a new mechanism that we call the flexible deferred acceptance mechanism. We show that, unlike other mechanisms, this mechanism finds a stable and efficient matching and is (group) strategy-proof for doctors.[5] In addition, the flexible deferred acceptance mechanism matches weakly more doctors to hospitals (in the sense of set inclusion) and makes every doctor weakly better off than the JRMP mechanism. These results suggest that replacing the current mechanisms with the flexible deferred acceptance mechanism will improve the performance of the matching market.

We also find that the structural properties of the stable matchings with distributional constraints are strikingly different from those in the standard matching models. First, there does not necessarily exist a doctor-optimal stable matching (a stable matching unanimously preferred to every stable matching by all doctors). Neither do there exist hospital-optimal nor doctor-pessimal nor hospital-pessimal stable matchings. Second,

---

[5]A mechanism being (group) strategy-proof for doctors means that telling the truth is a dominant strategy for each doctor (and even a coalition of doctors cannot jointly misreport preferences and benefit).

different stable matchings can leave different hospitals with unfilled positions, implying that the conclusion of the rural hospital theorem fails in our context. Based on these observations, we investigate whether the government can design a reasonable mechanism that selects a particular stable matching based on its policy goals such as minimizing the number of unmatched doctors.

Let us emphasize that analyzing abstract technical issues associated with distributional constraints is *not* the primary purpose of this paper. On the contrary, we study a model motivated by various real markets and offer practical solutions for these markets. Improving the existing markets is important by itself, such as the Japanese residency markets which produces around 8,000 medical doctors, the Chinese graduate admission which admits more than 400,000 new master students, and the UK medical match which involves around 7,000 doctors. Moreover, this paper tries to provide a general framework in which one can tackle problems arising in practical markets which have yet to be recognized or addressed. In these senses, our paper contributes to the general research agenda of market design, advocated by Roth (2002) for instance, that emphasizes the importance of addressing issues arising in practical allocation problems.

The rest of this paper proceeds as follows. In Section 2, we present the model of matching with regional caps. In Section 3 where we define the JRMP mechanism, we show that it does not necessarily produce an efficient matching and there is a sense in which the produced matching is not stable. Section 4 introduces and analyzes stability concepts under distributional constraints. In Section 5 we propose a new mechanism, the flexible deferred acceptance mechanism, and show that it produces a stable and efficient matching and is group strategy-proof. Section 6 discusses a number of further topics, Section 7 discusses the related literature, and Section 8 concludes. Proofs are in the Appendix unless stated otherwise.

2. Model

This section presents the model of matching with distributional constraints. Motivated by Japanese residency matching, we describe the model in terms of matching between doctors and hospitals, where there is a "regional cap," that is, an upper bound on the number of doctors that can be matched to hospitals in each region.[6] Later we discuss other matching problems in practice such as Chinese graduate school admission, U.K. medical

---

[6]In practice, there may be mulitple residency programs within a hospital. However, in such a case we use the word "a hospital" to mean a residency program for simplicity.

matching, Scottish teacher matching, and college admissions in Ukraine and Hungary. The model is also applicable to diverse contexts discussed in the Introduction, such as doctor-hospital matching with specialty constraints, school choice with building size constraints, and matching with affirmative action constraints.

Let there be a finite set of doctors D and a finite set of hospitals H. Each doctor d has a strict preference relation $\succ_d$ over the set of hospitals and being unmatched (being unmatched is denoted by $\emptyset$). For any $h, h' \in H \cup \{\emptyset\}$, we write $h \succeq_d h'$ if and only if $h \succ_d h'$ or $h = h'$. Each hospital h has a strict preference relation $\succ_h$ over the set of subsets of doctors. For any $D', D'' \subseteq D$, we write $D' \succeq_h D''$ if and only if $D' \succ_h D''$ or $D' = D''$. We denote by $\succ = (\succ_i)_{i \in D \cup H}$ the preference profile of all doctors and hospitals.

Doctor d is said to be **acceptable** to h if $d \succ_h \emptyset$.[7] Similarly, h is acceptable to d if $h \succ_d \emptyset$. It will turn out that only rankings of acceptable partners matter for our analysis, so we often write only acceptable partners to denote preferences. For example,

$$\succ_d: h, h'$$

means that hospital h is the most preferred, h' is the second most preferred, and h and h' are the only acceptable hospitals under preferences $\succ_d$ of doctor d.

Each hospital $h \in H$ is endowed with a (physical) **capacity** q_h, which is a nonnegative integer. We say that preference relation $\succ_h$ is **responsive with capacity** q_h (Roth, 1985) if

(1) For any $D' \subseteq D$ with $|D'| \leq q_h$, $d \in D \setminus D'$ and $d' \in D'$, $(D' \cup d) \setminus d' \succeq_h D'$ if and only if $d \succeq_h d'$,

(2) For any $D' \subseteq D$ with $|D'| \leq q_h$ and $d' \in D'$, $D' \succeq_h D' \setminus d'$ if and only if $d' \succeq_h \emptyset$, and

(3) $\emptyset \succ_h D'$ for any $D' \subseteq D$ with $|D'| > q_h$.

In words, preference relation $\succ_h$ is responsive with a capacity if the ranking of a doctor (or keeping a position vacant) is independent of her colleagues, and any set of doctors exceeding its capacity is unacceptable. We assume that preferences of each hospital h are responsive with capacity q_h throughout the paper.

There is a finite set R which we call the set of **regions**. The set of hospitals H is partitioned into hospitals in different regions, that is, $H_r \cap H_{r'} = \emptyset$ if $r \neq r'$ and $H = \cup_{r \in R} H_r$, where H_r denotes the set of hospitals in region $r \in R$. For each $h \in H$, let $r(h)$

---

[7]We denote singleton set $\{x\}$ by x when there is no confusion.

denote the region r such that $h \in H_r$. For each region $r \in R$, there is a **regional cap** q_r, which is a nonnegative integer.

A **matching** μ is a mapping that satisfies (i) $\mu_d \in H \cup \{\emptyset\}$ 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 $d \in \mu_h$. That is, a matching simply specifies which doctor is assigned to which hospital (if any). A matching is **feasible** if $|\mu_r| \le q_r$ for all $r \in R$, where $\mu_r = \cup_{h \in H_r} \mu_h$. In other words, feasibility requires that the regional cap for every region is satisfied. This requirement distinguishes the current environment from the standard model without regional caps: We allow for (though do not require) $q_r < \sum_{h \in H_r} q_h$, that is, the regional cap can be smaller than the sum of hospital capacities in the region.

Since regional caps are a primitive of the environment, we consider a constrained efficiency concept. A feasible matching μ is **(constrained) efficient** if there is no feasible matching μ' such that $\mu'_i \succeq_i \mu_i$ for all $i \in D \cup H$ and $\mu'_i \succ_i \mu_i$ for some $i \in D \cup H$.

To accommodate the regional caps, we introduce new stability concepts that generalize the standard notion. For that purpose, we first define two basic concepts. A matching μ is **individually rational** if (i) for each $d \in D$, $\mu_d \succeq_d \emptyset$, and (ii) for each $h \in H$, $d \succeq_h \emptyset$ for all $d \in \mu_h$, and $|\mu_h| \le q_h$. That is, no agent is matched with an unacceptable partner and each hospital's capacity is respected.

Given matching μ, a pair (d, h) of a doctor and a hospital is called a **blocking pair** if $h \succ_d \mu_d$ and either (i) $|\mu_h| < q_h$ and $d \succ_h \emptyset$, or (ii) $d \succ_h d'$ for some $d' \in \mu_h$. In words, a blocking pair is a pair of a doctor and a hospital who want to be matched with each other (possibly rejecting their partners in the prescribed matching) rather than following the proposed matching.

When there are no binding regional caps (in the sense that $q_r \ge \sum_{h \in H_r} q_h$ for every $r \in R$), a matching is said to be stable if it is individually rational and there is no blocking pair. Gale and Shapley (1962) show that there exists a stable matching in that setting. In the presence of binding regional caps, however, there may be no such matching that is feasible (in the sense that all regional caps are respected). Thus in some cases every feasible and individually rational matching may admit a blocking pair.

A **mechanism** φ is a function that maps preference profiles to matchings. The matching under φ at preference profile $\succ$ is denoted $\varphi(\succ)$ and agent i's match is denoted by $\varphi_i(\succ)$ for each $i \in D \cup H$.

A mechanism φ is said to be **strategy-proof** if there does not exist a preference profile $\succ$, an agent $i \in D \cup H$, and preferences $\succ'_i$ of agent i such that

$$\varphi_i(\succ'_i, \succ_{-i}) \succ_i \varphi_i(\succ).$$

That is, no agent has an incentive to misreport her preferences under the mechanism. Strategy-proofness is regarded as a very important property for a mechanism to be successful.[8]

Unfortunately, however, there is no mechanism that produces a stable matching for all possible preference profiles and is strategy-proof even in a market without regional caps, that is, $q_r > |D|$ for all $r \in R$ (Roth, 1982).[9] Given this limitation, we consider the following weakening of the concept requiring incentive compatibility only for doctors. A mechanism φ is said to be **strategy-proof for doctors** if there does not exist a preference profile $\succ$, a doctor $d \in D$, and preferences $\succ'_d$ of doctor d such that

$$\varphi_d(\succ'_d, \succ_{-d}) \succ_d \varphi_d(\succ).$$

A mechanism φ is said to be **group strategy-proof for doctors** if there is no preference profile $\succ$, a subset of doctors $D' \subseteq D$, and a preference profile $(\succ'_{d'})_{d' \in D'}$ of doctors in D' such that

$$\varphi_d((\succ'_{d'})_{d' \in D'}, (\succ_i)_{i \in D \cup H \setminus D'}) \succ_d \varphi_d(\succ) \text{ for all } d \in D'.$$

That is, no subset of doctors can jointly misreport their preferences to receive a strictly preferred outcome for every member of the coalition under the mechanism.

We do not necessarily regard (group) strategy-proofness for doctors as a minimum desirable property that our mechanism should satisfy (our criticism of the JRMP mechanism in Section 3 does not hinge on (group) strategy-proofness), but it will turn out that the flexible deferred acceptance mechanism we propose in Section 5 does have this property.

As this paper analyzes the effect of regional caps in matching markets, it is useful to compare it with the standard matching model without regional caps. Gale and Shapley (1962) consider a matching model without any binding regional cap, which corresponds

---

[8]One good aspect of having strategy-proofness is that the matching authority can actually state it as the property of the algorithm to encourage doctors to reveal their true preferences. For example, the current webpage of the JRMP (last accessed on May 25, 2010, http://www.jrmp.jp/01-ryui.htm) states, as advice for doctors, that "If you list as your first choice a program which is not actually your first choice, the probability that you end up being matched with some hospital does not increase [...] the probability that you are matched with your actual first choice decreases." In the context of student placement in Boston, strategy-proofness was regarded as a desirable fairness property, in the sense that it provides equal access for children and parents with different degrees of sophistication to strategize (Pathak and Sonmez, 2008).

[9]Remember that a special case of our model in which $q_r > |D|$ for all $r \in R$ is the standard matching model with no binding regional caps.

to a special case of our model in which $q_r > |D|$ for every $r \in R$. In that model, they propose the following **(doctor-proposing) deferred acceptance algorithm**:

- Step 1: Each doctor applies to her first choice hospital. Each hospital rejects the lowest-ranking doctors in excess of its capacity and all unacceptable doctors among those who applied to it, keeping the rest of the doctors temporarily (so doctors not rejected at this step may be rejected in later steps).

In general,

- Step t: Each doctor who was rejected in Step $(t-1)$ applies to her next highest choice (if any). Each hospital considers these doctors *and* doctors who are temporarily held from the previous step together, and rejects the lowest-ranking doctors in excess of its capacity and all unacceptable doctors, keeping the rest of the doctors temporarily (so doctors not rejected at this step may be rejected in later steps).

The algorithm terminates at a step in which no rejection occurs. The algorithm always terminates in a finite number of steps. Gale and Shapley (1962) show that the resulting matching is stable in the standard matching model without any binding regional cap.

Even though there exists no strategy-proof mechanism that produces a stable matching for all possible inputs, the deferred acceptance mechanism is (group) strategy-proof for doctors (Dubins and Freedman, 1981; Roth, 1982).[10] This result has been extended by many subsequent studies, suggesting that the incentive compatibility of the mechanism is quite robust and general.[11]

3. A Motivating Example of an Inefficient Mechanism

In this section we formulate the Japan Residency Matching Program (JRMP) mechanism, one of the mechanisms with distributional constraints used in practice, and point out its deficiencies. As we have mentioned in the Introduction, there are a number of practical markets besides the Japanese one in which distributional constraints are imposed, such as the UK medical match, the admission problem for the Chinese master's programs, and so forth. Here we chose the Japanese case because the JRMP mechanism

---

[10]Ergin (2002) defines a stronger version of group strategy-proofness. It requires that no group of doctors can misreport preferences jointly and make some of its members strictly better off without making any of its members strictly worse off. He identifies a necessary and sufficient condition for the deferred acceptance mechanism to satisfy this version of group strategy-proofness.

[11]Researches generalizing (group) strategy-proofness of the mechanism include Abdulkadiroğlu (2005), Hatfield and Milgrom (2005), Martinez, Masso, Neme, and Oviedo (2004), Hatfield and Kojima (2009, 2010), and Hatfield and Kominers (2009, 2010).

is easy to formulate so its problem can be transparently seen and a clear-cut comparison with our mechanism can be made. Other mechanisms are analyzed in subsequent sections.

In the JRMP mechanism, there is an exogenously given (government-imposed) non-negative integer $\bar{q}_h \leq q_h$, which we call **target capacity**, for each hospital h such that $\sum_{h \in H_r} \bar{q}_h \leq q_r$ for each region $r \in R$.[12,13] The **JRMP mechanism** is a rule that produces the matching resulting from the deferred acceptance algorithm except that, for each hospital h, it uses $\bar{q}_h$ instead of q_h as the hospital's capacity.

The JRMP mechanism is based on a simple idea: In order to satisfy regional caps, simply force hospitals to be matched to a smaller number of doctors than their real capacities, but otherwise use the standard deferred acceptance algorithm. Note, however, that target capacities are not feasibility constraints by themselves: the goal of Japanese policy makers is to satisfy regional caps and target capacities were introduced to achieve that goal.

Although the mechanism is a variant of the deferred acceptance algorithm, it suffers from at least one problem. Despite the government's intention, the result of the JRMP mechanism is not necessarily efficient, as seen in the following example.

Example 1 (JRMP mechanism does not necessarily produce an efficient matching)**.** There is one region r with regional cap $q_r = 10$, in which two hospitals, h_1 and h_2, reside. Each hospital h has a capacity of $q_h = 10$. Suppose that there are 10 doctors, $d_1, \ldots, d_{10}$. Preference profile $\succ$ is as follows:

$$\succ_{h_i} : d_1, d_2, \ldots, d_{10} \ \text{ for } i = 1, 2,$$

$$\succ_{d_j} : h_1 \text{ if } j \leq 3 \quad \text{ and } \quad \succ_{d_j} : h_2 \text{ if } j \geq 4.$$

---

[12]Note that we allow the sum of target capacities to be strictly smaller than the regional cap. This is necessary if the sum of hospital capacities is strictly smaller than the regional cap; we allow this possibility even otherwise. All results, including (counter)examples, hold when we assume that the sum of target capacities is equal to the regional cap.

[13]In our model, $\bar{q}_h$ is exogenously given for each hospital h. In the current Japanese system, if the sum of the hospitals' capacities exceeds the regional cap, then the target $\bar{q}_h$ of each hospital h is set at an integer close to $\frac{q_r}{\sum_{h' \in H_r} q_{h'}} \cdot q_h$. That is, each hospital's target is (roughly) proportional to its capacity. This might suggest that hospitals have incentives to misreport their capacities. As explained in footnote 4, however, the capacity can be considered exogenous in the Japanese context.

Thus, three doctors prefer hospital h_1 to being unmatched (the option $\emptyset$) to hospital h_2, while the other seven doctors prefer hospital h_2 to being unmatched to hospital h_1. Let the target capacities be $\bar{q}_{h_1} = \bar{q}_{h_2} = 5$.[14]

At the first round of the JRMP algorithm, doctors d_1, d_2 and d_3 apply to hospital h_1, and the rest of the doctors apply to hospital h_2. Hospital h_1 does not reject anyone at this round, as the number of applicants is less than its target capacity, and all applicants are acceptable. Hospital h_2 rejects d_9 and d_{10} and accepts other applicants, because the number of applicants exceeds the *target capacity* (not the hospital's capacity itself!), and it prefers doctors with smaller indices (and all doctors are acceptable). Since d_9 and d_{10} find h_1 unacceptable, they do not make further applications, so the algorithm terminates at this point. Hence the resulting matching μ is such that

$$\mu = \begin{pmatrix} h_1 & h_2 & \emptyset \\ d_1, d_2, d_3 & d_4, d_5, d_6, d_7, d_8 & d_9, d_{10} \end{pmatrix}.$$

Consider a matching μ' defined by,

$$\mu' = \begin{pmatrix} h_1 & h_2 \\ d_1, d_2, d_3 & d_4, d_5, d_6, d_7, d_8, d_9, d_{10} \end{pmatrix}.$$

Since the regional cap is still respected, μ' is feasible. Moreover, every agent is weakly better off with d_9, d_{10}, and h_2 being strictly better off than at μ. Hence we conclude that the JRMP mechanism results in an inefficient matching in this example.[15] $\qquad\square$

Remark 1. We note that there is a sense in which the matching μ is not stable. For example, hospital h_2 and doctor d_9 constitute a blocking pair while the regional cap for r is not binding. That is, even after d_9 is matched with h_2, the total number of doctors in the region is 9, which is less than the regional cap of 10.[16] Although this argument may

---

[14]The specification of target capacities follows the formula used in Japan that we mentioned earlier.

[15]In this example, not all hospitals are acceptable to all doctors. One may wonder whether this is an unrealistic assumption because doctors may be so willing to work that any hospital is acceptable. However, the example can be easily modified so that all hospitals are acceptable to all doctors while some doctors are unacceptable to some hospitals (which may be a natural assumption because, for instance, typically a hospital only lists doctors who they interviewed). Also, in many markets doctors apply to only a small subset of hospitals. In 2009, for instance, a doctor applied to only 3.3 hospitals on average (Japan Residency Matching Program, 2009a).

[16]One may wonder whether the "failure of stability" depends on the assumption that some agents find some of the potential partners unacceptable. However, a similar example can be constructed even if we require every agent finds every potential partner acceptable. For instance, modify the market in the example by introducing another hospital h_3 in another region with regional cap two; let h_3 find every doctor acceptable and have two positions; d_1, d_2 and d_3 prefer h_1 to h_3 to h_2 to being unmatched,

appear straightforward, defining stability in the presence of regional caps is not a trivial task. In Section 4, we define the notion of stability and show that the matching μ is not stable.[17]   □

The above example suggests that a problem of the JRMP mechanism is its lack of flexibility: The JRMP mechanism runs as if the target capacity is the actual capacity of hospitals, thus rejecting an application of a doctor to a hospital unnecessarily. The mechanism that we propose in Section 5 overcomes problems of inefficiency (and stability) by, intuitively speaking, making the target capacities flexible. Before formally introducing this mechanism, we define and discuss the goals that we try to achieve with the mechanism.

4. Goal Setting: Stability Concepts and Strategy-Proofness

As we discussed earlier, there may be no stable matching in the traditional sense that satisfies feasibility. Given this observation, this section defines two weaker stability concepts, in which certain types of blocking pairs are tolerated. The first notion, *strong stability*, is what we think is the most natural. Unfortunately this notion has several drawbacks related to non-existence. The second notion, *stability*, overcomes these drawbacks, and we use this notion from the next section onwards. *The objective in this section is not to discuss technical details of these stability concepts per se, but to set an explicit goal for constructing a new algorithm, which we introduce in Section 5.*

The first notion presented below is meant to capture the idea that any blocking pair that will not violate the regional cap should be considered legitimate, so the appropriate stability concept should require that no agents have incentives to form any such blocking pair. Recall that $r(h)$ is the region that hospital h belongs to.

Definition 1. A matching μ is **strongly stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then (i) $|\mu_{r(h)}| = q_{r(h)}$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and (iii) $\mu_d \notin H_{r(h)}$.

As stated in the definition, only certain blocking pairs are tolerated under strong stability. Any blocking pair that may remain would violate the regional cap since condition (i)

---

while all other doctors prefer h_2 to h_3 to h_1 to being unmatched (thus every doctor finds all hospitals acceptable).

[17]Moreover, in Appendix C, we define a weaker stability concept than the stability concept defined in Section 4, and show that μ does not satisfy this weaker notion of stability either.

implies that the cap for the blocking hospital's region is currently binding, condition (ii) implies that the only blocking involves filling a vacant position, and condition (iii) implies that the blocking doctor is not currently assigned in the hospital's region. In this sense, strong stability requires that any blocking pair is "caused" by regional caps. Indeed, this concept reduces to the standard stability concept of Gale and Shapley (1962) if there are no binding regional caps.

The implicit idea behind the definition is that the government or some authority can interfere and prohibit a blocking pair to be formed if regional caps are an issue. Indeed, in Japan, participants seem to be effectively forced to accept the matching announced by the clearinghouse because a severe punishment is imposed on deviators.[18] One might then wonder "If the government has the power to prohibit a blocking pair in certain cases, why doesn't it have the power to do so in all cases, so why do we care about stability in the first place?"

Our view is that even if the clearinghouse has power to enforce a matching (which may be the case in the Japanese residency match), an assignment that completely ignores participants' preferences would be undesirable. Indeed, as we discussed in Section 2, the introduction of a stable matching mechanism in this market was motivated by the criticism that the previous assignment system was "unfair" and "inefficient," rather than by a desire to prevent participants from circumventing the assignment by forming "blocking pairs."[19] In other words, we view minimizing blocking pairs as a normative criterion.[20] Given this observation, our strong stability captures the idea that it is desirable to minimize blocking pairs so that the only blocking pairs are "caused" by regional caps, which may be a legitimate reason to deny a blocking pair.[21]

Nonetheless, we will not pursue strong stability when we construct an algorithm in Section 5. There are at least two reasons for this. The first reason is that a strongly stable matching does not necessarily exist. The following example demonstrates this point.

---

[18]For example, violating hospitals can be excluded from participating in the matching mechanism in subsequent years (Japan Residency Matching Program, 2010).

[19]Another example of a labor market using a stable mechanism despite being heavily regulated is the labor market for junior academic positions in France (Haeringer and Iehle, 2010).

[20]"No justified envy" in the school choice literature corresponds to "no blocking pair" in our context, and it is viewed as a normative criterion.

[21]Another obvious normative criterion is (constrained) efficiency. Indeed, it will turn out that strong stability (as well as its weakenings that we will discuss in this paper) implies efficiency (Theorem 5). Thus strong stability (as well as its weakenings) has an additional normative appeal.

Example 2 (A strongly stable matching does not necessarily exist). There is one region r with regional cap $q_r = 1$, in which two hospitals, h_1 and h_2, reside. Each hospital h has a capacity of $q_h = 1$. Suppose that there are two doctors, d_1 and d_2. We assume the following preferences:

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

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

First, in any strongly stable matching, there is exactly one doctor matched to some hospital. This is because matching two doctors violate the regional cap, while (d_1, h_1) would constitute a blocking pair that is not tolerated in the definition of strong stability if no doctors is matched to any hospital. By symmetry, it suffices to consider the case in which d_1 is matched (and hence d_2 is unmatched). If d_1 is matched with h_1, then (d_1, h_2) is a non-tolerated blocking pair. On the other hand, if d_1 is matched with h_2, (d_2, h_2) is a non-tolerated blocking pair. Therefore, a strongly stable matching does not exist in this market.   □

Even if a strongly stable matching does not always exist, can we try to achieve a weaker desideratum? More specifically, does there exist a mechanism that selects a strongly stable matching whenever there exists one? We show that such a mechanism does not exist if we also require certain incentive compatibility: There is no mechanism that selects a strongly stable matching whenever there exists one and is strategy-proof for doctors. This is the second reason that we do not attempt to achieve strong stability as a natural desideratum. To see this point consider the following example.

Example 3 (No mechanism that is strategy-proof for doctors selects a strongly stable matching whenever there exists one). There is one region r with regional cap $q_r = 1$, in which two hospitals, h_1 and h_2, reside. Each hospital h has a capacity of $q_h = 1$. Suppose that there are two doctors, d_1 and d_2. We assume the following preferences:

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

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

In this market, there are two strongly stable matchings,

$$\mu = \begin{pmatrix} h_1 & h_2 & \emptyset \\ d_2 & \emptyset & d_1 \end{pmatrix},$$

$$\mu' = \begin{pmatrix} h_1 & h_2 & \emptyset \\ \emptyset & d_1 & d_2 \end{pmatrix}.$$

Now, suppose that a mechanism chooses μ under the above preference profile $\succ$. Then d_1 is unmatched. Consider reported preferences $\succ'_{d_1}$ of d_1,

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

Then μ' is a unique strongly stable matching, so the mechanism chooses μ' at $(\succ'_{d_1}, \succ_{-d_1})$. Doctor d_1 is better off at μ' than at μ since she is matched to h_2 at μ' while she is unmatched at μ. Hence, d_1 can profitably misreport her preferences when her true preferences are $\succ_{d_1}$.

If a mechanism chooses μ' under the above preference profile $\succ$, then by a symmetric argument, doctor d_2 can profitably misreport her preferences when her true preferences are $\succ_{d_2}$. Therefore there does not exist a mechanism that is strategy-proof for doctors and selects a strongly stable matching whenever there exists one. □

The above examples show that a strongly stable matching need not exist, and there exists no mechanism that is strategy-proof for doctors and selects a strongly stable matching whenever there exists one. These results suggest that the concept of strong stability is not appropriate as our desideratum.

Given that strong stability is "too strong" in the senses discussed above, a weaker stability notion that still takes regional caps into account is hoped for. Strong stability is *too* strong because *any* blocking pair is regarded as a legitimate deviation as long as it does not violate a regional cap. To define an appropriate stability concept, we need to further restrict blocking pairs that are regarded as legitimate. We do so by using the notion of target capacity. More specifically, we now regard target capacities $(\bar{q}_h)_{h \in H}$ as reflecting certain distributional goals (though not feasibility constraints) and define a stability concept that respects target capacities as much as possible.[22]

Definition 2. A matching μ is **stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then (i) $|\mu_{r(h)}| = q_{r(h)}$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and

(iii') either $\mu_d \notin H_{r(h)}$ or $|\mu'_h| - \bar{q}_h > |\mu'_{\mu_d}| - \bar{q}_{\mu_d}$,

where μ' is the matching such that $\mu'_d = h$ and $\mu'_{d'} = \mu_{d'}$ for all $d' \neq d$.

This concept is weaker than strong stability. Conditions (i) and (ii) in the definition of strong stability are also required in stability. Meanwhile stability is different from strong stability in that condition (iii) in strong stability is replaced by a condition (iii') and,

---

[22]Depending on the distributional goals, target capacities can be set differently from those specified in the description of the JRMP mechanism. Appendix B.4 discusses alternative ways to allocate target capacities.

since there are more possible cases in (iii') than in (iii), stability is weaker than strong stability.

The first part of condition (iii'), $\mu_d \notin H_{r(h)}$, is identical to condition (iii) and addresses the case in which the deviating doctor is currently assigned outside the region of the deviating hospital. The second part declares that certain types of blocking pairs within a region (note that $\mu_d \in H_{r(h)}$ holds in the remaining case) are not regarded as legitimate deviations (recall that our interpretation of stability concepts is normative). To see this point, consider the inequality in condition (iii'),

$$(4.1) \qquad\qquad |\mu'_h| - \bar{q}_h > |\mu'_{\mu_d}| - \bar{q}_{\mu_d}.$$

The left-hand side is the number of doctors matched to h in excess of its target $\bar{q}_h$ if d actually moves to h, realizing a new matching μ'. The right hand side is the number of doctors matched to the original hospital μ_d in excess of its target $\bar{q}_{\mu_d}$ if d moves out of μ_d. This property says that such a movement will not decrease the imbalance of over-target numbers of matching across hospitals. Intuitively, if the movement of the doctor in the blocking pair "equalizes" the excesses over the target capacities compared to the current matching (that is, $|\mu_h| - \bar{q}_h < |\mu'_h| - \bar{q}_h \leq |\mu'_{\mu_d}| - \bar{q}_{\mu_d} < |\mu_{\mu_d}| - \bar{q}_{\mu_d}$), then such a movement should be regarded as a legitimate deviation. Thus, the only blocking pair within a region that can remain under this definition should satisfy condition (4.1).

We note that there may be other natural definitions of stability. For example, it may be desirable to entitle a hospital with capacity 20 to twice as many doctors over the target as a hospital with capacity 10. There may also be other criteria that are deemed desirable, including even cases in which target capacities are not defined. To address this issue, in Section 6.5 and Appendix B we consider a class of stability concepts that includes the stability in Definition 2 as a special case and accommodates the above ideas.[23] For each stability notion in this class, we present a mechanism that generates a stable matching. In the main part of this paper, we assume that the policy goal is expressed as in condition (4.1). This particular policy goal is chosen here because it is expositionally simple and appears to be a reasonable starting point. However, it is not a necessary requirement for our analysis to work, as we will observe in Section 6.5 and Appendix B. The choice of a particular variant of stability should be in part the product of society's preferences. We restrict ourselves to proposing solutions that are flexible enough to meet as wide a range

---

[23]In Appendix I we consider a stability concept stronger than the stability concepts in this class (while weaker than strong stability) and show that this concept suffers from the same types of drawbacks (as in Examples 2 and 3) as those for strong stability.

of policy goals as possible. See Appendix B.3 for a partial list of other possible social preferences that we can accommodate.

Stability implies the following desirable property:

Theorem 1. *Any stable matching is (constrained) efficient.*

When there is no regional cap (in which case stability reduces to the standard concept of stability), a matching is stable if and only if it is in the core, and any core outcome is efficient. Without regional caps, Theorem 1 follows straightforwardly from these facts. With regional caps, however, there is no obvious way to define an appropriate cooperative game or a core concept. Theorem 1 states that efficiency of stable matchings still holds in our model.[24]

Remark 2. Since the outcome of the JRMP mechanism in Example 1 is not efficient, Theorem 1 implies that it is not stable either. This is easy to check by inspection as well. As hinted in Remark 1, (d_9, h_2) is a blocking pair that does not satisfy condition (i) in the definition of stability (Definition 2). That is, matching d_9 to h_2 does not violate the regional cap. □

A natural question is whether a stable matching exists in every market. This question will be answered in the affirmative in the next section, where we propose an algorithm that always generates a stable matching.

5. A New Mechanism: Flexible Deferred Acceptance

We present a new mechanism that, for any given input, results in a stable matching. To do so, we first define the **flexible deferred acceptance algorithm:**

For each $r \in R$, specify an order of hospitals in region r: Denote $H_r = \{h_1, h_2, \ldots, h_{|H_r|}\}$ and order h_i earlier than h_j if $i < j$. Given this order, consider the following algorithm.

(1) Begin with an empty matching, that is, a matching μ such that $\mu_d = \emptyset$ for all $d \in D$.

(2) Choose a doctor d who is currently not tentatively matched to any hospital and who has not applied to all acceptable hospitals yet. If such a doctor does not exist, then terminate the algorithm.

---

[24]To overcome the above difficulty, the proof presented in Appendix C shows this result directly rather than associating stability to the core in a cooperative game. The proof is general in the sense that it shows the (constrained) efficiency of "weak stability," the notion introduced in that section, which is weaker than stability.

(3) Let d apply to the most preferred hospital $\bar{h}$ at $\succ_d$ among the hospitals that have not rejected d so far. Let r be the region such that $\bar{h} \in H_r$.

(4) (a) For each $h \in H_r$, let D_h' be the entire set of doctors who have applied to but have not been rejected by h so far and are acceptable to h. For each hospital $h \in H_r$, choose the $\bar{q}_h$ best doctors according to $\succ_h$ from D_h' if they exist, and otherwise choose all doctors in D_h'. Formally, for each $h \in H_r$ choose D_h'' such that $D_h'' \subset D_h'$, $|D_h''| = \min\{\bar{q}_h, |D_h'|\}$, and $d \succ_h d'$ for any $d \in D_h''$ and $d' \in D_h' \setminus D_h''$.

   (b) Start with a tentative match D_h'' for each hospital $h \in H_r$. Hospitals take turns to choose (one doctor at a time) the best remaining doctor in their current applicant pool. Repeat the procedure (starting with h_1, proceeding to $h_2, h_3, \ldots$ and going back to h_1 after the last hospital) until the regional quota q_r is filled or the capacity of the hospital is filled or no doctor remains to be matched. All other applicants are rejected.[25]

We define the **flexible deferred acceptance mechanism** to be a mechanism that produces, for each input, the matching at the termination of the above algorithm.[26]

The flexible deferred acceptance mechanism is analogous to the deferred acceptance mechanism and the JRMP mechanism. What distinguishes the flexible deferred acceptance mechanism from the JRMP mechanism is that the former lets hospitals fill their capacities "more flexibly" than the latter. To see this point, first observe that the way that hospitals choose doctors who applied in Step 4a is essentially identical to the one in the JRMP algorithm. As seen before, the JRMP may result in an inefficient and unstable matching because this step does not let hospitals tentatively keep doctors beyond target capacities even if regional caps are not binding. This is addressed in Step 4b. In this step, hospitals in a region are allowed to keep more doctors than their target capacities

---

[25]Formally, let $\iota_i = 0$ for all $i \in \{1, 2, \ldots, |H_r|\}$. Let $i = 1$.

 (i) If either the number of doctors already chosen by the region r as a whole equals q_r, or $\iota_i = 1$, then reject the doctors who were not chosen throughout this step and go back to Step 2.

 (ii) Otherwise, let h_i choose the most preferred (acceptable) doctor in D_{h_i}' at $\succ_{h_i}$ among the doctors that have not been chosen by h_i so far, if such a doctor exists and the number of doctors chosen by h_i so far is strictly smaller than q_{h_i}.

 (iii) If no new doctor was chosen at Step 4(b)ii, then set $\iota_i = 1$. If a new doctor was chosen at Step 4(b)ii, then set $\iota_j = 0$ for all $j \in \{1, 2, \ldots, |H_r|\}$. If $i < |H_r|$ then increment i by one and if $i = |H_r|$ then set i to be 1 and go back to Step 4(b)i.

[26]We show in Theorem 2 that the algorithm stops in a finite number of steps.

if doing so keeps the regional caps respected. Thus there is a sense in which this algorithm corrects the deficiency of the JRMP mechanism while following closely the deferred acceptance algorithm.

In the flexible deferred acceptance algorithm, one needs to specify an ordering of hospitals. There are at least two reasons that this requirement may not cause problems such as conflicts among hospitals to get a "desirable position" in the order. The first is that, as we will discuss in Subsection 6.6, the effect of different ways of setting orders on the welfare of hospitals is ambiguous. More specifically, it may be the case that a hospital is better off being ordered later under some specification of preference profiles, while the opposite may be true under other specifications. Second, the flexible deferred acceptance algorithm can be modified without losing its desirable properties, by adding "Step 0" in which a particular ordering is chosen according to some probabilistic rule. The aforementioned problems can be resolved by, for example, choosing an order according to the uniform probability distribution.

The following example illustrates how the flexible deferred acceptance algorithm works.

Example 4 (The flexible deferred acceptance algorithm). Consider the same example as in Example 1. Recall that the JRMP mechanism can produce an inefficient and unstable matching. By contrast, the flexible deferred acceptance algorithm selects a matching that is efficient and stable. Precisely, let doctors apply to hospitals in the specified order. For doctors d_1 to d_8, the algorithm does not proceed to Step 4b, as the number of doctors in each hospital is no larger than its target. When d_9 applies, doctors $d_1, \ldots, d_8$ are still matched to hospitals in Step 4a, and d_9 is matched to h_2 in Step 4b. In the same way, when d_{10} applies, doctors $d_1, \ldots, d_8$ are still matched to hospitals in Step 4a, and d_9 and d_{10} are matched to h_2 in Step 4b. Hence an efficient and stable matching results. Intuitively, the algorithm treats doctors' applications in a more flexible manner than in the JRMP algorithm. This is the idea behind the name "*flexible* deferred acceptance."   □

The following is the main result of this section.

Theorem 2. *The flexible deferred acceptance algorithm stops in finite steps. The mechanism produces a stable matching for any input and is group strategy-proof for doctors.*

To see an intuition for the stability of the flexible deferred acceptance mechanism, recall that there is a sense in which hospitals fill their capacities "flexibly." More specifically, at each step of the algorithm hospitals can tentatively accept doctors beyond their target capacities as long as the regional cap is not violated. Then the kind of rejection that causes

instability in Example 1 does not occur in the flexible deferred acceptance algorithm.[27] Thus an acceptable doctor is rejected from a preferred hospital either because there are enough better doctors in that hospital, or the regional quota is filled by other doctors. So such a doctor cannot form a blocking pair, suggesting that the resulting matching is stable.[28]

The intuition for strategy-proofness for doctors is similar to the one for the deferred acceptance mechanism. A doctor does not need to give up trying for her first choice because, even if she is rejected, she will be able to apply to her second choice, and so forth. In other words, the "deferred" acceptance guarantees that she will be treated equally if she applies to a position later than others.

Although the above are rough intuitions of the results, the formal proof presented in Appendix B takes a different approach. It relates our model to the model of "(many-to-many) matching with contracts" (Hatfield and Milgrom, 2005). The basic idea of the proof is to regard each region as a consortium of hospitals that acts as one agent, and to define its choice function that selects a subset from any given collection of pairs (contracts) of a doctor and a hospital in the region. Once we successfully connect our model to the matching model with contracts, properties of the latter model can be invoked to show the theorem. In fact, the proof shows that a more general result (Theorem 4) holds which can be applicable to the class of stability concepts mentioned in Section 6.5 and that the current model is indeed a special case of the general model (Propositions 8 and 9). Theorem 2 then follows as a corollary of these results.

Theorems 1 and 2 imply the following appealing welfare property of the flexible deferred acceptance mechanism.

Corollary 1. *The flexible deferred acceptance mechanism produces an efficient matching for any input.*

Recall that the JRMP mechanism does not necessarily produce an efficient matching. In light of this observation, Corollary 1 implies that the flexible deferred acceptance mechanism has a better efficiency property than the JRMP mechanism.

The next two propositions formalize the idea that the flexible deferred acceptance mechanism respects target capacities and regional caps as much as possible.

---

[27]Indeed, in the market of Example 1, the result of the flexible deferred acceptance mechanism matches 10 doctors as in matching μ'. Thus in this example the flexible deferred acceptance mechanism strictly improves upon the JRMP mechanism.

[28]Since hospitals take turns one by one when they tentatively accept doctors, no blocking pair involving a doctor's movement within a region can "equalize" the distribution of doctors in the region.

Proposition 1. *If the number of doctors matched with $h \in H$ in the flexible deferred acceptance mechanism is strictly less than its target capacity, then for any $d \in D$ who are not matched with h, either d is unacceptable to h or d prefers its current match to h.*

In other words, the flexible deferred acceptance mechanism does not prevent a doctor from being matched to an underserved hospital, relative to the target capacity, in the name of respecting the regional caps.[29] This result suggests, as we argued informally when defining stability, that the choice of target capacities can be utilized as a means to achieve distributional goals.

Proposition 2.   (1) *If the number of doctors matched with $h \in H$ in the flexible deferred acceptance mechanism is strictly less than its target capacity, then the set of doctors matched with h under the (unconstrained) deferred acceptance mechanism is a subset of the one under the flexible deferred acceptance mechanism.*

(2) *If the number of doctors matched in $r \in R$ in the flexible deferred acceptance mechanism is strictly less than its regional cap, then each hospital h in r weakly prefers a matching produced by the flexible deferred acceptance mechanism to the one under the (unconstrained) deferred acceptance mechanism. Moreover, the number of doctors matched to any such h in the former matching is weakly larger than that in the latter.*

This result implies that, whenever a hospital or a region is underserved under the flexible deferred acceptance mechanism, the (unconstrained) deferred acceptance mechanism cannot improve the match at such a hospital or a region. This result offers a sense in which the flexible deferred acceptance mechanism avoids inflicting costs on underserved hospitals or regions.

6. Discussion

This section provides several discussions. Subsection 6.1 studies mechanisms in other contexts and countries, and shows that those existing mechanisms suffer from problems similar to those we pointed out for Japanese medical residency match. Subsection 6.2 considers an alternative mechanism that is often suggested to us, and shows that it is not strategy-proof for doctors. In Subsection 6.3, we consider the rural hospital theorem of Roth (1986) and a related concept of the "match rate," the ratio of the number of all matched doctors to the total number of doctors (matched plus unmatched). Subsection

---

[29]The conclusion of the theorem applies even if the regional cap is already binding, thus this property is not implied by the fact that the outcome of the flexible deferred acceptance algorithm is stable.

6.4 studies the existence issue of a side-optimal stable matching, that is, a matching that is preferred by all doctors or by all hospitals. Subsection 6.5 generalizes stability and the flexible deferred acceptance mechanism. Subsection 6.6 examines the welfare effect of different choices of picking orders over hospitals, target capacities, and regional caps, and Subsection 6.7 considers "floor constraints" instead of "ceiling constraints" (regional caps).

6.1. **Mechanisms in China and the United Kingdom.** In the main sections we showed that the flexible deferred acceptance mechanism has desirable properties such as efficiency, stability, and strategy-proofness for doctors, and we observed that it outperforms the JRMP mechanism. In this section we analyze several other mechanisms used in practice, by formulating them and pointing out their respective deficiencies. The analyses confirm the applicability of our mechanism in various markets around the world.

Specifically, in what follows we analyze the Chinese Graduate school admission and the medical matching problem in the United Kingdom. More comprehensive descriptions of the problems and the analyses can be found in Appendix A. Appendix A also discusses more examples mentioned in the Introduction such as college admission in Ukraine and matching of new teachers in Scotland.

6.1.1. *Chinese Graduate School Admission.* The first problem we study is the Chinese graduate school admission. As described briefly in the Introduction, master's programs are categorized as either academic or professional, and the Chinese government is currently trying to reduce the number of academic master students. To achieve this goal, the government decided to reduce the available seats of each academic master's program by about 25 percent by 2015.[30] In our framework, the policy goal of the Chinese government can be translated into imposing the "regional cap" on the set of all academic master's programs, where the regional cap is about 75 percent of the sum of the true capacities across all the academic programs. Then the government sets the target capacity of each academic master's program at about 75 percent of the true capacity.

Given the target capacities above, the main round of Chinese graduate admission runs as follows. Using an application website, each student applies to one graduate program. Given the set of applicants, each graduate school accepts its most preferred acceptable students up to its target capacity and rejects everyone else. All matches are final.[31]

---

[30]To achieve this goal gradually, the government plans to reduce the number of seats by about 5 percent every year until 2015.

[31]Here we are describing the main round of the admission process. There is a "guaranteed assignment" in which especially high-achieving students are admitted before the main round begins, and there is also

This mechanism suffers from several drawbacks. First, it is easy to see that this mechanism is not strategy-proof for students. Moreover, the mechanism may produce an unstable and inefficient matching. Here we focus on a particular source of inefficiency and instability that shares a certain feature with the JRMP mechanism: In the Chinese admission mechanism, the target capacities are used as rigid constraints, so the seats of a certain academic program beyond its target capacity *must* remain unfilled, even if some students prefer to be matched to these seats. This property holds true whatever mechanism the Chinese government uses as long as the target capacities are treated as rigid constraints, so the problem is orthogonal to the inefficiency coming from the fact that each student can list only one program. In fact, we show in Appendix A.2 that even a (complete information) pure-strategy subgame-perfect equilibrium outcome (of a game in which students submit their preference list first and then colleges admit students) can be unstable and inefficient, while we also show that the outcome is always stable and efficient if there is no binding regional cap (so that the target capacity of a program is equal to its physical capacity).

6.1.2. *United Kingdom.* As mentioned briefly in the Introduction, the mechanism used for medical match in the United Kingdom is based on a different idea from the mechanisms we have discussed so far. The process has two rounds, in which doctors are matched to a region first, and then to a program within their matched region. For a policy maker who desires to control the distribution of doctors across different regions, assigning doctors using such a two-round scheme may appear to be an appealing alternative to the JRMP mechanism or the Chinese mechanism. As we will see in Appendix A.4, however, this mechanism may also result in inefficiency and instability, and it entails incentive problems as well.

There have been several changes of the mechanism in recent years, and the matching mechanism in the second round varies across regions. To be specific, however, assume that both rounds use the serial dictatorship. Serial dictatorship is in use in the first round since 2012, and it is in use in the second round in Scotland since 2010. For simplicity we focus on this mechanism although the same points can be made in other mechanisms as well.[32]

---

a supplementary round called an "adjustment process" which is largely decentralized. We do not discuss these processes in the main text because they are not directly relevant to the topic of distributional constraints. See instead Appendix A.2.1 for detail of these rounds.

[32]This point will be shown more formally in Appendix A.4.2, in which another two-round mechanism is analyzed.

Suppose that a doctor whose first choice is a hospital h_1 in a region r_1, the second choice is h_2 in another region r_2, and her third choice is h_3 in region r_1. Assume that, in the first round, the doctor lists r_1 as her first choice and is matched to it. However, it is possible that she is matched to h_3 in the second round, while h_2 in region r_2 prefers her to one of the doctors matched to it. Then this doctor and h_2 form a blocking pair that is not tolerated under our stability concept (or even the weak stability concept as defined in Appendix C), implying that the resulting matching is not (weakly) stable.

Intuitively, instability can happen for the following reason: In the first round a student may apply to and is matched to a region where her preferred hospitals are located. But then in the second round she may end up being matched to a hospital that she prefers less to a hospital in another region which prefers her to one of its matched doctors. In other words, since the matching between doctors and regions are finalized before the ultimate matching to a hospital is decided, the resulting matching could result in instability. A similar example shows that the matching can be inefficient. Moreover, under this mechanism there does not generally exist a dominant strategy for doctors because a doctor's best report in the first round depends on which hospital in the region she will end up with. Concrete examples making these points are found in Appendix A.4.

Scotland's matching between new teachers and schools uses a similar two-round matching mechanism. As in the U.K. medical match, the matching clearinghouse first matches teachers to a local authority, which then assigns teachers matched to it to schools under its control. As such, this mechanism also suffers from similar instability, inefficiency, and strategic problems. See Appendix A.4 for detail.

6.2. **The Iterated Deferred Acceptance Mechanism.** As a solution to the efficiency and stability problems in the JRMP mechanism, we often encounter suggestions by government officials and matching theorists, saying that the **iterated deferred acceptance (iterated DA)** mechanism that uses the following algorithm may be useful: This algorithm consists of finite steps of rounds. In round 1, the deferred acceptance algorithm is run regarding the target capacities as the real capacities. If the resulting matching fills all the target capacities, then the algorithm stops. Otherwise, the algorithm proceeds to round 2 after the target capacities are modified as follows: hospitals set their new target capacities equal to their matched numbers of doctors if they have vacant seats relative to their target capacities; these vacant seats are reallocated to other hospitals in the same region according to a certain pre-specified rule. In round 2, the deferred acceptance algorithm is run with these modified target capacities. If the resulting matching fills all the new target capacities then the algorithm stops and otherwise it continues. We do the same

in all other rounds, with a restriction that once a hospital has reduced its target capacity then it never increases (and require that the algorithm stop if no further reallocation is possible).

As one might expect, this mechanism produces a (strongly) stable matching in Example 1 . However it turns out that this mechanism is not strategy-proof for doctors.

Example 5. Consider a market with two doctors, d_1 and d_2, and two hospitals h_1 and h_2 in a single region with regional cap 2. Each doctor prefers h_1 to h_2 to being unmatched. Each hospital is associated with a capacity of 2 and a target capacity of 1, and prefers d_1 to d_2 to being unmatched. In this market, the iterated DA ends in one round, resulting in the matching

$$\mu = \begin{pmatrix} h_1 & h_2 \\ d_1 & d_2 \end{pmatrix}.$$

Doctor d_2 has an incentive to misreport her preferences. For, if she reports that she prefers h_1 to being unmatched to h_2, then the iterated DA proceeds to the second round with one seat moving from h_2 to h_1, and in the second round the matching

$$\mu' = \begin{pmatrix} h_1 & h_2 \\ d_1, d_2 & \emptyset \end{pmatrix},$$

is realized and the algorithm stops. Since d_2 prefers μ'_{d_2} to μ_{d_2}, the iterated DA mechanism is not strategy-proof for doctors.   □

6.3. **The Rural Hospital Theorem and The Match Rate.** In this subsection, we show that the conclusion of the rural hospital theorem does not hold in our environment. Motivated by this finding, we study how the flexible deferred acceptance mechanism works in terms of the match rate, that is, the proportion of the number of all matched doctors to the total number of doctors (matched plus unmatched).

6.3.1. *The Rural Hospital Theorem.* The rural hospital theorem (Roth, 1986) states that, in a matching model without regional caps, any hospital that fails to fill all its positions in one stable matching is matched to an identical set of doctors in all stable matchings. It also states that the set of unmatched doctors is identical across all stable matchings.

The theorem is of particular interest when we consider allocating a sufficient number of doctors to rural areas. Although the rural hospital theorem might suggest that increasing the number of doctors in a particular set of hospitals is impossible, the conclusion of the theorem does not necessarily hold in our context with regional caps, even with the most stringent concept of strong stability. The following example makes this point clear.

Example 6 (The conclusion of the rural hospital theorem does not hold)**.** There is one region r with regional cap $q_r = 1$, in which two hospitals, h_1 and h_2, reside. Each hospital h has a capacity of $q_h = 1$. Suppose that there are two doctors, d_1 and d_2. We assume the following preferences:

$$\succ_{h_1}: d_1, \quad \succ_{h_2}: d_2,$$
$$\succ_{d_1}: h_1, \quad \succ_{d_2}: h_2.$$

It is straightforward to check that there are two strongly stable matchings,

$$\mu = \begin{pmatrix} h_1 & h_2 & \emptyset \\ d_1 & \emptyset & d_2 \end{pmatrix},$$

$$\mu' = \begin{pmatrix} h_1 & h_2 & \emptyset \\ \emptyset & d_2 & d_1 \end{pmatrix}.$$

Notice that hospital h_1 fills its capacity in matching μ while it does not do so in matching μ'. Also, d_1 is matched to a hospital in matching μ while unmatched in matching μ'. Hence both conclusions of the rural hospital theorem fail, even with the notion of strong stability. Since strong stability implies stability, this example also shows that the conclusions of the rural hospital theorem fail with stability (analogously, all negative conclusions of this subsection and the next hold under both stability and strong stability). $\qquad \square$

One might suspect that, although the rural hospital theorem does not apply, it might be the case that each region attracts the same number of doctors in any strongly stable matchings. The following example shows that this is not true.

Example 7 (The number of doctors matched to hospitals in a rural region may be different in different strongly stable matchings)**.** We modify Example 6 by adding one more region r', which we interpret here as a "rural region" for the sake of discussion. Region r' has the regional cap of $q_{r'} = 1$, and one hospital, h_3, resides in it. Suppose that h_3 has a capacity of $q_{h_3} = 1$. The preferences are modified as follows:

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

It is straightforward to check that there are two strongly stable matchings,

$$\mu = \begin{pmatrix} h_1 & h_2 & h_3 & \emptyset \\ d_1 & \emptyset & \emptyset & d_2 \end{pmatrix},$$

$$\mu' = \begin{pmatrix} h_1 & h_2 & h_3 \\ \emptyset & d_2 & d_1 \end{pmatrix}.$$

Thus the hospital in rural region r' does not attract any doctors in matching μ, while it attracts one doctor in matching μ'.   □

Hence, when the number of doctors matched to hospitals in rural regions matters, the choice of a mechanism *is* an important issue, in the presence of regional caps.

6.3.2. *The Match Rate.* Related to the rural hospital theorem is the notion of "match rate," which is the ratio of the number of all matched doctors to the total number of doctors (matched plus unmatched). The match rate seems to be a measure that many people care about. For example, match rates are listed on the annual reports published by the NRMP and the JRMP.[33] This is perhaps because the match rate is an easy measure for participants to understand.[34]

Although it would be desirable to select a matching that has the maximum match rate among the stable matchings, the following example shows that the flexible deferred acceptance mechanism fails to do so.

Example 8 (The flexible deferred acceptance mechanism does not necessarily select a matching with the highest match rate among stable matchings). Take the same example as in Example 7. Also, let the target profile be $(\bar{q}_{h_1}, \bar{q}_{h_2}, \bar{q}_{h_3}) = (1, 0, 1)$. Then, the flexible deferred acceptance mechanism always selects a matching μ defined in Example 7. But this has a match rate of $1/2$, while the other matching, namely μ' defined in Example 7, has a match rate of 1.   □

It is unfortunate that the flexible deferred acceptance mechanism does not necessarily maximize the match rate within stable matchings, but the following example shows that this is a necessary consequence of requiring stratregy-proofness for doctors.

Example 9 (No mechanism that is strategy-proof for doctors can always select a matching with the highest match rate among stable matchings). Modify the environment in Example 7 as follows:

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

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

---

[33]For instance, see National Resident Matching Market (2010) and Japan Residency Matching Program (2009b).

[34]The ease of understanding may not be a persuasive reason for economic theorists to care about the match rates, but it seems to be a crucial issue for market designers. For a mechanism to work well in practice, it is essential that people are willing to participate in the mechanism. To this end, providing information in an accessible manner, as in the form of the match rates, seems to be of great importance.

with everything else unchanged (thus hospitals h_1 and h_2 are in one region and h_3 is in the other, each region has a regional cap of one, and each hospital has capacity of one). Let $(\bar{q}_{h_1}, \bar{q}_{h_2}, \bar{q}_{h_3}) = (1, 0, 1)$. Notice that, given these preferences, there are two stable matchings, namely μ with $\mu_{d_1} = h_1$ and $\mu_{d_2} = h_3$, and μ' with $\mu'_{d_1} = h_3$ and $\mu'_{d_2} = h_2$. Take a mechanism that always selects a matching with the highest match rate among the stable matchings. We show that this mechanism cannot be strategy-proof. Since both μ and μ' have match rate of 1, both can potentially be chosen by the mechanism. Suppose that the mechanism chooses μ. Then, doctor d_2 has an incentive to misreport her preferences: If she reports that hospital h_2 is the only acceptable match, then given the new profile of the preferences, the only stable matching that maximizes the match rate among stable matchings is μ'. Since $\mu'_{d_2} \succ_{d_2} \mu_{d_2}$, doctor d_2 indeed has an incentive to misreport. A symmetric argument can be made for the case in which the mechanism chooses μ' given the true preference profile. Hence, there does not exist a mechanism that is strategy-proof for doctors and always selects a matching with the highest match rate among stable matchings. $\qquad\square$

Despite the above negative results, there are bounds on the match rates in the matchings produced by the flexible deferred acceptance mechanism. More specifically, the following comparison can be made with the JRMP mechanism as well as with the (unconstrained) deferred acceptance algorithm without regional caps:

Theorem 3. *For any preference profile,*

(1) *Each doctor $d \in D$ weakly prefers a matching produced by the deferred acceptance mechanism to the one produced by the flexible deferred acceptance mechanism to the one produced by the JRMP mechanism.*

(2) *If a doctor is unmatched in the deferred acceptance mechanism, she is unmatched in the flexible deferred acceptance mechanism. If a doctor is unmatched in the flexible deferred acceptance mechanism, she is unmatched in the JRMP mechanism.*

Notice that part (2) of the above result, which is a direct corollary of part (1), implies that the match rate is weakly higher in the deferred acceptance mechanism than in the flexible deferred acceptance mechanism, which in turn has a weakly higher match rate than the JRMP mechanism.[35]

---

[35]For an example in which the deferred acceptance mechanism and the flexible deferred acceptance mechanism differ in terms of match rates, see Example 2 (with an arbitrary target capacity profile). For the flexible deferred acceptance mechanism and the JRMP mechanism, see Example 1.

In Appendix E, we present simulations results to compare the performances of the deferred acceptance mechanism, the JRMP mechanism, and the flexible deferred acceptance mechanism. The effects highlighted by the above theorem are quite substantial. For instance, almost 600 additional doctors become unmatched under the JRMP mechanism compared to the unconstrained deferred acceptance mechanism (1396 versus 805), but this figure is reduced to almost 200 (1010 versus 805) if the flexible deferred acceptance mechanism is used instead of the JRMP mechanism, while satisfying all the regional caps just like the JRMP mechanism does. Moreover, almost 1000 doctors–about 12 percent of the total number of doctors– prefer the outcomes of the flexible deferred acceptance mechanism *strictly* to the JRMP outcomes, while every doctor prefers the former *weakly* to the latter, as predicted by Theorem 3. See Appendix E for detail of these results, as well as our simulation methods and other results.

Theorem 3 suggests that the flexible deferred acceptance mechanism matches reasonably many doctors. Characterizing stable mechanisms that achieve strategy-proofness for doctors and match "as many doctors as possible," as well as studying their relationship with the flexible deferred acceptance mechanism, is an interesting open question.

6.4. **Nonexistence of Side-Optimal Stable Matchings.** There does not necessarily exist a doctor-optimal stable matching (a stable matching unanimously preferred to every stable matching by all doctors). Neither does there exist a hospital-optimal stable matching. To see this point, consider the environment presented in Example 6, and suppose that $(\bar{q}_{h_1}, \bar{q}_{h_2}) = (1, 0)$. There are two stable matchings, μ and μ' specified in Example 6, where only d_1 and h_1 are matched at μ while only d_2 and h_2 are matched at μ'. Clearly, d_1 and h_1 strictly prefer μ to μ' while d_2 and h_2 strictly prefer μ' to μ. Thus there exists neither a doctor-optimal stable matching nor a hospital-optimal stable matching. Moreover, this example shows that there exists neither a doctor-pessimal stable matching nor a hospital-pessimal stable matching in general.

6.5. **Generalizations.** As mentioned in Section 4, the notion of stability is based on the idea that if the result of a move of a doctor within a region does not equalize the excesses over the target capacities compared to the current matching, it is not deemed as a legitimate deviation. We argued that this is not the only reasonable definition as, for example, it may be natural to suppose that a hospital with capacity 20 is entitled to twice as many doctors (over the target) as a hospital with capacity 10. There may be other criteria, and the Appendix B explores the extent to which our analysis goes through. More specifically, we present a model in which each region is endowed with "regional preferences" over the set of distributions of doctors within the region. One special case of

the regional preferences is when the region prefers to have more equal number of doctors in excess of targets. We define a stability concept that takes the regional preferences into consideration. We provide a condition on regional preferences under which a generalized version of the flexible deferred acceptance algorithm finds a stable matching as defined more generally, and it is group strategy-proof. The criteria mentioned so far satisfy our condition.

Appendix F provides a further generalization: we consider the situation where there is a hierarchy of regional caps. We show that a generalization of the flexible deferred acceptance mechanism induces a stable matching appropriately defined. This generalization not only has a theoretical appeal but also is practically important. For instance, one could consider a hierarchy of regional caps, say one cap for a prefecture and one for each district within the prefecture. Or the policy maker may desire to regulate the total number of doctors practicing in each specialty in each prefecture.

6.6. **Welfare Effects of Picking Orders, Targets, and Regional Caps.** The flexible deferred acceptance algorithm follows a certain picking order of hospitals in each region when there are some doctors remaining to be tentatively matched after hospitals have kept doctors up to their target capacities. One issue is how to decide the picking order. One natural conjecture may be that choosing earlier (that is, having an earlier order in the flexible deferred acceptance algorithm) benefits a hospital. As we have mentioned earlier, this would be a problematic property: If choosing earlier benefits a hospital, then how to order hospitals will be a sensitive policy issue to cope with because each hospital would have incentives to be granted an early picking order. Fortunately, the conjecture is not true, as shown in the following example. The example also shows that the different choices of orders result in different stable matchings, thus the choice of an order does matter for the algorithm's outcome.

Example 10 (Ordering a hospital earlier may make it worse off). Let there be hospitals h_1, h_2 and h_3 in region r_1, and h_4 in region r_2. Suppose that $(q_{h_1}, q_{h_2}, q_{h_3}, q_{h_4}) = (2, 1, 1, 1)$ and $(\bar{q}_{h_1}, \bar{q}_{h_2}, \bar{q}_{h_3}, \bar{q}_{h_4}) = (1, 0, 1, 1)$. The regional cap of r_1 is 2 and that for r_2 is 1. Preferences are

$$\succ_{h_1}: d_1, d_4, d_2, \quad \succ_{h_2}: d_3, \quad \succ_{h_3}: \text{arbitrary}, \quad \succ_{h_4}: d_2, d_1,$$

$$\succ_{d_1}: h_4, h_1, \quad \succ_{d_2}: h_1, h_4, \quad \succ_{d_3}: h_2, \quad \succ_{d_4}: h_1.$$

(1) Assume that h_1 is ordered earlier than h_2. In that case, in the flexible deferred acceptance mechanism, d_1 applies to h_4, d_2 and d_4 apply to h_1, and d_3 applies to h_2. d_2 and d_4 are accepted while d_3 is rejected. The matching finalizes.

(2) Assume that h_1 is ordered after h_2. In that case, in the flexible deferred acceptance mechanism, d_1 applies to h_4, d_2 and d_4 apply to h_1, and d_3 applies to h_2. But now d_2 is rejected while d_3 is accepted. Then d_2 applies to h_4, displacing d_1 from h_4. Then d_1 applies to h_1. d_1 is accepted, displacing d_4 from h_1. The matching finalizes.

First, notice that hospital h_2 is better off in case (2). Thus being ordered earlier helps h_2 in this example. However, if h_1 prefers $\{d_1\}$ to $\{d_2, d_4\}$ (which is consistent with the assumption that hospital preferences are responsive with capacities), then h_1 is also made better off in case (2). Thus being ordered later helps h_1 if she prefers $\{d_1\}$ to $\{d_2, d_4\}$. Therefore, the effect of a picking order on hospitals' welfare is not monotone.   □

A related concern is about what could be called "target monotonicity." That is, keeping everything else constant, does an increase of the target of a hospital make it better off under the flexible deferred acceptance mechanism? If so, then hospitals would have strong incentives to influence policy makers to give them large targets. The following example shows that target monotonicity is not necessarily true.

Example 11 (Target monotonicity may fail)**.** Consider a market that is identical to the one in Example 10, except that the target of h_1 is now decreased to 0, with the order such that h_1 chooses before h_2.[36] Then h_1 is matched to $\{d_1\}$ under the flexible deferred acceptance mechanism. Therefore, if h_1 prefers $\{d_1\}$ to $\{d_2, d_4\}$, then h_1 is made better off when its target capacity is smaller.   □

Note that both hospital and doctor preferences are heterogeneous in Examples 10 and 11. However, similar failures can occur even when hospitals or doctors have homogeneous preferences. Moreover, splitting or merging regions also has ambiguous welfare effects. These points are made by examples in Supplementary Appendix H.

By contrast, there exist natural comparative statics results regarding welfare effects of the regional caps.

Proposition 3. *Fix a picking order in the flexible deferred acceptance mechanism. Let $(q_r)_{r \in R}$ and $(q'_r)_{r \in R}$ be regional caps such that $q'_r \leq q_r$ for each $r \in R$. Then the following statements hold.*

---

[36]When the target capacity of h_1 is decreased, the sum of the target capacities becomes strictly smaller than the regional cap (note that such a situation is allowed in our model). If one wishes to keep the sum equal to the regional cap, the example can be modified by increasing the target capacity of h_3 by 1, and the conclusion of the example continues to hold.

(1) *Each doctor $d \in D$ weakly prefers a matching produced by the flexible deferred acceptance mechanism under regional caps $(q_r)_{r \in R}$ to the one under $(q'_r)_{r \in R}$.*

(2) *For each region r such that $q_r = q'_r$, the number of doctors matched in r at a matching produced by the flexible deferred acceptance mechanism under regional caps $(q'_r)_{r \in R}$ is weakly larger than at the matching under $(q_r)_{r \in R}$.*

Thus all doctors are made weakly worse off when the regional caps become more stringent. Meanwhile, the number of doctors matched in a region whose regional cap is unchanged weakly increases when the regional caps of other regions become more stringent. This result highlights the tradeoff that a policy maker faces in using the flexible deferred acceptance mechanism: If the regional caps of urban regions are reduced, then the number of doctors matched to other regions weakly increases. However this change weakly decreases the welfare of doctors.

Remark 3. We obtain Proposition 3 as a corollary of a general comparative statics result that we prove in the Appendix D (Lemma 1). This result can be useful in analyzing matching with distributional constraints. For example, the following comparative statics about the JRMP mechanism can be shown using this result.

Proposition 4. *Let $(\bar{q}_h)_{h \in H}$ and $(\bar{q}'_h)_{h \in H}$ be target capacities such that $\bar{q}'_h \leq \bar{q}_h$ for each $h \in H$. Then the following statements hold.*[37]

(1) *Each doctor $d \in D$ weakly prefers a matching produced by the JRMP mechanism under target capacities $(\bar{q}_h)_{h \in H}$ to the one under $(\bar{q}'_h)_{h \in H}$.*

(2) *Each hospital $h \in H$ such that $\bar{q}_h = \bar{q}'_h$ weakly prefers a matching produced by the JRMP mechanism under target capacities $(\bar{q}'_h)_{h \in H}$ to the one under $(\bar{q}_h)_{h \in H}$. Moreover, the number of doctors matched to any such h in the former matching is weakly larger than that in the latter.*

□

6.7. **Floor Constraints.** The present paper offers a practical solution for the Japanese resident matching problem with regional caps. However, the regional cap may not be an ultimate objective per se, but a means to allocate medical residents "evenly" to different areas. Setting a cap –a ceiling constraint on the number of residents in a region– is an obvious approach to this desideratum, but there may be other possible regulations. For example, one might wonder if setting floor constraints, as opposed to cap constraints, would

---

[37] Since the JRMP mechanism is equivalent to the deferred acceptance mechanism with respect to the target capacities, this result can also be obtained by appealing to the "Capacity Lemma" by Konishi and Ünver (2006), although we obtain these results as corollaries of a more general result, Lemma 1.

be an easier and more direct solution. However, there are reasons that floor constraints may be difficult to use. First, even the existence of an individually rational matching that respects floor constraints is not guaranteed. For example, if no doctor finds any hospital in a certain region to be acceptable, then satisfying a positive floor constraint for the region results in an individually irrational matching (doctors matched with hospitals in the region would just reject taking the job). Second, even if an individually rational matching exists, it is not clear whether a stable matching exists. In fact, an appropriate definition of stability in the presence of floor constraints is unclear.[38]

7. Related literature

In the one-to-one matching setting, McVitie and Wilson (1970) show that a doctor or a hospital that is unmatched at one stable matching is unmatched in every stable matching. This is the first statement of the rural hospital theorem to our knowledge, and its variants and extensions have been established in increasingly general settings by Gale and Sotomayor (1985a,b), Roth (1984, 1986), Martinez, Masso, Neme, and Oviedo (2000), and Hatfield and Milgrom (2005), among others. As recent results are quite general, it seems that placing more doctors in rural areas has been believed to be a difficult (if not impossible) task, and thus there are few studies offering solutions to this problem. The current paper explores possible ways to offer some positive results.

Roth (1991) points out that some hospitals in the United Kingdom prefer to hire no more than one female doctor while offering multiple positions. Similarly, some schools (or school districts) desire certain diversity characteristics of their incoming classes such as ethnicity and academic performance (Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu, 2005; Ergin and Sönmez, 2006). Westkamp (2010) considers a college admission problem in which colleges have admission criteria based on trait-specific quotas. If one regards a region (instead of a hospital) as a single agent in our model, these models and ours appear similar in that an agent in both models has certain "preferences" over distributions more complex than responsive ones. However, the above models are different from ours. For instance, in our model, a distinction should be made between a matching of a doctor to one hospital in a region and a matching of the same doctor to a different hospital in the same region, but such a distinction cannot be even described in the former models. This distinction is essential in the context of residency matching because a doctor may have

---

[38]Similar points are made in the context of school choice by Ehlers (2010), Ehlers, Hafalir, Yenmez, and Yildirim (2011), and Fragiadakis, Iwasaki, Troyan, Ueda, and Yokoo (2012).

incentives to deviate by moving between hospitals within a single region. Thus results from these papers cannot be applied in this paper's environment.

Despite the above-mentioned difficulty, there is a way to make an association between our model and an existing model, namely the model of matching with contracts as defined by Hatfield and Milgrom (2005).[39] More specifically, given a matching market with regional caps, one can define an associated matching model with contracts such that a stable allocation in the latter model induces a stable matching in the former. This correspondence allows us to show some of our results by using properties of the matching with contracts model established by Hatfield and Milgrom (2005), Hatfield and Kojima (2009, 2010), and Hatfield and Kominers (2009, 2010).[40] On the other hand, it is also worth noting that these models are still different. The reason is that certain types of blocks allowed in the matching model with contracts are considered infeasible in our context. Thus stable allocations in a matching model with contracts can induce only a subset of stable matchings in our model. For this reason, the structural properties of the set of stable matchings in our model are strikingly different from those in the matching model with contracts. For instance, a doctor-optimal stable allocation exists and the conclusion of the rural hospital theorem holds in their model but not in ours.[41]

Abraham, Irving, and Manlove (2007) study allocation of students to projects where a lecturer may offer multiple projects. Both projects and lecturers have capacity constraints. Sönmez and Ünver (2006) analyze a related model in the context of school choice in which there may be multiple school programs in a school building. Motivated by the matching system for higher education in Hungary, Biró, Fleiner, Irving, and Manlove (2010) extend these models to cases in which capacity constraints are imposed on a nested system of sets. Their models are analogous to ours if we associate a lecturer and a project – and a school building and a school, respectively– in their models to a region and a hospital in

---

[39]Fleiner (2003) considers a framework that generalizes various mathematical results. A special case of his model corresponds to the model of Hatfield and Milgrom (2005), although not all results of the latter (e.g., those concerning incentives) are obtained in the former. See also Crawford and Knoer (1981) who observe that wages can represent general job descriptions in their model, given their assumption that firm preferences satisfy separability.

[40]Note that residency matching and school choice with balance requirements mentioned in the last paragraph (Roth, 1991; Abdulkadiroğlu and Sönmez, 2003) can be modeled as special cases of this paper's model. A related issue appears in the National Resident Matching Program where a hospital may have multiple types of residency positions (Roth and Peranson, 1999; Niederle, 2007).

[41]More specifically, the former result holds under the property called the substitute condition, and the latter under the substitute condition and another property called the law of aggregate demand or size (or cardinal) monotonicity (Alkan, 2002; Alkan and Gale, 2003).

our model, respectively. However, there are two notable differences. First, they assume that preferences of all projects provided by the same lecturer (school programs in the same building) are identical while such a restriction is not imposed in our model.[42] Second, the stability concepts employed in their models are different from ours, thus our results do not reduce to theirs even in their more specialized settings.

Milgrom (2009) and Budish, Che, Kojima, and Milgrom (2010) consider object allocation mechanisms with restrictions similar to the regional caps in our model. While their models are independent of ours (most notably, their analysis is primarily about object allocation, and stability is not studied), they share motivations with ours in that they consider flexible assignment in the face of complex constraints.

More broadly, this paper is part of a rapidly growing literature on matching market design. As advocated by Roth (2002), much of recent market design theory advanced by tackling problems arising in practical markets.[43] For instance, practical considerations in designing school choice mechanisms in Boston and New York City are discussed by Abdulkadiroğlu, Pathak, and Roth (2005, 2009) and Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005, 2006). Abdulkadiroğlu, Che, and Yasuda (2008, 2009), Erdil and Ergin (2008), and Kesten (2009) analyze alternative mechanisms that may produce more efficient student placements than those that are currently used in New York City and Boston. Design issues motivated by an anti-trust lawsuit against the American medical resident matching clearinghouse are investigated by Bulow and Levin (2006), Kojima (2007), Konishi and Sapozhnikov (2008), Niederle (2007), and Niederle and Roth (2003). A classical resource allocation problem with multi-unit demand has attracted renewed attention in the context of practical course allocation at business schools as studied by Sönmez and Ünver (2010), Budish and Cantillon (2010), and Budish (2010). Initiated by Roth, Sönmez, and Ünver (2004, 2005, 2007), even the organ transplantation problem has become a subject of market design research in recent years. See Roth and Sotomayor (1990) for a comprehensive survey of the matching literature in the first three decades, and Roth (2007a) and Sönmez and Ünver (2008) for discussion of more recent studies.

8. Conclusion

We showed that the current matching mechanisms used in various contexts around the world may result in avoidable inefficiency and instability even though some of them are

---

[42]In our context, it is important to allow different hospitals in the same region to have different preferences because two hospitals rarely have identical preferences in practice.

[43]Literature on auction market design also emphasizes the importance of solving practical problems (see Milgrom (2000, 2004) for instance).

similar to the celebrated deferred acceptance mechanism. We proposed a new mechanism, called the flexible deferred acceptance mechanism. This mechanism is (group) strategy-proof and generates a stable and efficient matching.

With regional caps, defining stability is not a trivial task, and it seems that the right notion depends on the welfare and distributional goals that the policy maker wants to achieve. Hence there may not necessarily exist a unique choice of the mechanism, and there is room for the policy maker to select a particular stable matching based on such goals. We hope that this paper serves as a basis for achieving such goals and, more broadly, that it contributes to the general agenda of matching/market design theory to address specific issues arising in practical problems.

We intentionally refrained from judging the merits of imposing regional caps itself (except for certain welfare results mentioned below). We took this approach because our model does not explicitly include patients or ethical concerns of the general populace, which may be underlying arguments for increasing doctors in rural areas. Similarly, we did not analyze other policies such as subsidies to incentivize residents to work in rural areas.[44] Instead, we took an approach in the new tradition of market design research, in which one regards constraints such as fairness and repugnance as requirements to be respected and offers solutions consistent with them.[45] That is, as regional caps seem to be a strong political reality, we believe that it is important to take them as given and provide a practical solution. To help the policy maker make informed judgements about the tradeoffs involved in imposing regional caps, we provided a number of comparative statics results.

The paper opens new avenues for further research topics. First, as mentioned before, strategy-proofness for every agent including hospitals is impossible even without regional caps if we also require stability. However, truthtelling is an approximately optimal strategy even for hospitals under the deferred acceptance mechanism in large markets under some assumptions (Roth and Peranson, 1999; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009). Although such an analysis requires a much more specialized model

---

[44]This is not because subsidies are not important. In fact, subsidies are used to attract residents to rural areas in many countries such as the United States and Japan. However, there are political pressures to restrict the use of subsidies in the Japanese medical market. Beginning in 2011, for instance, the government will reduce subsidies to residency programs that pay annual salaries of more than 7,200,000 yen (about 85,000 U.S. dollars) to residents. In any case, our analysis is applicable given participants' preferences which reflect subsidies, thus our method can be employed on top of subsidies.

[45]This approach is eloquently advocated by Roth (2007b).

structure than what this paper has and is outside the scope of this paper, approximate incentive compatibility similar to these papers may hold.

Second, studying more general constraint structures may be interesting. As mentioned in Appendix F, we analyzed the case in which there is a hierarchy of regional caps, and showed that a stable matching can be found by a generalization of our flexible deferred acceptance mechanism. By contrast, we show by example that if the regions do not form a hierarchy, a stable matching does not necessarily exist.[46] A general recipe for defining a stability concept and finding a stable matching by an algorithm is an open question.

Third, it would be desirable to quantify the effect of using the flexible deferred acceptance mechanism instead of the existing mechanisms. As briefly discussed in Section 6.3, we conducted simulations to compare the outcome of the flexible deferred acceptance mechanism with that of the JRMP mechanism and the deferred acceptance mechanism. Since the data of submitted preferences have been unavailable to us so far, we used randomly generated preferences (while using publicly available data to mimic several aspects of the real market). However, a better prediction would be possible if we could simulate the performance of our mechanism based on actual data of preferences. In a new project joint with Jun Wako, we have started talking with the matching organizers to discuss such issues and put our mechanism in real use.

Finally, it would be nice to study markets that have similar structures to the ones in this paper. We explored a wide range of applications both in terms of geography (such as Japan, China, the United Kingdom, Hungary, and Ukraine) and in terms of the context (such as medical match, teacher allocation, college admission, and graduate school admission). We expect some general insights will carry over to other applications, while market-specific details may need to be carefully taken into account when we consider different markets in different political or cultural environments.

References

Abdulkadiroğlu, A. (2005): "College admissions with affirmative action," *International Journal of Game Theory*, 33(4), 535–549.

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2008): "Expanding Choice in School Choice," Economic Research Initiatives at Duke Research Paper No. 20.

——— (2009): "Resolving Conflicting Preferences in School Choice: the Boston Mechanism Reconsidered," *American Economic Review*, forthcoming.

---

[46]Situations with non-hierarchies may be relevant if, for instance, a government would like to impose caps based on prefectures and specialities independently.

ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2005): "The New York City High School Match," *American Economic Review Papers and Proceedings*, 95, 364–367.

ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match," *American Economic Review*, forthcoming.

ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *American Economic Review*, 99, 1954–1978.

ABDULKADIROĞLU, A., P. A. PATHAK, A. E. ROTH, AND T. SÖNMEZ (2005): "The Boston Public School Match," *American Economic Review Papers and Proceedings*, 95, 368–372.

——— (2006): "Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access," Harvard University and Boston College, unpublished mimeo.

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

ABRAHAM, D. J., R. IRVING, AND D. MANLOVE (2007): "Two algorithms for the student-project allocation problem," *Journal of Discrete Algorithms*, 5, 73–90.

ALCOBA, N. (2009): "Hitting the road to alleviate India's rural doctor shortage," *the Canadian Medical Association Journal*,
http://www.cmaj.ca/cgi/reprint/180/10/E34.pdf.

ALKAN, A. (2002): "A class of multipartner matching markets with a strong lattice structure," *Economic Theory*, 19(4), 737–746.

ALKAN, A., AND D. GALE (2003): "Stable Schedule Matching Under Revealed Preferences," *Journal of Economic Theory*, 84, 73–94.

ASSOCIATION OF MEDICAL STUDENTS (2009): "Public comment on the proposed revision to the residency match (in Japanese)," http://students.umin.jp/.

AYGÜN, O., AND T. SÖNMEZ (2012): "Matching with Contracts: The Critical Role of Irrelevance of Rejected Contracts," Discussion paper, Boston College Department of Economics.

BIRÓ, P., T. FLEINER, R. W. IRVING, AND D. F. MANLOVE (2010): "The College Admissions problem with lower and common quotas," *Theoretical Computer Science*, 411, 3136–3153.

BUDISH, E. (2010): "The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes," *Harvard University, Unpublished mimeo*.

BUDISH, E., AND E. CANTILLON (2010): "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," Discussion paper, mimeo.

BUDISH, E., Y.-K. CHE, F. KOJIMA, AND P. R. MILGROM (2010): "Designing Random Allocation Mechanisms: Theory and Applications," forthcoming, *American Economic Review*.

BULOW, J., AND J. LEVIN (2006): "Matching and Price Competition," *American Economic Review*, 96, 652–668.

CRAWFORD, V. (1991): "Comparative statics in matching markets," *Journal of Economic Theory*, 54(2), 389–400.

CRAWFORD, V., AND E. M. KNOER (1981): "Job Matching with Heterogeneous Firms and Workers," *Econometrica*, 49, 437–450.

DUBINS, L. E., AND D. A. FREEDMAN (1981): "Machiavelli and the Gale-Shapley algorithm," *American Mathematical Monthly*, 88, 485–494.

ECHENIQUE, F., AND J. OVIEDO (2006): "A Theory of Stability in Many-to-Many Matching," *Theoretical Economics*, 1, 233–273.

EHLERS, L. (2010): "Controlled School Choice," Unpublished mimeo.

EHLERS, L., I. HAFALIR, M. YENMEZ, AND M. YILDIRIM (2011): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," *mimeo*.

ERDIL, A., AND H. ERGIN (2008): "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice," *American Economic Review*, 98, 669–689.

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

ERGIN, H., AND T. SÖNMEZ (2006): "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 90, 215–237.

FLEINER, T. (2003): "A Fixed-Point Approach to Stable Matchings and Some Applications," *Mathematics of Operations Research*, 28, 103–126.

FRAGIADAKIS, D., A. IWASAKI, P. TROYAN, S. UEDA, AND M. YOKOO (2012): "Strategyproof Matching with Minimum Quotas," *mimeo*.

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

GALE, D., AND M. A. O. SOTOMAYOR (1985a): "Ms Machiavelli and the Stable matching problem," *American Mathematical Monthly*, 92, 261–268.

——— (1985b): "Some Remarks on the Stable Matching Problem," *Discrete Applied Mathematics*, 11, 223–232.

HAERINGER, G., AND V. IEHLE (2010): "Knocking out the supply and sorting in centralized job markets," mimeo.

HATFIELD, J. W., AND F. KOJIMA (2009): "Group Incentive Compatibility for Matching with Contracts," *Games and Economic Behavior*, pp. 745–749.

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

HATFIELD, J. W., AND S. D. KOMINERS (2009): "Contract Design and Stability in Matching Markets," mimeo.

——— (2010): "Matching in Networks with Bilateral Contracts," mimeo.

HATFIELD, J. W., AND P. MILGROM (2005): "Matching with Contracts," *American Economic Review*, 95, 913–935.

IMMORLICA, N., AND M. MAHDIAN (2005): "Marriage, Honesty, and Stability," *SODA*, pp. 53–62.

IRVING, R., AND D. MANLOVE (2009): "Finding large stable matchings," *ACM Journal of Experimental Algorithmics*, 14, Section 1 article no. 2.

JAPAN RESIDENCY MATCHING PROGRAM (2009a): "Result of the Residency Match," http://www.jrmp.jp/toukei/2009/2009toukei-5.pdf.

——— (2009b): "Result of the Residency Match," http://www.jrmp.jp/koho/2009/2009press.htm.

——— (2010): "Rules for Matching Participation (residency hospitals and university hospitals)," http://www.jrmp.jp/kiyaku-byouin.htm.

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

KESTEN, O. (2009): "School Choice with Consent," *Quarterly Journal of Economics*, forthcoming.

KISELGOF, S. (2012): "Matching Practices for Universities–Ukraine," http://www.matching-in-practice.eu/index.php/matching-in-practice/universities/ukraine.

KOJIMA, F. (2007): "Matching and price competition: comment," *The American Economic Review*, 97(3), 1027–1031.

KOJIMA, F., AND M. MANEA (2009): "Axioms for Deferred Acceptance," forthcoming, *Econometrica*.

KOJIMA, F., AND P. PATHAK (2009): "Incentives and stability in large two-sided matching markets," *American Economic Review*, 99(3), 608–627.

KOMINERS, S., AND T. SÖNMEZ (2012): "Designing for Diversity: Matching with Slot-Specific Priorities," *Boston College and University of Chicago working paper.*

KONISHI, H., AND M. SAPOZHNIKOV (2008): "Decentralized matching markets with endogenous salaries," *Games and Economic Behavior*, 64(1), 193–218.

KONISHI, H., AND M. U. ÜNVER (2006): "Games of Capacity Manipulation in the Hospital-Intern Market," *Social Choice and Welfare*, 27, 3–24.

KOZU, T. (2006): "Medical education in Japan," *Academic Medicine*, 81(12), 1069–1075.

LIN, H. (2011): "Conference Minutes for Ministry of Education," http://www.mpacc.net.cn/ReadNews.asp?NewsID=1689&BigClassName=%CA%D4%B5%E3%D4%BA%D0%A3&SmallClassName=%D4%BA%D0%A3%B6%AF%CC%AC&SpecialID=34.

MARTINEZ, R., J. MASSO, A. NEME, AND J. OVIEDO (2000): "Single agents and the set of many-to-one stable matchings," *Journal of Economic Theory*, 91, 91–105.

——— (2004): "On group strategy-proof mechanisms for a many-to-one matching model," *International Journal of Game Theory*, 33, 115–128.

MCVITIE, D. G., AND L. WILSON (1970): "Stable Marriage Assignments for Unequal Sets," *BIT*, 10, 295–309.

MILGROM, P. (2000): "Putting Auction Theory to Work: The Simulteneous Ascending Auction," *Journal of Political Economy*, 108(2), 245–272.

——— (2004): *Putting Auction Theory to Work.* Cambridge University Press, Cambridge.

——— (2009): "Assignment messages and exchanges," *American Economic Journal: Microeconomics*, 1(2), 95–113.

MINISTRY OF EDUCATION OF CHINA (2010): "Notification about Master's Program Enrollment of the year 2010 (in Chinese)," http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s180/201007/xxgk_91621.html.

MINISTRY OF HEALTH, LABOUR AND WELFARE (2005a): "Proposed survey on no-doctor districts (in Japanese)," http://www.mhlw.go.jp/shingi/2005/01/s0124-11c.html.

——— (2005b): "Report on no-doctor districts (in Japanese)," http://www.mhlw.go.jp/toukei/list/76-16.html.

——— (2009a): "On opinions expressed in the public comment procedure (in Japanese)," http://www.mhlw.go.jp/shingi/2009/04/dl/s0423-8a.pdf.

——— (2009b): "On the outcome of the residency match (in Japanese)," http://www.mhlw.go.jp/topics/bukyoku/isei/rinsyo/matching/dl/091030-1a.pdf.

NAMBIAR, M., AND J. BAVAS (2010): "Rudd plan won't fix rural doctor shortage: RDA," http://www.abc.net.au/news/stories/2010/03/16/2846827.htm.

National Resident Matching Market (2010): "Results and Data, 2010 Main Residency Match," http://www.nrmp.org/data/resultsanddata2010.pdf.

Nicholson, S. (2003): "Barriers to Entering Medical Specialties," NBER Working Paper 9649.

Niederle, M. (2007): "Competitive wages in a match with ordered contracts," *The American Economic Review*, 97(5), 1957–1969.

Niederle, M., and A. Roth (2003): "Relationship between wages and presence of a match in medical fellowships," *JAMA*, 290(9), 1153.

Onishi, H., and I. Yoshida (2004): "Rapid change in Japanese medical education," *Medical Teacher*, 26(5), 403–408.

Pathak, P., and T. Sonmez (2008): "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism," *The American Economic Review*, 98(4), 1636–1652.

People's Republic of China (2010): "National Long-Term Plan for Education Reform and Development (2010-2020) (in Chinese)," http://www.gov.cn/jrzg/2010-07/29/content_1667143.htm.

Pronovost, P., D. Angus, T. Dorman, K. Robinson, T. Dremsizov, and T. Young (2002): "Physician staffing patterns and clinical outcomes in critically ill patients: a systematic review," *JAMA, the Journal of the American Medical Association*, 288(17), 2151.

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

——— (1984): "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 92, 991–1016.

——— (1985): "The College Admission Problem is not Equivalent to the Marriage Problem," *Journal of Economic Theory*, 36, 277–288.

——— (1986): "On the allocation of residents to rural hospitals: a general property of two-sided matching markets," *Econometrica*, 54(2), 425–427.

——— (1991): "A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians and Surgeons in the U.K.," *American Economic Review*, 81, 415–440.

——— (2002): "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics, Fisher-Schultz Lecture," *Econometrica*, 70, 1341–1378.

——— (2007a): "Deferred Acceptance Algorithms: History, Theory, Practice and Open Questions," forthcoming, *International Journal of Game Theory*.

——— (2007b): "Repugnance as a Constraint on Markets," *The Journal of Economic Perspectives*, pp. 37–58.

ROTH, A. E., AND E. PERANSON (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 748–780.

ROTH, A. E., T. SÖNMEZ, AND U. ÜNVER (2004): "Kidney Exchange," *Quarterly Journal of Economics*, 119, 457–488.

——— (2005): "Pairwise Kidney Exchange," *Journal of Economic Theory*, 125, 151–188.

——— (2007): "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review*, 97, 828–851.

ROTH, A. E., AND M. A. O. SOTOMAYOR (1990): *Two-sided Matching: a Study in Game-theoretic Modeling and Analysis*. Econometric Society monographs, Cambridge.

ROTH, A. E., AND X. XING (1994): "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions," *American Economic Review*, 84, 992–1044.

——— (1997): "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists," *Journal of Political Economy*, 105, 284–329.

SHALLCROSS, T. (2005): "General Medicine in a Remote & Rural Hospital: Problems and Proposals – a personal view," http://www.rcpe.ac.uk/publications/rarm/shalcross.pdf.

SÖNMEZ, T., AND M. ÜNVER (2006): "School Choice with Aggregate Capacity Constraints," mimeo.

——— (2008): "Matching, Allocation, and Exchange of Discrete Resources," *Handbook of Social Economics, Elseview, forthcoming*.

——— (2010): "Course bidding at business schools," *International Economic Review*, 51(1), 99–123.

SOTOMAYOR, M. A. O. (2004): "Implementation in the many-to-many matching market," *Games and Economic Behavior*, 46(1), 199–212.

TALBOTT, C. (2007): "Shortage of Doctors Affects Rural U.S.," *the Washington Post*, http://www.washingtonpost.com/wp-dyn/content/article/2007/07/21/AR2007072100432.html.

TEO, A. (2007): "The current state of medical education in Japan: a system under reform," *Medical education*, 41(3), 302–308.

THOMSON, W. (2003): "Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: a survey," *Mathematical Social Sciences*, 45(3), 249–297.

TOTTORI PREFECTURE (2009): "Emergency request to the proposed revision to the residency match by four prefectural governors (in Japanese)," http://www.pref.tottori.lg.jp/dd.aspx?menuid=102291.

TOYABE, S. (2009): "Trend in geographic distribution of physicians in Japan," *International Journal for Equity in Health*, 8(1), 5.

WESTKAMP, A. (2010): "An Analysis of the German University Admissions System," mimeo.

WONGRUANG, P. (2010): "Diagnosis bleak for rural doctors," *the Bangkok Post*, http://www.bangkokpost.com/news/investigation/34445/diagnosis-bleak-for-rural-doctors.

ZHAI, Y. (2011): "Development of Professional Education is Left Behind," *Chinese News of Science*, http://edu.qq.com/a/20111117/000077.htm.

APPENDIX

TABLE OF CONTENTS: SECTIONS F TO I ARE RELEGATED TO THE ONLINE SUPPLEMENTARY APPENDIX.

Appendix A. Matching Markets in Practice

A.1. **Residency Matching in Japan.** In Japan, about 8,000 doctors and 1,500 residency programs participate in the matching process each year. This section describes how this process has evolved and how it has affected the debate on the geographical distribution of residents. For further details of Japanese medical education written in English, see Teo (2007) and Kozu (2006). Also, information about the matching program written in Japanese is available at the websites of the government ministry and the matching organizer.[47]

The Japanese residency matching started in 2003 as part of a comprehensive reform of the medical residency program. Prior to the reform, clinical departments in university hospitals, called *ikyoku*, had de facto authority to allocate doctors. The system was criticized because it was seen to have given clinical departments too much power and resulted in opaque, inefficient, and unfair allocations of doctors against their will.[48] Describing the situation, Onishi and Yoshida (2004) write "This clinical-department-centred system was often compared to the feudal hierarchy."

To cope with the above problem a new system, the Japan Residency Matching Program (JRMP), introduced a centralized matching procedure using the (doctor-proposing) deferred acceptance algorithm by Gale and Shapley (1962). Unlike its U.S. counterpart, the National Resident Matching Program (NRMP), the system has no "match variation" (Roth and Peranson, 1999) such as married couples, which would cause many of the good properties of the deferred acceptance algorithm to fail.

Although the matching system was welcomed by many, it has also received a lot of criticisms. This is because some hospitals, especially university hospitals in rural areas, felt that they attracted fewer residents under the new matching mechanism. They argued that the new system provided too much opportunity for doctors to work for urban hospitals rather than rural hospitals, resulting in severe doctor shortages in rural areas. While there is no conclusive evidence supporting their claim, an empirical study by Toyabe (2009) finds that the geographical imbalance of doctors has increased in recent years according to several measures (the Gini coefficient, Atkinson index, and Theil index of the per-capita number of doctors across regions). By contrast, he also finds that the

---

[47]See the websites of the Ministry of Health, Labor and Welfare (http://www.mhlw.go.jp/topics/bukyoku/isei/rinsyo/) and the Japan Residency Matching Program (http://www.jrmp.jp/).

[48]The criticism appears to have some justification. For instance, Niederle and Roth (2003) offer empirical evidence that a system without a centralized matching procedure reduces mobility and efficiency of resident allocation in the context of the U.S. gastroenterologist match.

imbalance is lower when residents are excluded from the calculation. Based on these findings, he suggests that the matching system introduced in 2003 may have contributed to the widening regional imbalance of doctors.

To put such criticisms into context, we note that the regional imbalance of doctors has been a long-standing and serious problem in Japan. As of 2004, there were over 160,000 people living in the so-called *mui-chiku*, which means "districts with no doctors" (Ministry of Health, Labour and Welfare, 2005b)[49] and many more who were allegedly underserved. One government official told one of the authors (personal communication) that the regional imbalance is one of the most important problems in the government's health care policy, together with financing health care cost. Popular media regularly report stories of doctor shortages, often in a very sensational tone.[50] There is evidence that the sufficient staffing of doctors in hospitals is positively correlated with the quality of medical care such as lower mortality (see Pronovost, Angus, Dorman, Robinson, Dremsizov, and Young (2002) for instance); thus the doctor shortage in rural areas may lead to bad medical care.

In response to the criticisms against the matching mechanism, the Japanese government introduced a new system with regional caps beginning with the matching conducted in 2009. More specifically, a regional cap was imposed on the number of residents in each of the 47 prefectures that partition the country. If the sum of the hospital capacities in a region exceeds its regional cap, then the capacity of each hospital is reduced to equalize the total capacity with the regional cap.[51] Then the deferred acceptance algorithm is implemented under the reduced capacities. We call this mechanism the Japan Residency Matching Program (JRMP) mechanism. The basic intuition behind this policy is that if

---

[49]A mui-chiku is defined by various criteria such as the ease of access to hospitals, the population, the regularity of clinic openings, and so forth (Ministry of Health, Labour and Welfare, 2005a).

[50]For instance, the Yomiuri Shimbun newspaper, with circulation of over 10,000,000, recently provoked a controversy by its article about the only doctor in Kamikoani-mura village, where 2,800 people live (Yomiuri Shimbun newspaper, 03/19/2010). Although the doctor, aged 65, took only 18 days off a year, she was persistently criticized by some "unreasonable demanding" patients. When she announced that she wanted to quit (which means that the village will be left with no doctor) because she was "exhausted," 600 signatures were collected in only 10 days, to change her mind.

[51]The capacity of a hospital is reduced proportionately to its original capacity in principle (subject to integrality constraints) although there are a number of fine adjustments and exceptions. This rule might suggest that hospitals have incentives to misreport their true capacities, but in Japan, the government regulates how many positions each hospital can offer so that the capacity can be considered exogenous. More specifically, the government decides the physical capacity of a hospital based on verifiable information such as the number of beds in it.

residents are denied from urban hospitals because of the reduced capacities, then some of them will work for rural hospitals.
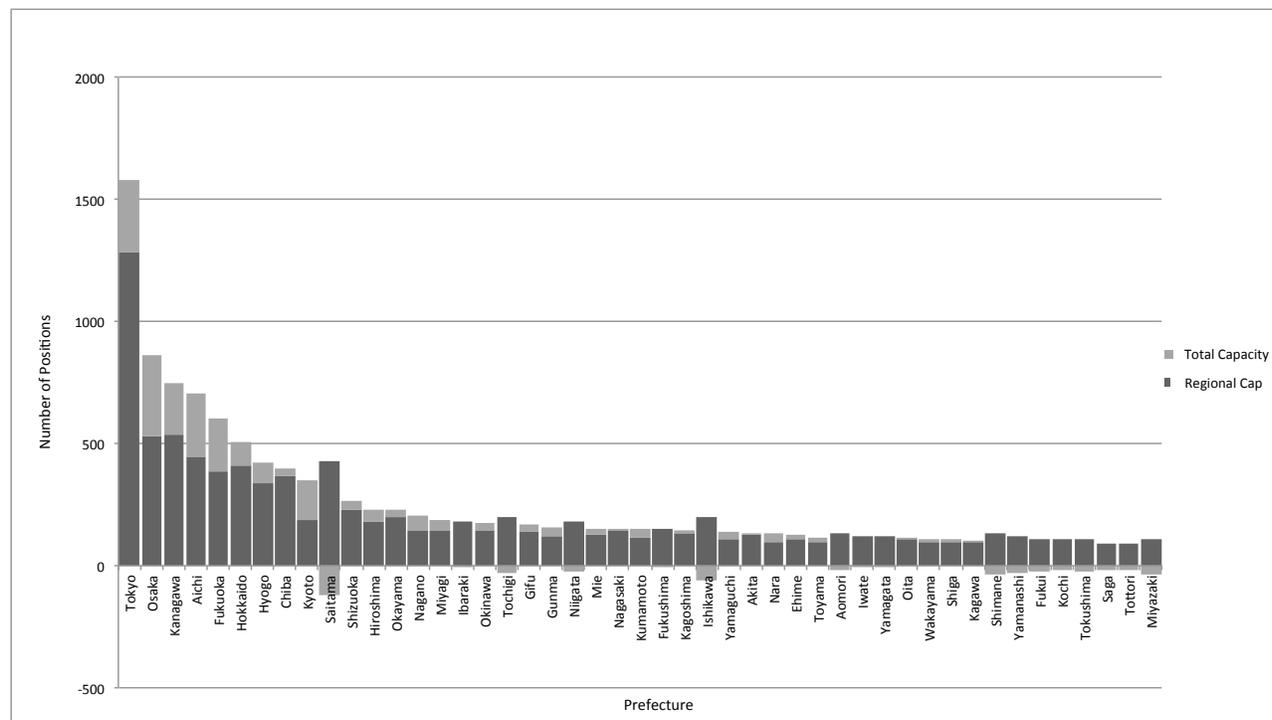

FIGURE 1.   For each prefecture, the total capacity is the sum of advertised positions in hospitals located in the prefecture in 2008. The regional caps are based on the government's plan in 2008 (Ministry of Health, Labour and Welfare, 2009a). Negative values of total capacities in some prefectures indicate the excess amount of regional caps beyond the advertised positions.

The magnitude of the regional caps is illustrated in Figure 1. Relatively large reductions are imposed on urban areas. For instance, hospitals in Tokyo and Osaka advertised 1,582 and 860 positions in 2008, respectively, but the government set the regional caps of 1,287 and 533, the largest reductions in the number of positions. The largest reduction in proportion is imposed on Kyoto, which offered 353 positions in 2008 but the number is dropped to 190, a reduction of about 46 percent. Indeed, the projected changes were so large that the government provided a temporary measure that limits per-year reductions within a certain bound in the first years of operation, though the plan is to reach the planned regional cap eventually. In total, 34 out of 47 prefectures are given regional caps smaller than the numbers of advertised positions in 2008.

The new JRMP mechanism with regional caps was used in 2009 for the first time. The government claims that the change alleviated the regional imbalance of residents:

It reports that the proportion of residents matched to hospitals in rural areas has risen to 52.3 percent, an increase of one percentage point from the previous year (Ministry of Health, Labour and Welfare, 2009b).[52] However, there is mounting criticism to the JRMP mechanism as well. For instance, a number of governors of rural prefectures (see Tottori Prefecture (2009) for instance) and a student group (Association of Medical Students, 2009) have demanded that the government modify or abolish the JRMP mechanism with regional caps.[53] Among other things, a commonly expressed concern is that the current system with regional caps causes efficiency losses, for instance by preventing residents from learning their desired skills for practicing medical treatments.

In the main text we formalized the JRMP mechanism (Section 3), explored its properties (Example 1, Remark 1, and Proposition 4), and compared it to the flexible deferred acceptance mechanism (Theorem 3). Our analysis suggests that the current JRMP mechanism needs to be changed to the flexible deferred acceptance mechanism.

A.2. **Chinese Graduate Admission.** This section describes the Chinese graduate admission in detail, and formally shows that the mechanism may result in an unstable and inefficient matching.[54]

A.2.1. *Institutional Background.* Chinese society is changing rapidly, and it is widely believed that there is need for more workers with professional master's degrees.[55] However, professional master's degrees have traditionally been regarded as inferior to academic master's degrees by many.[56] And there are not as many students in professional master's programs as the government aims.[57]

---

[52]Ministry of Health, Labour and Welfare (2009b) defines "rural areas" as all prefectures except for 6 prefectures, Tokyo, Kyoto, Osaka, Kanagawa, Aichi, and Fukuoka, which have large cities.

[53]Interestingly, even regional governments in rural areas such as Tokushima and Tottori were opposed to the JRMP mechanism. They were worried that since the system reduces capacities of *each hospital* in the region, some of which could hire more residents, it can reduce the number of residents allocated in the regions even further. This feature - inflexibility of the way capacities are reduced - is one of the problems of the current JRMP mechanism, which we try to remedy by our alternative mechanism.

[54]We greatly benefited from discussing Chinese graduate school admission with Jin Chen.

[55]Ministry of Education of China (2010) states that "[the education authority and graduate schools] should put emphasis on the promotion of education for advanced professionals, especially full-time professional master's degree."

[56]Chinese government has been emphasizing that the only differences between professional and academic master's programs are in the types and goals of education, and not in standards, but the reputation of a professional master's degree is still not as good as an academic one (Zhai, 2011).

[57]China's master-level education system traditionally emphasized academic training, rather than professional training. However, most graduates from master's programs pursue a professional career instead

To address this issue, Chinese government started a new regulation to increase enrollment in professional master's programs in 2010 (People's Republic of China, 2010). More specifically, the government began to impose constraints on the total number of academic master students, while increasing the number of professional master students. To achieve this goal, the government decided to reduce the available seats of each academic master's program by about 25 percent by 2015, while increasing capacities for professional programs.[58]

Although Chinese graduate school admission is different from Japanese residency match in many ways, there is a clear isomorphism between the structures of the problems that these markets are faced with. Just as there is demand for increasing resident allocation in rural Japan, there is demand for increasing professional master students in Chinese graduate education. Moreover, both in Japanese and Chinese cases, the feasibility requirements are placed on the total numbers of allocations for a subset of institutions (hospitals in each prefecture in the Japanese case, and the academic master's programs in the Chinese case). Lastly, when implementing the requirement, both governments place rigid restrictions on the allowed seats of each institution (each hospital in Japan, and each graduate school in China).

Remark 4. As mentioned in Section 6.1.1 of the main text, the matching mechanism for Chinese graduate school admission has additional rounds. First, there is a recommendation-based admission for high achievers before the main round. Second, there is a round called an "adjustment process" for those who have not been matched by the end of the main round. We do not formally analyze these rounds because they are not directly related to the issue of distributional constraints, and processes similar to them and the associated problems have been analyzed by other works.[59] Nevertheless, we describe these rounds for completeness.

---

of an academic one: in 2009, for instance, enrollment for master's programs was around 415,000 while that for PhD programs was around 60,000 (in China, the master's degree is similar to M.Phil. in countries like the U.K. in that a student seeking PhD first attend a master's program). Between 2008 and 2011, the proportion of professional masters' enrollment has increased from just 7 percent to 30 percent of the total enrollment. Ministry of Education aims at 50 percent by 2015 (Lin, 2011).

[58]To achieve this goal gradually, the government plans to reduce the number of seats by about 5 percent every year until 2015 (to our knowledge, the government has not disclosed whether it will continue imposing the reduction beyond 2015).

[59]Abdulkadiroğlu, Pathak, and Roth (2005); Abdulkadiroğlu, Pathak, and Roth (2009) study New York City's high school match. In NYC, top 2 percent students are automatically admitted to certain schools if they prefer, similarly to top performers who can be admitted to some schools in China's recommendation-based admission. Roth and Xing (1994, 1997) study labor markets that proceed in real

In the recommendation-based admission, students are recommended to graduate schools directly even before the graduate entrance examination that everyone else should take for admission. This round works as a shortcut for excellent students to enter graduate schools. Most recommended students are admitted to and attend the same university's graduate school where he or she attended college, and few of the recommended students transfer to another graduate school. This round is decentralized, and how to evaluate and admit students is largely up to each school, and thus the admission policy varies from school to school.

Students who were unmatched in the recommendation-based admission and the main round, as well as graduate schools which were not full in these rounds, enter the adjustment process. This round proceeds in real time, and each student maintains an application list which consists of at most two schools at any moment during this process. More specifically, at the beginning of this round, each student enters at most two schools into an online system.[60] Each school sees students who listed it, decides whether to give each applicant the chance for interview or not, and sends out interview invitations to them. Upon receiving an interview invitation, a student chooses whether to accept it or not. If a student fails to receive an interview invitation in 48 hours or is rejected after the interview from a school, the student can remove the school and add another school into her online application list. Once changed, a student must keep the new school in the list for at least 48 hours unless it interviews her. On the interview day, a graduate school admits or rejects students. If the capacity becomes full, the graduate school completes the admission. Each student can confirm admission from at most one school. Once she confirms, she exits the matching process.

The adjustment process happens in real time: In 2012, for instance, the process was in session from April 1st to May 5th. Given this time constraint, it is widely believed that students and schools contact each other ahead of the official adjustment process.[61]   □

A.2.2. *Formal Analysis.* Let us use the same notation as in the main text, although now we call h a program instead of a hospital, and d a student instead of a doctor. Further assume that the set of all programs is partitioned into the set of all academic programs r

---

time, and highlight the time constraint and associated strategic behavior and inefficiency of the resulting matching.

[60]The application is maintained at http://yz.chsi.com.cn/.

[61]See "Experts: 2012 graduate school entrance examination, advice and strategies for adjustment" from *Kuakao Education*, a counseling agency in China for students applying for graduate schools, available at http://yz.chsi.com.cn/kyzx/fstj/201203/20120322/293594478.html.

and the set of all professional programs r'. Throughout, assume $q_{r'} > \sum_{h \in H_{r'}} q_h$ so that the cap for professional programs is not binding.

We define the main round of the Chinese graduate admission formally.[62] As described in the main text, given the cap q_r, the main round of Chinese graduate admission runs as follows (we describe the mechanism for a general value of q_r, although in China q_r is an integer close to 75 percent of the sum of academic program capacities). Set $\bar{q}_h \leq q_h$ for each h in such a way that $\sum_{h \in H_r} \bar{q}_h \leq q_r$ (in Chinese graduate admission, $\bar{q}_h$ is an integer that is at most 75 percent of q_h for each $h \in H_r$). Each student applies to at most one program. Given the set of applicants, each program h accepts its most preferred students up to its target capacity $\bar{q}_h$ and rejects everyone else. All matchings are final.

We model the behavior in this mechanism by considering the following two-stage extensive-form game. In the first stage, students simultaneously apply to programs, one for each student. Then in the second stage, each program admits students from those who applied to it up to its target capacity. In this game, the following result holds:

Result 1. *Suppose that $q_r > \sum_{h \in H_r} q_h$ and $\bar{q}_h = q_h$ for all h. Then the set of the pure-straetegy subgame-perfect equilibrium outcomes in the game induced by the Chinese graduate admission coincides with the set of stable matchings.*

Proof. When $q_r > \sum_{h \in H_r} q_h$ and $\bar{q}_h = q_h$ for all h, the stability concept of this paper is equivalent to the standard stability concept (as in Roth and Sotomayor (1990) for example). By Sotomayor (2004) and Echenique and Oviedo (2006), the set of subgame perfect equilibrium outcomes of this game is equivalent to the set of stable matchings in the standard sense. These two observations complete the proof.   □

Thus if there is no binding cap on academic programs, then the equilibrium outcomes are stable and hence efficient. When the cap on academic programs is binding as in the Chinese admission mechanism, however, neither of these good properties hold even for equilibrium outcomes. The following example, which is an adaptation of Example 1, is such a case.

Example 12 (Equilibrium Outocomes under the Chinese Mechanism Can Be Unstable and Inefficient)**.** The "regional cap" for academic programs r is $q_r = 10$. There are two academic programs h_1 and h_2 and no professional program. Each program h has a capacity

---

[62]The description is based on the website of the "National Graduate Admissions Information Network" (http://yz.chsi.com.cn/), which provides information on graduate admission and host online applications for graduate schools.

of $q_h = 10$. Let the target capacities be $\bar{q}_{h_1} = \bar{q}_{h_2} = 5$. There are 10 students, $d_1, \ldots, d_{10}$. Preference profile $\succ$ is as follows:

$$\succ_{h_i}: d_1, d_2, \ldots, d_{10} \text{ for } i = 1, 2,$$
$$\succ_{d_j}: h_1 \text{ if } j \leq 3 \quad \text{and} \quad \succ_{d_j}: h_2 \text{ if } j \geq 4.$$

It is easy to see that the only pure-strategy subgame perfect equilibrium outcome is

$$\mu = \begin{pmatrix} h_1 & h_2 & \emptyset \\ d_1, d_2, d_3 & d_4, d_5, d_6, d_7, d_8 & d_9, d_{10} \end{pmatrix}.$$

Consider a matching μ' defined by,

$$\mu' = \begin{pmatrix} h_1 & h_2 \\ d_1, d_2, d_3 & d_4, d_5, d_6, d_7, d_8, d_9, d_{10} \end{pmatrix}.$$

Since the cap for academic programs is still respected, μ' is feasible. Moreover, every student is weakly better off with students d_9 and d_{10} being strictly better off than at μ. Hence we conclude that the Chinese mechanism can result in an inefficient matching. We also note that μ is not stable: For example, program h_2 and student d_9 constitute a blocking pair while the cap for r is not binding. $\qquad\square$

A.3. **College Admission in Ukraine.** A problem similar to Japanese residency match and Chinese graduate school admission is found in college admission in Ukraine as well. In Ukraine, some of the seats are financed by the state, while other "open-enrollment" seats require that students pay tuition (Kiselgof, 2012).[63] There is a cap on the number of state-financed seats, apparently as there is a limit on the budget that can be used to finance college study. The government implements the cap on the number of state-financed seats by imposing a cap on each program as in Japanese residency match and the Chinese graduate admission. Although the specific mechanism the Ukrainian college admission system uses is different from JRMP and Chinese graduate school admissions (see Kiselgof (2012) for detail), instability and inefficiency because of the constraints similarly result.

A.4. **Medical Matching in the United Kingdom.**

---

[63]A similar cap on the number of state-financed college seats exists in Hungarian college admission as well. The situation is somewhat different here, however, as ranking by colleges are based on a common exam and hence is common for different university programs on the same subject. Biró, Fleiner, Irving, and Manlove (2010) propose an elegant matching mechanism in such an environment.

A.4.1. *Institutional Backgrounds.* In recent years, how to organize medical training has been a contentious topic in the U.K., and the system has undergone a number of drastic changes. This section describes the current system, whose basic structure was set up in 2005.[64]

In order to practice medicine in the U.K., graduates from medical schools must undertake two years of training. The organization is called the Foundation Programme, and places about 7,000 medical school graduates to training programs every year.[65] In the first round of the matching scheme of the Foundation Programme, applicants are matched to one of 25 "foundation schools" by a national matching process. A foundation school is a consortium made of medical schools and other organizations, and each foundation school largely corresponds to a region of the country. Upon being matched to a foundation school, students are matched to individual training programs within that foundation school in the second round of the matching process. In these processes, applicants are assigned a numerical score, which may result in ties. Until 2011, the Boston mechanism (also known as the "first-choice-first" mechanism in the U.K.) was in use, based on the numerical score and a random tie-breaking.[66] Beginning in 2012, a serial dictatorship algorithm based on the score with random tie-breaking is used for allocation to foundation schools.[67] It is up to individual foundation schools as to how they match their assigned applicants to programs in their region. In Scotland, for example, a stable mechanism was in use to allocate students to programs within the region until serial dictatorship based on applicant scores and random tie-breaking replaced it in 2010.[68]

For our purposes, an especially interesting point is that the mechanism used in U.K. medical matching has two rounds, in which students are assigned to a region first, and then to a program within their assigned region. Although the specific mechanisms used in the second round vary from region to region, the United Kingdom as a whole uses a two-round mechanism.

---

[64]We are grateful to Peter Biró, Rob Irving, and David Manlove for answering our questions about medical match in the U.K.

[65]Some institutional details and the statistics reported here can be found in the Foundation Programme's webiste, especially in its annual reports: see for example its 2011 annual report at http://www.foundationprogramme.nhs.uk/download.asp?file=Foundation_Programme_Annual_Report_Nov11_FINAL.pdf.

[66]See Abdulkadiroğlu and Sönmez (2003) who study the Boston mechanism in the school choice context.

[67]The algorithm is described at http://www.foundationprogramme.nhs.uk/pages/medical-students/faqs#answer39

[68]There are applicants who participate as couples, and the algorithms handle these couples in certain manners. See Irving and Manlove (2009) for details.

A.4.2. *Formal Analysis.* As indicated above, mechanisms used in U.K. medical match (and Scottish teacher matching as mentioned in Appendix A.5) have many variations, but their basic structure is common in the sense that applicants are matched by a two-round procedure. Formally, we consider a mechanism in which applicants are matched to a region (up to its regional cap) in the first round, and then they are matched to a hospital within the assigned region in the second round. For concreteness we focus on the mechanism in which serial dictatorship is used in both rounds, but the main conclusions can be obtained for other mechanisms as well (we describe the details later in this section).

The first example shows that the outcome of this two-round mechanism may be unstable.

Example 13. There are two regions r_1 and r_2 with regional caps $q_{r_1} = q_{r_2} = 2$. There are two hospitals h_1 and h_3 in r_1 while there is one hospital h_2 in r_2. Hospital capacities are $q_{h_1} = 1$, $q_{h_2} = 2$, and $q_{h_3} = 1$. Suppose that there are 3 doctors, d_1, d_2, and d_3. Preference profile $\succ$ is as follows:

$$\succ_{h_i}: d_1, d_2, d_3 \text{ for all } i,$$

$$\succ_{d_j}: h_1, h_2, h_3 \text{ for all } j.$$

And let us assume that, in both rounds, the serial dictatorship is used with respect to the ordering d_1, d_2, and d_3. That is, d_1 is matched to her most preferred region (or hospital), d_2 is matched to the most preferred region (or hospital) that are still available, and so on. Note that we assume that hospital preferences are common and coincide with the applicant ordering in the serial dictatorship. This assumption is meant to make stability as easy to obtain as possible, because if serial order and hospital preferences are different, it is almost trivial to obtain unstable matchings (indeed, under the original serial dictatorship, the resulting matching is stable under this assumption, but not otherwise). Assume that each docotor prefers r_1 most and r_2 second.[69] Let target capacities be arbitrary.

At the first round of this mechanism, d_1 and d_2 are matched to r_1, while d_3 is matched to r_2. In the second round, d_1 is matched to her first choice h_1, d_2 is matched to h_3, which is the only remaining hospital in region r_1, and d_3 is matched to hospital h_2, resulting in

$$\mu = \begin{pmatrix} h_1 & h_2 & h_3 \\ d_1 & d_3 & d_2 \end{pmatrix}.$$

This matching μ is unstable, because d_2 and h_2 form a legitimate blocking pair: d_2 prefers h_2 to its match h_3 and h_2 prefers d_2 to its match d_3. □

---

[69]Such reported preferences may arise if, for instance, she believes that there is nonzero probability to be matched with h_1 and her cardinal utility from h_1 is sufficiently high.

Although we phrased the above example in the context of a mechanism both of whose rounds employ the serial dictatorship, the same point can be made for other two-round mechanisms. Consider, for instance, the case in which the first-round procedure is a Boston mechanism (as was the case in the U.K. until 2011) based on the above ordering, while the second round is a serial dictatorship.[70] In the above market, under this procedure d_1 and d_2 are still matched to region r_1 in the first round, and then d_2 is matched to h_3 in the second round, leading to the same matching μ of Example 13, thus to instability. This observation shows that the problem of instability is not restricted to the detail of the current mechanism using serial dictatorship in both rounds, but rather a general feature of two-round systems that have been the basic framework of the U.K. medical match.

The following example shows that the matching resulting from the U.K. medical match can be inefficient.

Example 14. There are two regions r_1 and r_2 with regional caps $q_{r_1} = 2$ and $q_{r_2} = 1$. There are two hospitals h_1 and h_3 in r_1 while there is one hospital h_2 in r_2. Each hospital h has a capacity of $q_h = 1$. Suppose that there are 3 doctors, d_1, d_2, d_3. Preference profile $\succ$ is as follows:

$$\begin{aligned}
&\succ_{h_i}: d_1, d_2, d_3 \text{ for all } i, \\
&\succ_{d_1}: h_1, h_2, h_3, \\
&\succ_{d_2}: h_1, h_2, \\
&\succ_{d_3}: h_1, h_3, h_2.
\end{aligned}$$

As before, let us assume that the serial dictatorship with ordering d_1, d_2, and d_3 is used in both rounds. Assume further that doctors' preferences over the regions are induced in the manner specified in Example 13.

At the first round of this mechanism, d_1 and d_2 are matched to r_1, while d_3 is matched to r_2. In the second round, d_1 is matched to her first choice h_1, d_2 is unmatched, and d_3 is matched to hospital h_2, resulting in matching

$$\mu = \begin{pmatrix} h_1 & h_2 & h_3 & \emptyset \\ d_1 & d_3 & \emptyset & d_2 \end{pmatrix}.$$

---

[70]Recall that the first round algorithm was changed from the Boston mechanism to the serial dictatorship only beginning in 2012 in the U.K. medical match, while serial dictatorship was already in use in Scotland in 2009.

Consider a matching μ' defined by,

$$\mu' = \begin{pmatrix} h_1 & h_2 & h_3 \\ d_1 & d_2 & d_3 \end{pmatrix}.$$

The latter matching satisfies all regional caps and Pareto dominates the former matching μ: d_1 and h_1 are indifferent between μ and μ', while every other agent is made strictly better off at μ' than at μ. Therefore the matching μ is inefficient. $\square$

The last drawback of the two-round mechanism we point out involves incentives. Serial dictatorship is strategy-proof, and this property is often regarded as one of the main advantages of this mechanism. In a two-round mechanism, however, there exists no dominant strategy even if both rounds employ serial dictatorship.

Example 15. Consider the market defined in Example 13. Note that it is a weakly dominant strategy to report true preferences in any subgame of the second round, i.e., once doctors are matched to regions, so assume that all doctors report true preferences in the second round. Suppose that doctors report preferences over regions as in Example 13. Then doctor d_2 is assigned to her third choice hospital h_3. However, if d_2 reports region r_2 to be her most preferred region while no other doctor changes his reported preference, then she is matched to h_2, which is the optimal matching possible for any of her reported preferences. In other words, reporting r_2 as the most preferred region is a best response while reporting r_1 is not. Next, consider a report of d_1 that reports r_2 to be his most preferred region. Then d_2 is matched to h_1 if she reports r_1 to be her most preferred region while she is matched to her less preferred hospital h_2 if she reports r_2 to be her most preferred region. In other words, reporting r_1 as the most preferred region is a best response while reporting r_2 is not. Therefore there is no dominant strategy.[71] $\square$

A.5. **Probationary Teacher Matching in Scotland.** Another problem of interest is the matching of new teachers (called probationary teachers) to schools. Teachers in Scotland need to get a training as probationers for one year. The General Teaching Council for Scotland (GTCS) runs a procedure called the Teacher Induction Scheme, which allocates probationary teachers to training posts in Scottish schools.[72] Scotland has 32 local authorities, and probationary teachers and these local authorities are matched in the first round of the mechanism. Information about the algorithm used is unavailable to our knowledge, but some documents suggest that a slight variant of the random serial

---

[71]It is trivial, and hence omitted, to show that reporting no region to be acceptable is weakly dominated, so the above argument is enough to establish the claim.

[72]See http://www.gtcs.org.uk/home/students/teacher-induction-scheme-faq.aspx.

dictatorship is used.[73] Then each local authority decides which probationers matched to it are sent to which schools under its control, and that round occurs subsequently to the first round. The mechanism that local authorities use in this round is up to each local authority, and appears to vary widely from one local authority to another.

This scheme has a lot in common with the previous examples. As in the U.K. medical match, the matching clearinghouse first assigns teachers to a local authority, who then assigns them to schools under its control.

Appendix B. A General Model

Let **regional preferences** $\succeq_r$ be a weak ordering over nonnegative-valued integer vectors $W_r := \{w = (w_h)_{h \in H_r} | w_h \in \mathbb{Z}_+\}$. That is, $\succeq_r$ is a binary relation that is complete and transitive (but not necessarily antisymmetric). We write $w \succ_r w'$ if and only if $w \succeq_r w'$ holds but $w' \succeq_r w$ does not. Vectors such as w and w' are interpreted to be supplies of acceptable doctors to the hospitals in region r, but they only specify how many acceptable doctors apply to each hospital and no information is given as to who these doctors are. Given $\succeq_r$, a function $\tilde{\text{Ch}}_r : W_r \to W_r$ is an **associated quasi choice rule** if $\tilde{\text{Ch}}_r(w) \in \arg\max_{\succeq_r} \{w' | w' \leq w\}$ for any non-negative integer vector $w = (w_h)_{h \in H_r}$.[74] We require that the quasi choice rule $\tilde{\text{Ch}}_r$ be **consistent**, that is, $\tilde{\text{Ch}}_r(w) \leq w' \leq w \Rightarrow \tilde{\text{Ch}}_r(w') = \tilde{\text{Ch}}_r(w)$.[75] This condition requires that, if $\tilde{\text{Ch}}_r(w)$ is chosen at w and the supply decreases to $w' \leq w$ but $\tilde{\text{Ch}}_r(w)$ is still available under w', then the same choice $\tilde{\text{Ch}}_r(w)$ should be made under w' as well. Note that there may be more than one quasi choice rule associated with a given weak ordering $\succeq_r$ because the set $\arg\max_{\succeq_r} \{w' | w' \leq w\}$ may

---

[73]"Teacher Induction Scheme 2008/2009," http://www.scotland.gov.uk/Resource/Doc/200891/0053701.pdf states that "A computer system will match and allocate students to local authorities using each local authority's vacancy list and student's preference list. You will be chosen at random and matched against your five preferences, beginning with your first preference. Where an appropriate vacancy is unavailable, you will be matched against your second preference, and so on until an appropriate match is found." As indicated above, a probationary teacher is asked to only rank 5 local authorities, unlike the exact random serial dictatorship. Another complication is that a student can alternatively tick a preference waiver box indicating that they are happy to work anywhere in Scotland. Those who choose the preference waiver option are paid additional compensation.

[74]For any two vectors $w = (w_h)_{h \in H_r}$ and $w' = (w'_h)_{h \in H_r}$, we write $w \leq w'$ if and only if $w_h \leq w'_h$ for all $h \in H_r$. We write $w \lneqq w'$ if and only if $w \leq w'$ and $w_h < w'_h$ for at least one $h \in H_r$. For any $W'_r \subseteq W_r$, $\arg\max_{\succeq_r} W'_r$ is the set of vectors $w \in W'_r$ such that $w \succeq_r w'$ for all $w' \in W'_r$.

[75]In Appendix G, we show that if a regional preference satisfies substitutability and its associated quasi choice rule is acceptant, as defined later, then the quasi choice rule satisfies consistency. Aygün and Sönmez (2012) independently prove analogous results although they do not work on substitutability defined over the space of integer vectors.

not be a singleton for some $\succeq_r$ and w. Note also that there always exists a consistent quasi choice rule.[76] We assume that the regional preferences $\succeq_r$ satisfy the following mild regularity conditions:

(1) $w' \succ_r w$ if $w_h > q_h \geq w'_h$ for some $h \in H_r$ and $w'_{h'} = w_{h'}$ for all $h' \neq h$.

This property says that the region desires no hospital to be forced to be assigned more doctors than its real capacity. This condition implies that, for any w, the component $[\tilde{\text{Ch}}_r(w)]_h$ of $\tilde{\text{Ch}}_r(w)$ for h satisfies $[\tilde{\text{Ch}}_r(w)]_h \leq q_h$ for each $h \in H_r$, that is, the capacity constraint for each hospital is respected by the (quasi) choice of the region.

(2) $w' \succ_r w$ if $\sum_{h \in H_r} w_h > q_r \geq \sum_{h \in H_r} w'_h$.

This property simply says that region r prefers the total number of doctors in the region to be at most its regional cap. This condition implies that $\sum_{h \in H_r} (\tilde{\text{Ch}}_r(w))_h \leq q_r$ for any w, that is, the regional cap is respected by the (quasi) choice of the region.

(3) If $w' \lneq w \leq q_{H_r} := (q_h)_{h \in H_r}$ and $\sum_{h \in H_r} w_h \leq q_r$, then $w \succ_r w'$.

This condition formalizes the idea that region r prefers to fill as many positions of hospitals in the region as possible so long as doing so does not lead to a violation of the hospitals' real capacities or the regional cap. This requirement implies that any associated quasi choice rule is **acceptant** (Kojima and Manea, 2009), that is, for each w, if there exists h such that $[\tilde{\text{Ch}}_r(w)]_h < \min\{q_h, w_h\}$, then $\sum_{h' \in H_r} [\tilde{\text{Ch}}_r(w)]_{h'} = q_r$. This captures the idea that the social planner should not waste caps allocated to the region: If some doctor is rejected by a hospital even though she is acceptable to the hospital and the hospital's capacity is not binding, then the regional cap should be binding.

Definition 3. The regional preferences $\succeq_r$ are **substitutable** if there exists an associated quasi choice rule $\tilde{\text{Ch}}_r$ that satisfies $w \leq w' \Rightarrow \tilde{\text{Ch}}_r(w) \geq \tilde{\text{Ch}}_r(w') \wedge w$.

Notice that the condition in this definition is equivalent to

$$(B.1) \qquad w \leq w' \Rightarrow [\tilde{\text{Ch}}_r(w)]_h \geq \min\{[\tilde{\text{Ch}}_r(w')]_h, w_h\} \text{ for every } h \in H_r.$$

This condition says that, when the supply of doctors is increased, the number of accepted doctors at a hospital can increase only when the hospital has accepted all acceptable

---

[76]To see this point consider preferences $\succeq'_r$ such that $w \succ'_r w'$ if $w \succ_r w'$ and $w = w'$ if $w \succeq'_r w'$ and $w' \succeq'_r w$. The quasi choice rule that chooses (the unique element of) $\arg\max_{\succeq'_r}\{w'|w' \leq w\}$ for each w is clearly consistent with $\succeq_r$.

doctors under the original supply profile. Formally, condition (B.1) is equivalent to

$$\text{(B.2)} \qquad w \leq w' \text{ and } [\tilde{\text{Ch}}_r(w)]_h < [\tilde{\text{Ch}}_r(w')]_h \Rightarrow [\tilde{\text{Ch}}_r(w)]_h = w_h.$$

To see that condition (B.1) implies condition (B.2), suppose that $w \leq w'$ and $[\tilde{\text{Ch}}_r(w)]_h < [\tilde{\text{Ch}}_r(w')]_h$. These assumptions and condition (B.1) imply $[\tilde{\text{Ch}}_r(w)]_h \geq w_h$. Since $[\tilde{\text{Ch}}_r(w)]_h \leq w_h$ holds by the definition of $\tilde{\text{Ch}}_r$, this implies $[\tilde{\text{Ch}}_r(w)]_h = w_h$. To see that condition (B.2) implies condition (B.1), suppose that $w \leq w'$. If $[\tilde{\text{Ch}}_r(w)]_h \geq [\tilde{\text{Ch}}_r(w')]_h$, the conclusion of (B.1) is trivially satisfied. If $[\tilde{\text{Ch}}_r(w)]_h < [\tilde{\text{Ch}}_r(w')]_h$, then condition (B.2) implies $[\tilde{\text{Ch}}_r(w)]_h = w_h$, thus the conclusion of (B.1) is satisfied.

This definition of substitutability is analogous to *persistence* by Alkan and Gale (2003), who define the condition on a choice function in a slightly different context. While our definition is similar to substitutability as defined in standard matching models (see Chapter 6 of Roth and Sotomayor (1990) for instance), there are two differences: (i) it is now defined on a region as opposed to a hospital, and (ii) it is defined over vectors that only specify how many doctors apply to hospitals in the region, and it does not distinguish different doctors.

Given $(\succeq_r)_{r \in R}$, stability is defined as follows.

Definition 4. A matching μ is **stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then (i) $|\mu_{r(h)}| = q_{r(h)}$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and

(iii') either $\mu_d \notin H_{r(h)}$ or $w \succeq_{r(h)} w'$,

where $w_{h'} = |\mu_{h'}|$ for all $h' \in H_{r(h)}$ and $w'_h = w_h + 1$, $w'_{\mu_d} = w_{\mu_d} - 1$ and $w'_{h'} = w_{h'}$ for all other $h' \in H_{r(h)}$.

Given the above properties, we can think of the following (generalized) flexible deferred acceptance algorithm:

The (Generalized) Flexible Deferred Acceptance Algorithm For each region r, fix an associated quasi choice rule $\tilde{\text{Ch}}_r$ which satisfies condition (B.1). Note that the assumption that $\succeq_r$ is substitutable assures the existence of such a quasi choice rule.

(1) Begin with an empty matching, that is, a matching μ such that $\mu_d = \emptyset$ for all $d \in D$.

(2) Choose a doctor d arbitrarily who is currently not tentatively matched to any hospital and who has not applied to all acceptable hospitals yet. If such a doctor does not exist, then terminate the algorithm.

(3) Let d apply to the most preferred hospital $\bar{h}$ at $\succ_d$ among the hospitals that have not rejected d so far. If d is unacceptable to $\bar{h}$, then reject this doctor and go

back to Step 2. Otherwise, let r be the region such that $\bar{h} \in H_r$ and define vector $w = (w_h)_{h \in H_r}$ by

(a) $w_{\bar{h}}$ is the number of doctors currently held at $\bar{h}$ plus one, and

(b) w_h is the number of doctors currently held at h if $h \neq \bar{h}$.

(4) Each hospital $h \in H_r$ considers the new applicant d (if $h = \bar{h}$) *and* doctors who are temporarily held from the previous step together. It holds its $(\tilde{\mathrm{Ch}}_r(w))_h$ most preferred applicants among them temporarily and rejects the rest (so doctors held at this step may be rejected in later steps). Go back to Step 2.

We define the **(generalized) flexible deferred acceptance mechanism** to be a mechanism that produces, for each input, the matching given at the termination of the above algorithm.

B.1. **Associated Matching Model with Contracts.** It is useful to relate our model to a (many-to-many) matching model with contracts (Hatfield and Milgrom, 2005).[77] Let there be two types of agents, doctors in D and regions in R. Note that we regard a region, instead of a hospital, as an agent in this model. There is a set of contracts $X = D \times H$.

We assume that, for each doctor d, any set of contracts with cardinality two or more is unacceptable, that is, a doctor wants to sign at most one contract. For each doctor d, her preferences $\succ_d$ over $(\{d\} \times H) \cup \{\emptyset\}$ are given as follows.[78] We assume $(d, h) \succ_d (d, h')$ in this model if and only if $h \succ_d h'$ in the original model, and $(d, h) \succ_d \emptyset$ in this model if and only if $h \succ_d \emptyset$ in the original model.

For each region $r \in R$, we assume that the region has preferences $\succeq_r$ and its associated choice rule $\mathrm{Ch}_r(\cdot)$ over all subsets of $D \times H_r$. For any $X' \subset D \times H_r$, let $w(X') := (w_h(X'))_{h \in H_r}$ be the vector such that $w_h(X') = |\{(d, h) \in X' | d \succ_h \emptyset\}|$. For each X', the chosen set of contracts $\mathrm{Ch}_r(X')$ is defined by

(B.3)
$$\mathrm{Ch}_r(X') = \bigcup_{h \in H_r} \left\{ (d, h) \in X' \ \middle| \ |\{d' \in D | (d', h) \in X', d' \succeq_h d\}| \leq (\tilde{\mathrm{Ch}}_r(w(X')))_h \right\}.$$

That is, each hospital $h \in H_r$ chooses its $(\tilde{\mathrm{Ch}}_r(w(X')))_h$ most preferred contracts available in X'.

We extend the domain of the choice rule to the collection of all subsets of X by setting $\mathrm{Ch}_r(X') = \mathrm{Ch}_r(\{(d, h) \in X' | h \in H_r\})$ for any $X' \subseteq X$.

---

[77]See Fleiner (2003) for a related analysis.

[78]We abuse notation and use the same notation $\succ_d$ for preferences of doctor d both in the original model and in the associated model with contracts.

Definition 5 (Hatfield and Milgrom (2005))**.** Choice rule $\text{Ch}_r(\cdot)$ satisfies the **substitutes condition** if there does not exist contracts $x, x' \in X$ and a set of contracts $X' \subseteq X$ such that $x' \notin \text{Ch}_r(X' \cup \{x'\})$ and $x' \in \text{Ch}_r(X' \cup \{x, x'\})$.

In other words, contracts are substitutes if adding a contract to the choice set never induces a region to choose a contract it previously rejected. Hatfield and Milgrom (2005) show that there exists a stable allocation (defined in Definition 7) when contracts are substitutes for every region.

Definition 6 (Hatfield and Milgrom (2005))**.** Choice rule $\text{Ch}_r(\cdot)$ satisfies the **law of aggregate demand** if for all $X' \subseteq X'' \subseteq X$, $|\text{Ch}_r(X')| \leq |\text{Ch}_r(X'')|$.

Proposition 5. *Suppose that $\succeq_r$ is substitutable. Then choice rule $\text{Ch}_r(\cdot)$ defined above satisfies the substitutes condition and the law of aggregate demand.*

Proof. Fix a region $r \in R$. Let $X' \subseteq X$ be a subset of contracts and $x = (d, h) \in X \setminus X'$ where $h \in H_r$. Let $w = w(X')$ and $w' = w(X' \cup x)$. To show that Ch_r satisfies the substitutes condition, we consider a number of cases as follows.

(1) Suppose that $\emptyset \succ_h d$. Then $w' = w$ and, for each $h' \in H_r$, the set of acceptable doctors available at $X' \cup x$ is identical to the one at X'. Therefore, by inspection of the definition of Ch_r, we have $\text{Ch}_r(X' \cup x) = \text{Ch}_r(X')$, satisfying the conclusion of the substitutes condition in this case.

(2) Suppose that $d \succ_h \emptyset$.

   (a) Consider a hospital $h' \in H_r \setminus h$. Note that we have $w'_{h'} = w_{h'}$. This and the inequality $[\tilde{\text{Ch}}_r(w')]_{h'} \leq w'_{h'}$ (which always holds by the definition of $\tilde{\text{Ch}}_r$) imply that $[\tilde{\text{Ch}}_r(w')]_{h'} \leq w_{h'}$. Thus we obtain $\min\{[\tilde{\text{Ch}}_r(w')]_{h'}, w_{h'}\} = [\tilde{\text{Ch}}_r(w')]_{h'}$. Since $w' \geq w$ and condition (B.1) holds, this implies that

$$\text{(B.4)} \qquad\qquad [\tilde{\text{Ch}}_r(w)]_{h'} \geq [\tilde{\text{Ch}}_r(w')]_{h'}.$$

     Also observe that the set $\{d' \in D | (d', h') \in X'\}$ is identical to $\{d' \in D | (d', h') \in X' \cup x\}$, that is, the sets of doctors that are available to hospital h' are identical under X' and $X' \cup x$. This fact, inequality (B.4), and the definition of Ch_r imply that if $x' = (d', h') \notin \text{Ch}_r(X')$, then $x' \notin \text{Ch}_r(X' \cup x)$, obtaining the conclusion for the substitute condition in this case.

   (b) Consider hospital h.

     (i) Suppose that $[\tilde{\text{Ch}}_r(w)]_h \geq [\tilde{\text{Ch}}_r(w')]_h$. In this case we follow an argument similar to (but slightly different from) Case (2a): Note that the set $\{d' \in D | (d', h) \in X'\}$ is a subset of $\{d' \in D | (d', h) \in X' \cup x\}$, that

is, the set of doctors that are available to hospital h under X' is smaller than under $X' \cup x$. These properties and the definition of Ch_r imply that if $x' = (d', h) \in X' \setminus \mathrm{Ch}_r(X')$, then $x' \in X' \setminus \mathrm{Ch}_r(X' \cup x)$, obtaining the conclusion for the substitute condition in this case.

(ii) Suppose that $[\tilde{\mathrm{Ch}}_r(w)]_h < [\tilde{\mathrm{Ch}}_r(w')]_h$. This assumption and (B.2) imply $[\tilde{\mathrm{Ch}}_r(w)]_h = w_h$. Thus, by the definition of Ch_r, any contract $(d', h) \in X'$ such that $d' \succ_h \emptyset$ is in $\mathrm{Ch}_r(X')$. Equivalently, if $x' = (d', h) \in X' \setminus \mathrm{Ch}_r(X')$, then $\emptyset \succ_h d'$. Then, again by the definition of Ch_r, it follows that $x' \notin \mathrm{Ch}_r(X' \cup x)$ for any contract $x' = (d', h) \in X' \setminus \mathrm{Ch}_r(X')$. Thus we obtain the conclusion of the substitute condition in this case.

To show that Ch_r satisfies the law of aggregate demand, simply note that $\tilde{\mathrm{Ch}}_r$ is acceptant by assumption. This leads to the desired conclusion. $\qquad\square$

A subset X' of $X = D \times H$ is said to be **individually rational** if (1) for any $d \in D$, $|\{(d, h) \in X' | h \in H\}| \leq 1$, and if $(d, h) \in X'$ then $h \succ_d \emptyset$, and (2) for any $r \in R$, $\mathrm{Ch}_r(X') = X' \cap (D \times H_r)$.

Definition 7. A set of contracts $X' \subseteq X$ is a **stable allocation** if

(1) it is individually rational, and
(2) there exists no region $r \in R$, hospital $h \in H_r$, and a doctor $d \in D$ such that $(d, h) \succ_d x$ and $(d, h) \in \mathrm{Ch}_r(X' \cup \{(d, h)\})$, where x is the contract that d receives at X' if any and $\emptyset$ otherwise.

When condition (2) is violated by some (d, h), we say that (d, h) is a **block** of X'.

Given any individually rational set of contracts X', define a **corresponding matching** $\mu(X')$ in the original model by setting $\mu_d(X') = h$ if and only if $(d, h) \in X'$ and $\mu_d(X') = \emptyset$ if and only if no contract associated with d is in X'. Since each doctor regards any set of contracts with cardinality of at least two as unacceptable, each doctor receives at most one contract at X' and hence $\mu(X')$ is well defined for any individually rational X'.

Proposition 6. *If X' is a stable allocation in the associated model with contracts, then the corresponding matching $\mu(X')$ is a stable matching in the original model.*

Proof. Suppose that X' is a stable allocation in the associated model with contracts and denote $\mu := \mu(X')$. Individual rationality of μ is obvious from the construction of μ. Suppose that (d, h) is a blocking pair of μ. Denoting $r := r(h)$, by the definition of stability, it suffices to show that the following conditions (B.5) and (B.6) hold if $\mu_d \notin H_r$,

and (B.5), (B.6) and (B.7) hold if $\mu_d \in H_r$:

(B.5) $$|\mu_{H_r}| = q_r,$$

(B.6) $$d' \succ_h d \text{ for all } d' \in \mu_h,$$

(B.7) $$w \succeq_r w',$$

where $w = (w_h)_{h \in H_r}$ is defined by $w_{h'} = |\mu_{h'}|$ for all $h' \in H_r$ while $w' = (w'_h)_{h \in H_r}$ is defined by $w'_h = w_h + 1$, $w'_{\mu_d} = w_{\mu_d} - 1$ (if $\mu_d \in H_r$) and $w'_{h'} = w_{h'}$ for all other $h' \in H_r$.

Claim 1. *Conditions (B.5) and (B.6) hold (irrespectively of whether $\mu_d \in H_r$ or not).*

Proof. First note that the assumption that $h \succ_d \mu_d$ implies that $(d, h) \succ_d x$ where x denotes the (possibly empty) contract that d signs under X'. Let $w'' = (w''_h)_{h \in H_r}$ be defined by $w''_h = w_h + 1$ and $w''_{h'} = w_{h'}$ for all other $h' \in H_r$.

(1) Assume by contradiction that condition (B.6) is violated, that is, $d \succ_h d'$ for some $d' \in \mu_h$. First, by consistency of $\tilde{\text{Ch}}_r$, we have $[\tilde{\text{Ch}}_r(w'')]_h \geq [\tilde{\text{Ch}}_r(w)]_h$.[79] That is, weakly more contracts involving h are signed at $X' \cup (d, h)$ than at X'. This property, together with the assumptions that $d \succ_h d'$ and that $(d', h) \in X'$ imply that $(d, h) \in \text{Ch}_r(X' \cup (d, h))$.[80] Thus, together with the above-mentioned property that $(d, h) \succ_d x$, (d, h) is a block of X' in the associated model of matching with contracts, contradicting the assumption that X' is a stable allocation.

(2) Assume by contradiction that condition (B.5) is violated, so that $|\mu_{H_r}| \neq q_r$. Then, since $|\mu_{H_r}| \leq q_r$ by the construction of μ and the assumption that X' is individually rational, it follows that $|\mu_{H_r}| < q_r$. Then $(d, h) \in \text{Ch}_r(X' \cup (d, h))$ because,

   (a) $d \succ_h \emptyset$ by assumption,
   (b) since $\sum_{h \in H_r} w_h = \sum_{h \in H_r} |\mu_h| = |\mu_{H_r}| < q_r$, it follows that $\sum_{h \in H_r} w''_h = \sum_{h \in H_r} w_h + 1 \leq q_r$. Moreover, $|\mu_h| < q_h$ because (d, h) is a blocking pair by

---

[79]To show this claim, assume for contradiction that $[\tilde{\text{Ch}}_r(w'')]_h < [\tilde{\text{Ch}}_r(w)]_h$. Then, $[\tilde{\text{Ch}}_r(w'')]_h < [\tilde{\text{Ch}}_r(w)]_h \leq w_h$. Moreover, since $w''_{h'} = w_{h'}$ for every $h' \neq h$ by construction of w'', it follows that $[\tilde{\text{Ch}}_r(w'')]_{h'} \leq w''_{h'} = w_{h'}$. Combining these inequalities, we have that $\tilde{\text{Ch}}_r(w'') \leq w$. Also we have $w \leq w''$ by the definition of w'', so it follows that $\tilde{\text{Ch}}_r(w'') \leq w \leq w''$. Thus, by consistency of $\tilde{\text{Ch}}_r$, we obtain $\tilde{\text{Ch}}_r(w'') = \tilde{\text{Ch}}_r(w)$, a contradiction to the assumption $[\tilde{\text{Ch}}_r(w'')]_h < [\tilde{\text{Ch}}_r(w)]_h$.

[80]The proof of this claim is as follows. $\text{Ch}_r(X')$ induces hospital h to select its $[\tilde{\text{Ch}}_r(w)]_h$ most preferred contracts while $\text{Ch}_r(X' \cup (d, h))$ induces h to select a weakly larger number $[\text{Ch}_r(w'')]_h$ of its most preferred contracts. Since (d', h) is selected as one of the $[\tilde{\text{Ch}}_r(w)]_h$ most preferred contracts for h at X' and $d \succ_h d'$, we conclude that (d, h) should be one of the $[\text{Ch}_r(w'')]_h \geq [\tilde{\text{Ch}}_r(w)]_h$ most preferred contracts at $X' \cup (d, h)$, thus selected at $X' \cup (d, h)$.

assumption and (B.6) holds, so $w_h'' = |\mu_h| + 1 \leq q_h$. These properties and the assumption that $\tilde{\mathrm{Ch}}_r$ is acceptant imply that $\tilde{\mathrm{Ch}}_r(w'') = w''$. In particular, this implies that all contracts $(d', h) \in X' \cup (d, h)$ such that $d' \succ_h \emptyset$ is chosen at $\mathrm{Ch}_r(X' \cup (d, h))$.

Thus, together with the above-mentioned property that $(d, h) \succ_d x$, (d, h) is a block of X' in the associated model of matching with contract, contradicting the assumption that X' is a stable allocation.

$\square$

To finish the proof of the proposition suppose that $\mu_d \in H_r$ and by contradiction that (B.7) fails, that is, $w' \succ_r w$. Then it should be the case that $[\tilde{\mathrm{Ch}}_r(w'')]_h = w_h'' = w_h + 1 = |\mu_h| + 1$.[81] Also we have $|\mu_h| < q_h$ and hence $|\mu_h| + 1 \leq q_h$ and $d \succ_h \emptyset$, so

$$(d, h) \in \mathrm{Ch}_r(X' \cup (d, h)).$$

This relationship, together with the assumption that $h \succ_d \mu_d$, and hence $(d, h) \succ_d x$, is a contradiction to the assumption that X' is stable in the associated model with contracts.

$\square$

A **doctor-optimal stable allocation** in the matching model with contracts is a stable allocation that every doctor weakly prefers to every other stable allocation (Hatfield and Milgrom, 2005). We will show that the flexible deferred acceptance mechanism is "isomorphic" to the doctor-optimal stable mechanism in the associated matching model with contracts.

Proposition 7. *Suppose that $\succeq_r$ is substitutable for every $r \in R$. Then the doctor-optimal stable allocation in the associated matching model with contracts, X', exists. In the original model, the flexible deferred acceptance mechanism produces matching $\mu(X')$ in a finite number of steps.*

Proof. First observe that the doctor-optimal stable allocation in matching with contracts can be found by the cumulative offer process in a finite number of steps (Hatfield and Milgrom, 2005; Hatfield and Kojima, 2010). Then, we observe that each step of the flexible deferred acceptance algorithm corresponds to a step of the cumulative offer process, that

---

[81]To show this claim, assume by contradiction that $[\tilde{\mathrm{Ch}}_r(w'')]_h \leq w_h$. Then, since $w_{h'}'' = w_{h'}$ for any $h' \neq h$ by the definition of w'', it follows that $\tilde{\mathrm{Ch}}_r(w'') \leq w \leq w''$. Thus by consistency of $\tilde{\mathrm{Ch}}_r$, we obtain $\tilde{\mathrm{Ch}}_r(w'') = \tilde{\mathrm{Ch}}_r(w)$. But $\tilde{\mathrm{Ch}}_r(w) = w$ because X' is a stable allocation in the associated model of matching with contracts, so $\tilde{\mathrm{Ch}}_r(w'') = w$. This is a contradiction because $w' \leq w''$ and $w' \succ_r w$ while $\tilde{\mathrm{Ch}}_r(w'') \in \arg\max_{\succeq_r} \{w''' | w''' \leq w''\}$.

is, at each step, if d proposes to h in the flexible deferred acceptance algorithm, then at the same step of the cumulative offer process, contract (d, h) is proposed. Moreover, for each region, the set of doctors accepted for hospitals in the region at a step of the flexible deferred acceptance algorithm corresponds to the set of contracts held by the region at the corresponding step of the cumulative offer process.   □

Theorem 4. *Suppose that $\succeq_r$ is substitutable for every $r \in R$. Then the flexible deferred acceptance algorithm stops in finite steps. The mechanism produces a stable matching for any input and is group strategy-proof for doctors.*

Proof. Propositions 6 and 7 imply that the flexible deferred acceptance algorithm finds a stable matching in a finite number of steps. Also, Propositions 5 and 7 imply that the flexible deferred acceptance mechanism is (group) strategy-proof for doctors, as the substitutes condition and the law of aggregate demand imply that any mechanism that selects the doctor-optimal stable allocation is (group) strategy-proof (Hatfield and Milgrom, 2005; Hatfield and Kojima, 2009; Hatfield and Kominers, 2010).[82]   □

B.2. **Stability in The Main Text.** In this section we are going to establish Theorem 2 in the main text by showing that the stability concept in the main text can be rewritten by using a substitutable regional preferences.

Fix a region r. Given the target capacity profile $(\bar{q}_h)_{h \in H_r}$ and the vector $w \in W_r$, define the **ordered excess weight vector** $\eta(w) = (\eta_1(w), ..., \eta_{|H_r|}(w))$ by setting $\eta_i(w)$ to be the i'th lowest value (allowing repetition) of $\{w_h - \bar{q}_h | h \in H_r\}$ (we suppress dependence of η on target capacities). For example, if $w = (w_{h_1}, w_{h_2}, w_{h_3}, w_{h_4}) = (2, 4, 7, 2)$ and $(\bar{q}_{h_1}, \bar{q}_{h_2}, \bar{q}_{h_3}, \bar{q}_{h_4}) = (3, 2, 3, 0)$, then $\eta_1(w) = -1, \eta_2(w) = \eta_3(w) = 2, \eta_4(w) = 4$.

Consider the regional preferences $\succeq_r$ that compare the excess weights lexicographically. More specifically, let $\succeq_r$ be such that $w \succ_r w'$ if and only if there exists an index $i \in \{1, 2, \ldots, |H_r|\}$ such that $\eta_j(w) = \eta_j(w')$ for all $j < i$ and $\eta_i(w) > \eta_i(w')$. The associated weak regional preferences $\succeq_r$ are defined by $w \succeq_r w'$ if and only if $w \succ_r w'$ or $\eta(w) = \eta(w')$. We call such regional preferences **Rawlsian**.

---

[82]Aygün and Sönmez (2012) point out that a condition called path-independence (Fleiner, 2003) or irrelevance of rejected contracts (Aygün and Sönmez, 2012) is needed for these conclusions. Aygün and Sönmez (2012) show that the substitutes condition and the law of aggregate demand imply this condition. Since the choice rules in our context satisfy the substitutes condition and the law of aggregate demand, the conclusions go through.

Proposition 8. *Stability defined in the main text (Definition 2) is a special case of the general concept of stability in the Appendix (Definition 4) such that the regional preferences of each region are Rawlsian.*

Proof. Let μ be a matching and w be defined by $w_{h'} = |\mu_{h'}|$ for each $h' \in H_r$ and w' by $w'_h = w_h + 1$, $w'_{\mu_d} = w_{\mu_d} - 1$, and $w'_{h'} = w_{h'}$ for all $h' \in H_r \setminus \{h, \mu_d\}$. It suffices to show that $w \succeq_r w'$ if and only if $|\mu_h| + 1 - \bar{q}_h > |\mu_{\mu_d}| - 1 - \bar{q}_{\mu_d}$.

Suppose that $|\mu_h| + 1 - \bar{q}_h > |\mu_{\mu_d}| - 1 - \bar{q}_{\mu_d}$. This means that $w_h + 1 - \bar{q}_h > w_{\mu_d} - 1 - \bar{q}_{\mu_d}$, which is equivalent to either $w_h - \bar{q}_h = w_{\mu_d} - 1 - \bar{q}_{\mu_d}$ or $w_h - \bar{q}_h \geq w_{\mu_d} - \bar{q}_{\mu_d}$. In the former case, obviously $\eta(w) = \eta(w')$, so $w \succeq_r w'$. In the latter case, $\{h'|w'_{h'} - \bar{q}_{h'} < |\mu_{\mu_d}| - \bar{q}_{\mu_d}\} = \{h'|w_{h'} - \bar{q}_{h'} < |\mu_{\mu_d}| - \bar{q}_{\mu_d}\} \cup \{\mu_d\}$, and $w_{h'} = w'_{h'}$ for all $h' \in \{h'|w_{h'} - \bar{q}_{h'} < |\mu_{\mu_d}| - \bar{q}_{\mu_d}\}$. Thus we obtain $w \succ_r w'$.

If $|\mu_h| + 1 - \bar{q}_h \leq |\mu_{\mu_d}| - 1 - \bar{q}_{\mu_d}$, then obviously $w' \succ_r w$. This completes the proof. $\square$

Consider the (generalized) flexible deferred acceptance algorithm in a previous subsection. With the following quasi choice rule, this algorithm is equivalent to the flexible deferred acceptance algorithm in the main text: For each $w' \in W_r$,

(B.8)
$$\tilde{\text{Ch}}_r(w') = \max_{\substack{w=w^k \text{ for some } k \\ \sum_{h \in H_r} w_h \leq q_r}} w,$$

where $w^0 = (\min\{w'_h, \bar{q}_h\})_{h \in H_r}$ and $w^k \in W_r$ $(k = 1, 2, \dots)$ is defined by

$$w^k_{h_j} = \min\{w'_{h_j}, q_{h_j}, w^{k-1}_{h_j} + \mathbb{I}_{j \equiv k \pmod{|H_r|}}\} \quad \text{for each } j = 1, 2, \dots, |H_r|.$$

Proposition 9. *Rawlsian preferences are substitutable with the associated quasi choice rule (B.8) that corresponds to the flexible deferred acceptance algorithm in the main text.*

Proof. It is clear that the quasi choice rule $\tilde{\text{Ch}}_r$ defined in (B.8) satisfies the condition (B.1) for substitutability (as well as consistency and acceptance). Thus in the following, we will show that $\tilde{\text{Ch}}_r$ indeed satisfies $\tilde{\text{Ch}}_r(w) \in \arg\max_{\succeq_r}\{x|x \leq w\}$ for each w. Let $w' = \tilde{\text{Ch}}_r(w)$. Assume by contradiction that $w' \notin \arg\max_{\succeq_r}\{x|x \leq w\}$ and consider an arbitrary $w'' \in \arg\max_{\succeq_r}\{x|x \leq w\}$. Then we have $w'' \succ_r w'$, so there exists i such that $\eta_j(w'') = \eta_j(w')$ for every $j < i$ and $\eta_i(w'') > \eta_i(w')$. Consider the following cases.

(1) Suppose $\sum_j \eta_j(w'') > \sum_j \eta_j(w')$. First note that $\sum_j \eta_j(w'') + \sum_h \bar{q}_h = \sum_h w''_h \leq q_r$ because $w'' \in \arg\max_{\succeq_r}\{x|x \leq w\}$. Thus $\sum_h w'_h = \sum_j \eta_j(w') + \sum_h \bar{q}_h < \sum_j \eta_j(w'') + \sum_h \bar{q}_h \leq q_r$. Moreover, the assumption implies that there exists a hospital h such that $w'_h < w''_h \leq \min\{q_h, w_h\}$. These properties contradict the construction of $\tilde{\text{Ch}}_r$.

(2) Suppose $\sum_j \eta_j(w'') < \sum_j \eta_j(w')$. First note that $\sum_j \eta_j(w') + \sum_h \bar{q}_h = \sum_h w'_h \leq q_r$ by construction of $\tilde{\text{Ch}}_r$. Thus $\sum_h w''_h = \sum_j \eta_j(w'') + \sum_h \bar{q}_h < \sum_j \eta_j(w') + \sum_h \bar{q}_h \leq q_r$. Moreover, the assumption implies that there exists a hospital h such that $w''_h < w'_h \leq \min\{q_h, w_h\}$. Then, w''' defined by $w'''_h = w''_h + 1$ and $w'''_{h'} = w''_{h'}$ for all $h' \neq h$ satisfies $w''' \leq w$ and $w''' \succ_r w''$, contradicting the assumption that $w'' \in \arg\max_{\succeq_r}\{x | x \leq w\}$.

(3) Suppose that $\sum_j \eta_j(w'') = \sum_j \eta_j(w')$. Then there exists some k such that $\eta_k(w'') < \eta_k(w')$. Let $l = \min\{k | \eta_k(w'') < \eta_k(w')\}$ be the smallest of such indices. Then since $l > i$, we have $\eta_i(w') < \eta_i(w'') \leq \eta_l(w'') < \eta_l(w')$. Thus it should be the case that $\eta_i(w') + 2 \leq \eta_l(w')$. By the construction of $\tilde{\text{Ch}}_r$, this inequality holds only if $w'_h = \min\{q_h, w_h\}$, where h is an arbitrarily chosen hospital such that $w'_h - \bar{q}_h = \eta_i(w')$. Now it should be the case that $w''_h = \min\{q_h, w_h\}$ as well, because otherwise $w'' \notin \arg\max_{\succeq_r}\{x | x \leq w\}$.[83] Thus $w'_h = w''_h$. Now consider the modified vectors of both w' and w'' that delete the entries corresponding to h. All the properties described above hold for these new vectors. Proceeding inductively, we obtain $w'_h = w''_h$ for all h, that is, $w' = w''$. This is a contradiction to the assumption that $w' \notin \arg\max_{\succeq_r}\{x | x \leq w\}$ and $w'' \in \arg\max_{\succeq_r}\{x | x \leq w\}$.

The above cases complete the proof. □

Theorem 4 and Propositions 8 and 9 imply Theorem 2 in the main text.

B.3. **Alternative Criteria.** Although the main text focuses on a particular stability concept and corresponding regional preferences, called Rawlsian preferences, it is quite plausible that some societies may prefer to impose different criteria from the Rawlsian preferences. This section proposes other criteria that seem to be appealing. The following are examples of regional preferences that satisfy substitutability defined in Definition 3. In the following, we assume that $0 \succ_r w$ for any weight vector w such that $\sum_{h \in H_r} w_h > q_r$ or $w_h > q_h$ for some $h \in H_r$. Thus in (1) - (4) below, we assume that any weight vector w satisfies $\sum_{h \in H_r} w_h \leq q_r$ and $w_h \leq q_h$ for all $h \in H_r$.

(1) "Equal gains": Let the region prefer a distribution that equalizes the weights across hospitals in the region as much as possible. Formally, such a preference, which we call the **equal gains** preferences, can be expressed as the Rawlsian preferences

---

[83]The proof that $w'' \notin \arg\max_{\succeq_r}\{x | x \leq w\}$ if $w''_h < \min\{q_h, w_h\}$ is as follows. Suppose that $w''_h < \min\{q_h, w_h\}$. Consider w''' defined by $w'''_h = w''_h + 1$, $w'''_{h'} = w''_{h'} - 1$ for some h' such that $w''_{h'} - \bar{q}_{h'} = \eta_i(w'')$, and $w'''_{h''} = w''_{h''}$ for all $h'' \in H_r \setminus \{h, h'\}$. Then we have $w'''_h - \bar{q}_h = w''_h - \bar{q}_h + 1 \leq w'_h - \bar{q}_h < w''_{h'} - \bar{q}_{h'}$, where the weak inequality follows because $w''_h < \min\{q_h, w_h\} = w'_h$. The strict inequality implies that $w'_h - \bar{q}_h \leq w''_{h'} - 1 - \bar{q}_{h'} = w'''_{h'} - \bar{q}_{h'}$. Hence $w'''_h - \bar{q}_h \leq w'''_{h'} - \bar{q}_{h'}$, which implies $w''' \succ_r w''$.

for the special case in which the target capacity for every hospital is set at zero. Since Proposition 9 shows that the Rawlsian preferences are substitutable for any target capacity profile, the equal gains preferences satisfy substitutability.

(2) "Equal Losses": Let the region prefer to equalize the "losses," that is, the differences between the (physical) capacities and the weights across hospitals in the region. More generally, one could consider the preferences for **equal losses above target capacities**, that is, the regional preferences first prefer to fill as many positions as possible to meet target capacities and then (lexicographically less importantly) prefer to equalize the losses. To formally define such preferences $\succ_r$, recall that $\eta(w)$ denotes the ordered excess weight vector as defined in Appendix B.2, and define $\hat{\eta}(w)$ as a $|H_r|$-dimensional vector whose i'th component $\hat{\eta}_i(w)$ is the i'th highest value (allowing repetition) of $\{q_h - w_h | h \in H_r\}$. We let $w \succ_r w'$ if and only if

   (a) there exists an index $i \in \{1, 2, \ldots, |H_r|\}$ such that $\min\{\eta_j(w), 0\} = \min\{\eta_j(w'), 0\}$ for all $j < i$ and $\min\{\eta_i(w), 0\} > \min\{\eta_i(w'), 0\}$, or

   (b) $\min\{\eta_i(w), 0\} = \min\{\eta_i(w'), 0\}$ for every index $i \in \{1, 2, \ldots, |H_r|\}$, and there exists an index $i \in \{1, 2, \ldots, |H_r|\}$ such that $\hat{\eta}_j(w), = \hat{\eta}_j(w')$ for all $j < i$ and $\hat{\eta}_i(w) < \hat{\eta}_i(w')$.

(3) "Proportional": The **proportional** regional preferences prefer to allocate positions to hospitals in a proportional manner subject to integer constraints. More precisely, define $\tilde{\eta}(w)$ as a $|H_r|$-dimensional vector whose i'th component $\tilde{\eta}_i(w)$ is the i'th lowest value (allowing repetition) of $\{w_h/q_h | h \in H_r\}$. We let $w \succ_r w'$ if there exists an index $i \in \{1, 2, \ldots, |H_r|\}$ such that $\tilde{\eta}_j(w), = \tilde{\eta}_j(w')$ for all $j < i$ and $\tilde{\eta}_i(w) > \tilde{\eta}_i(w')$. As above, one could consider preferences for proportional losses as well. Also, these preferences can be generalized so that these concerns enter only above target capacities (this generalization is somewhat tedious but straightforward, and can be done as in Item 2). Finally, when constructing $\tilde{\eta}_i$, we can use a denominator different from q_h.[84]

(4) "Hospital-lexicographic": Let there be a pre-specified order over hospitals, and the region lexicographically prefers filling a slot in a higher-ranked hospital to filling that of a lower-ranked hospital. For instance, the region may desire to fill positions of hospitals that are underserved within the region (say, a prefecture may desire to fill positions of a hospital in a remote island within the prefecture

---

[84]Moreover, the generalizations mentioned above can be combined. For example, the region may desire to fill capacities above targets proportionally to $q_h - \bar{q}_h$.

before other hospitals). Formally, **hospital-lexicographic** regional preferences $\succ_r$ are defined as follows. Fix an order over hospitals in r, denoted by $h_1, h_2, \ldots,$ and $h_{|H_r|}$. Let $w \succ_r w'$ if and only if there exists an index $i \in \{1, 2, \ldots, |H_r|\}$ such that $w_{h_j} = w'_{h_j}$ for all $j < i$ and $w_{h_i} > w'_{h_i}$. We note that one can also consider hospital-lexicographic preferences above targets by using the criterion for hospital-lexicographic preferences for weights above targets.

All the above regional preferences have associated quasi choice rules that satisfy the property that we call "order-respecting." To define this property, let there be a finite sequence of hospitals in region r such that each hospital h appears, potentially repeatedly, q_h times in the sequence, and the total size of the sequence is $\sum_{h \in H_r} q_h$. Consider a quasi choice rule that increases the weights of hospitals one by one following the specified order.[85] Formally, fix a vector $(h_1, h_2, \ldots, h_{\sum_{h \in H_r} q_h}) \in (H_r)^{\sum_{h \in H_r} q_h}$ such that $\#\{i \in \{1, 2, \ldots, \sum_{h \in H_r} q_h\} | h_i = h\} = q_h$ for each $h \in H_r$, and define $\tilde{\mathrm{Ch}}_r(w)$ through the following algorithm:

(1) Let w^0 be the $|H_r|$-dimensional zero vector, indexed by hospitals in H_r.
(2) For any $t \geq 0$, if $\sum_{h \in H_r} w_h^t = q_r$ or $w_h^t = \min\{q_h, w_h\}$ for all $h \in H_r$, then stop the algorithm and define $\tilde{\mathrm{Ch}}_r(w) = w^t$. If not, define w^{t+1} by:
   (a) If $w_{h_{t+1}}^t < \min\{q_{h_{t+1}}, w_{h_{t+1}}\}$, then let $w_{h_{t+1}}^{t+1} = w_{h_{t+1}}^t + 1$; otherwise, let $w_{h_{t+1}}^{t+1} = w_{h_{t+1}}^t$.
   (b) For every $h \neq h_{t+1}$, let $w_h^{t+1} = w_h^t$.

It is easy to see that any order-respecting quasi choice rule satisfies the condition in the definition of substitutability. Also it is easy to see that, for each of the above regional preferences (1) - (4), there exists an associated quasi choice rule that is order-respecting. By these observations, all of the above regional preferences are substitutable.

Remark 5. In addition, it may be of interest to consider regional preferences involving "subregions": The region prefers to assign no more doctors than a certain number to a subset of the hospitals in the region. Such preferences may arise if the society desires to impose a hierarchy of regional caps, say one cap for a prefecture and one for each district within the prefecture. Or the policy maker may desire to regulate the total number of doctors practicing in each specialty in each prefecture. In general, this type of preferences is outside of the current framework because if a cap of a district in a prefecture is filled while there are remaining seats in the prefecture as a whole, then no more doctor can be accepted to hospitals in the district and this violates the assumption that the associated

---

[85]Order-respecting quasi choice rules are similar to choice functions based on the precedence order of Kominers and Sönmez (2012), although we find no logical relationship between these two concepts.

quasi choice rule of the prefecture is acceptant. For this reason, a further generalization of our model is called for. Such a generalization is done in Online Supplementary Appendix F.

B.4. **Allocating Target Capacities.** A problem related to, but distinct from, our discussion in Appendix B.3 is how to allocate target capacities among hospitals in a region. We will not try to provide a final answer to the normative question of how to do so for several reasons. First, there may be different ways to specify the quasi choice rule even given the same target capacity profile, as we have seen in this section, and in fact there may be reasonable quasi choice rules that do not even presuppose the existence of target capacities. Second, even if we fix a quasi choice rule, the relation between target capacities and the desirability of the resulting outcome is ambiguous. For instance, Example 11 in the main text shows that the effect on hospital welfare is ambiguous. In fact, Example 20 in Appendix H shows that the same conclusion holds even if hospitals or doctors have homogeneous preferences, which are strong restrictions that often lead to strong conclusions in matching.

Despite these reservations, hospitals may still find having higher targets intuitively appealing in practice, so the problem seems to be practically important. Motivated by this observation, we present several methods to allocate target capacities that seem to be reasonable.

To do so, our starting point is to point out that the problem of allocating target capacities is similar to the celebrated "bankruptcy problem" (see Thomson (2003)). This is a useful association in the sense that, in the bankruptcy problem, there are known analyses (e.g., axiomatic characterizations) of various allocation rules, which can be utilized to judge which rule is appropriate for a given application.

To make this association, recall that in the standard bankruptcy problem, there is a divisible asset and agents whose claims sum up to (weakly) more than the amount of the available asset. Our problem is a discrete analogue of the bankruptcy problem. The regional cap q_r is an asset, hospitals in region r are agents, and physical capacity q_h is the claim of hospital h. Just as in the bankruptcy problem, the sum of the physical capacities may exceed the available regional cap, so the target capacity profile $(\bar{q}_h)_{h \in H_r}$ needs to be decided subject to the constraint $\sum_{h \in H_r} \bar{q}_h \leq q_r$.

This association suggests adaptations of well-known solutions in the bankruptcy problem to our problem, with the modification due to the fact that both the asset and the claims are discrete in our problem. The following are a few examples (in the following, we assume $\sum_{h \in H_r} q_h \geq q_r$; otherwise set $\bar{q}_h = q_h$ for all h).

(1) "Constrained Equal Awards Rule": This solution allocates the targets as equally as possible except that, for any hospital, it does not allocate a target larger than the capacity. This rule is called the **constrained equal awards rule** in the literature on the bankruptcy problem. In our context, this solution should be modified because all the targets need to be integers. Formally, a constrained equal awards rule in our context can be defined as follows:

   (a) Find λ that satisfies $\sum_{h \in H_r} \min\{\lambda, q_h\} = q_r$.

   (b) For each $h \in H_r$, if $\lambda > q_h$, then set $\bar{q}_h = q_h$. Otherwise, set $\bar{q}_h$ to be either $\lfloor \lambda \rfloor$ (the largest integer no larger than λ) or $\lfloor \lambda \rfloor + 1$, subject to the constraint that $\sum_{h \in H_r} \bar{q}_h = q_r$.

   The rule to decide which hospital receives $\lfloor \lambda \rfloor$ or $\lfloor \lambda \rfloor + 1$ is arbitrary: For any decision rule, the resulting target profiles satisfy conditions assumed in the main text. The decision can also use randomization, which may help achieve ex ante fairness.

(2) "Constrained Equal Losses Rule": This solution allocates the targets in such a way that it equates losses (that is, differences between the capacities and the targets) as much as possible, except that none of the allocated targets can be strictly smaller than zero. This rule is called the **constrained equal losses rule** in the literature on the bankruptcy problem. As in the case of the constrained equal awards rule, this solution should be modified because all the targets need to be integers. Formally, a constrained equal losses rule in our context can be defined as follows:

   (a) Find λ that satisfies $\sum_{h \in H_r} \max\{q_h - \lambda, 0\} = q_r$.

   (b) For each $h \in H_r$, if $q_h - \lambda < 0$, then set $\bar{q}_h = 0$. Otherwise, set $\bar{q}_h$ to be either $q_h - \lfloor \lambda \rfloor$ or $q_h - \lfloor \lambda \rfloor - 1$, subject to the constraint that $\sum_{h \in H_r} \bar{q}_h = q_r$.

   As in the constrained equal awards rule, the rule to decide which hospital receives $q_h - \lfloor \lambda \rfloor$ or $q_h - \lfloor \lambda \rfloor - 1$ is arbitrary: For any decision rule, the resulting target profiles satisfy conditions assumed in the main text. The decision can also use randomization, which may help achieve ex ante fairness.

(3) "Proportional Rule": This solution allocates the targets in a manner that is as proportional as possible to the hospitals' capacities. This rule is called the **proportional rule** in the literature on the bankruptcy problem. As in the case of the previous rules, this solution should be modified because all the targets need to be integers. Formally, a proportional rule in our context can be defined as follows:

   (a) Find λ that satisfies $\sum_{h \in H_r} \lambda q_h = q_r$.

(b) For each $h \in H_r$, set $\bar{q}_h$ be either $\lfloor \lambda q_h \rfloor$ or $\lfloor \lambda q_h \rfloor + 1$, subject to the constraint that $\sum_{h \in H_r} \bar{q}_h = q_r$.

As in the previous cases, the rule to decide which a hospital receives $\lfloor \lambda q_h \rfloor$ or $\lfloor \lambda q_h \rfloor + 1$ is arbitrary: For any decision rule, the resulting target profiles satisfy conditions assumed in the main text. The decision can also use randomization, which may help achieve ex ante fairness.

The proportional rule seems to be fairly appealing in practice. This rule is used in Japanese residency match and Chinese graduate school admission, for example.

Appendix C. Weak Stability

In this section, we define a weaker stability concept than strong stability and stability. Our objective here is to generalize the efficiency theorem (Theorem 1), and also to make statements of impossibility results that do not depend on the particular way of our defining stability (such as the introduction of target capacities or even regional preferences in the definition; see Remark 2, Example 6 and Section 6.1.2).

Recall that $r(h)$ is the region that h belongs to.

Definition 8. A matching μ is **weakly stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then (i) $|\mu_{r(h)}| = q_{r(h)}$ and (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$.

The difference of weak stability from strong stability defined in Definition 1 is the deletion of condition (iii), "$\mu_d \notin H_{r(h)}$." Thus, a blocking pair such that the doctor in the pair moves between two hospitals in the same region is tolerated. Moreover, since (iii') must be satisfied to be deemed as a tolerated blocking pair under stability while it need not be in weak stability, weak stability is weaker than stability.[86]

Since weak stability tolerates too many blocking pairs that do not violate regional caps, we do not necessarily claim that weak stability is the most natural stability concept. The main point that we want to make here is that, although this is a weak notion and it does not even involve the concept of target capacities (or regional preferences in general), the JRMP mechanism does not even guarantee weak stability. This point can be seen by Remark 2 in the main text.

We demonstrate that this weak notion of stability does imply a desirable property, namely efficiency:

---

[86]For an example in which the three stability concepts – weak stability, stability, and strong stability – lead to different choices of matchings, consider Example 2 with the additional specification of a target capacity profile $(\bar{q}_{h_1}, \bar{q}_{h_2}) = (1, 0)$.

Theorem 5. *Any weakly stable matching is (constrained) efficient.*

Proof. Let μ be a weakly stable matching and assume, for contradiction, that μ is not efficient. Then there exists a feasible matching μ' that Pareto dominates μ, that is, there is a feasible matching μ' such that $\mu'_i \succeq_i \mu_i$ for all $i \in D \cup H$, with at least one being strict. Noting that matching is bilateral, this implies that there exists a doctor $d \in D$ with $\mu'_d \succ_d \mu_d$. Since μ is a weakly stable matching, $\mu_d \succeq_d \emptyset$ and hence $\mu'_d \neq \emptyset$, so $\mu'_d \in H$. Denote $h = \mu'_d$. Since μ is a weakly stable matching, $h \succ_d \mu_d$ implies one of the following (cases (1) and (2) correspond to a situation in which (d, h) is not a blocking pair of μ. Case (3) covers, by the definition of weak stability, the case in which (d, h) blocks μ):

(1) $\emptyset \succ_h d$.
(2) $|\mu_h| = q_h$ and $d' \succ_h d$ for all $d' \in \mu_h$.
(3) $|\mu_{H_r}| = q_r$ for r such that $h \in H_r$ and $d' \succ_h d$ for all $d' \in \mu_h$.

Suppose $\emptyset \succ_h d$. Then, if $|\mu_h| = q_h$, then there is a doctor $d'' \in \mu'_h \setminus \mu_h$ such that $d'' \succ_h d'$ for some $d' \in \mu_h$ (otherwise, by responsiveness of the preference of h, it follows that $\mu_h \succ_h \mu'_h$). Then, since μ is weakly stable, $\mu_{d''} \succ_{d''} h = \mu'_{d''}$, contradicting the assumption that μ' Pareto dominates μ. If $|\mu_h| < q_h$, then there should be a doctor $d'' \in \mu'_h \setminus \mu_h$ such that $d'' \succ_h \emptyset$ (otherwise, by responsiveness of the preference of h, it follows that $\mu_h \succ_h \mu'_h$). Then, since μ is weakly stable, $\mu_{d''} \succ_{d''} h = \mu'_{d''}$, contradicting the assumption that μ' Pareto dominates μ.

Suppose $|\mu_h| = q_h$ and $d' \succ_h d$ for all $d' \in \mu_h$. Then there should be a doctor $d'' \in \mu'_h \setminus \mu_h$ such that $d'' \succ_h d'$ for some $d' \in \mu_h$ (otherwise, by responsiveness of the preference of h, it follows that $\mu_h \succ_h \mu'_h$). Then, since μ is weakly stable, $\mu_{d''} \succ_{d''} h = \mu'_{d''}$, contradicting the assumption that μ' Pareto dominates μ.

Suppose $|\mu_{H_r}| = q_r$ for r such that $h \in H_r$ and $d' \succ_h d$ for all $d' \in \mu_h$. Then, if $|\mu'_h| \leq |\mu_h|$, then there should be a doctor $d'' \in \mu'_h \setminus \mu_h$ such that $d'' \succ_h d'$ for some $d' \in \mu_h$ (otherwise, by responsiveness of the preference of h, it follows that $\mu_h \succ_h \mu'_h$). Then, since μ is weakly stable, $\mu_{d''} \succ_{d''} h = \mu'_{d''}$, contradicting the assumption that μ' Pareto dominates μ. If $|\mu'_h| > |\mu_h|$, then since $|\mu_{H_r}| = q_r$, there exists a hospital $h' \in H_r$ with $|\mu'_{h'}| < |\mu_{h'}|$. This, since $\mu'_{h'} \succ_{h'} \mu_{h'}$ as μ' Pareto dominates μ, implies that there should be a doctor $d'' \in \mu'_{h'} \setminus \mu_{h'}$ such that $d'' \succ_{h'} d'$ for some $d' \in \mu_{h'}$ (otherwise, by responsiveness of the preference of h', it follows that $\mu_{h'} \succ_{h'} \mu'_{h'}$). Then, since μ is weakly stable, $\mu_{d''} \succ_{d''} h' = \mu'_{d''}$, contradicting the assumption that μ' Pareto dominates μ. □

Since stability implies weak stability, Theorem 5 implies that any stable matching is efficient, as stated in Theorem 1 in the main text.

Appendix D. Remaining Proofs for the Main Text

Proof of Proposition 1. Assume that d prefers h to her outcome under the flexible deferred acceptance mechanism. Then d has applied to h and was rejected under the flexible deferred acceptance algorithm. If the number of doctors matched with h in the flexible deferred acceptance mechanism is strictly less than its target capacity, then the number of doctors who have ever applied to h and are acceptable to h is strictly smaller than the target capacity of h. This implies that any doctor who applied to h and was rejected in the flexible deferred acceptance algorithm is unacceptable to h. In particular d is unacceptable, completing the proof. □

The following result, which applies not only to matching with contract models defined over the set of contracts $D \times H$ but also to those defined over general environments, proves useful.

Lemma 1. *Consider a model of matching with contracts. Fix the set of doctors and regions as well as doctor preferences. Assume that choice rules $Ch := (Ch_r)_{r \in R}$ and $Ch' := (Ch'_r)_{r \in R}$ satisfy $Ch'_r(X') \subseteq Ch_r(X')$ for every subset of contracts X' and region r. Then the following two statements hold:*

(1) *Each doctor weakly prefers the outcome of the cumulative offer process with respect to Ch to the result with respect to Ch'. Hence each doctor weakly prefers the doctor-optimal stable allocation under Ch to the doctor-optimal stable allocation under Ch'.*

(2) *The set of contracts that have been offered up to and including the terminal step of the cumulative offer process under Ch is a subset of the corresponding set under Ch'.*

Proof. Let Y_d and Y'_d be the contracts allocated to d by the cumulative offer processes under Ch and Ch', respectively. Also, let $C(t)$ be the set of contracts that have been offered up to and including step t of the cumulative offer process under Ch, and $C'(t)$ be the corresponding set for the cumulative offer process under Ch'. Let T and T' be the terminal steps for the cumulative offer processes under Ch and Ch', respectively. We first prove Part 2 of the lemma, and then show Part 1.

Part 2: Suppose the contrary, i.e., that $C(T) \nsubseteq C'(T')$. Then there exists a step t' such that $C(t) \subseteq C'(T')$ for all $t < t'$ and $C(t') \nsubseteq C'(T')$ holds. That is, t' is the first step such that an application not made in the cumulative offer process under Ch' is made in the cumulative offer process under Ch. Let x be the contract that d offers in this step under Ch. Notice that $Y'_d \succ_d x$. This implies that $Y'_d \neq \emptyset$ and that Y'_d is rejected by r' in some

steps of the cumulative offer process under Ch, where r' is the region associated with Y_d'. Let the first of such steps be t''. Since in the cumulative offer process doctors make offers in order of their preferences, $Y_d' \succ_d x$ implies that $t'' < t'$, which in turn implies $C(t'') \subseteq C'(T')$ by the definition of t'.

Now, we show that the set of contracts accepted by r' at step t'' of the cumulative offer process under Ch is a superset of the set of contracts accepted by r' from the application pool $C(t'')$ (which is a subset of $C'(T')$) at step T' of the cumulative offer process under Ch'. To see this, note that if the same application pool $C'(T')$ is given, the set of contracts accepted by r' in the cumulative offer process under Ch is weakly larger than that under Ch' by the assumption that $Ch_r'(X') \subseteq Ch_r(X')$ for all $X' \subseteq X$ and $r \in R$. Since Ch is substitutable, subtracting applications in $C'(T') \setminus C(t'')$ does not shrink the set of contracts accepted by r' within $C(t'')$ at step t'' of the cumulative offer process under Ch, which establishes our claim.

However, this contradicts our earlier conclusion that Y_d' is rejected by r' at step t'' of the cumulative offer process under Ch while she is allocated Y_d' in the cumulative offer process under Ch'. Hence we conclude that $C(T) \subseteq C'(T')$.

Part 1: Now, since in the cumulative offer process each doctor d make offers of contracts in order of her preferences, Y_d is $\emptyset$ or the worst contract for d in the set of contracts associated with d in $C(T)$. Similarly, for each doctor d, Y_d' is $\emptyset$ or the worst contract for d in the set of contracts associated with d in $C'(T')$. If $Y_d \neq \emptyset$, this and $C(T) \subseteq C'(T')$ imply that $Y_d \succeq_d Y_d'$. If $Y_d = \emptyset$, d has applied to all acceptable contracts in the cumulative offer process under Ch. Thus $C(T) \subseteq C'(T')$ implies that she has applied to all acceptable contracts in the algorithm under Ch', too. Let x' be the worst acceptable contract in X for d, and r be a region associated with x'. At this point we already know that Y_d' is either x' or $\emptyset$, and we will show that $Y_d' = \emptyset$ in what follows. Again, $C(T) \subseteq C'(T')$ implies that all applications associated with r in $C(T)$ is in $C'(T')$. In particular, d's application to x' is in $C'(T')$. Since Ch is substitutable, subtracting applications in $C'(T') \setminus C(T)$ does not shrink the set of doctors accepted by r within $C(T)$ at step T of the deferred acceptance, so d not being accepted by r from $C(T)$ at step T of the cumulative offer process under Ch implies that she is not accepted by r from $C'(T')$ in step T' of the process under Ch' either. But since we have shown that d's offer of contract x' to r is in $C'(T')$, this implies that in the cumulative offer process under Ch', x' is rejected by r. Because x' is the worst acceptable contract for d and d's applications are made in order of her preferences, we conclude that $Y_d' = \emptyset$, thus in particular $Y_d \succeq_d Y_d'$.

This shows that each doctor $d \in D$ weakly prefers a contract allocated by the cumulative offer process under Ch to the one under Ch'.

Since the outcome of the cumulative offer process is the doctor-optimal stable allocation, the preceding proof has also shown that the doctor-optimal stable allocation under Ch is weakly more preferred to the doctor-optimal stable allocation under Ch'. □

Lemma 1 is a generalization of a number of existing results. Gale and Sotomayor (1985a,b) establish comparative statics results in one-to-one and many-to-one matching with respect to the extension of an agent's list of her acceptable partners or an addition of an agent to the market, and Crawford (1991) generalizes the results to many-to-many matching. Konishi and Ünver (2006) consider many-to-one matching and obtain a comparative statics result with respect to the changes of hospital capacities.[87] All these changes are special cases of changes in the choice rules, so these results are corollaries of Lemma 1.

Lemma 1 may be of independent interest as the most general comparative statics result known to date. In addition, the lemma implies various results that are directly relevant to the current study of regional caps, such as Theorem 3 and Propositions 2, 3, and 4 in the main text.

Proof of Theorem 3. **Part 1:** Let $\mathrm{Ch}^F = (\mathrm{Ch}_r^F)_{r \in R}$ be the choice rule associated with the flexible deferred acceptance as defined earlier, that is, for each region $r \in R$ and subset of contracts $X' \subseteq X = D \times H$, the chosen set of contracts $\mathrm{Ch}_r^F(X')$ is defined by

$$\mathrm{Ch}_r^F(X') = \bigcup_{h \in H_r} \left\{ (d, h) \in X' \ \Big| \ |\{d' \in D | (d', h) \in X', d' \succeq_h d\}| \leq (\tilde{\mathrm{Ch}}_r(w(X')))_h \right\},$$

where $\tilde{\mathrm{Ch}}_r$ corresponds to a Rawlsian regional preference of region r and $w(X') = (w_h(X'))_{h \in H_r}$ is the vector such that $w_h(X') = |\{(d, h) \in X' | d \succ_h \emptyset\}|$ (this is a special case of the choice rule (B.3)).

Moreover, consider choice rules $Ch^D = (Ch_r^D)_{r \in R}$ and $Ch^J = (Ch_r^J)_{r \in R}$ such that, for each X' and r,

$$\mathrm{Ch}_r^D(X') = \bigcup_{h \in H_r} \left\{ (d, h) \in X' \ \Big| \ |\{d' \in D | (d', h) \in X', d' \succeq_h d\}| \leq q_h \right\},$$

$$\mathrm{Ch}_r^J(X') = \bigcup_{h \in H_r} \left\{ (d, h) \in X' \ \Big| \ |\{d' \in D | (d', h) \in X', d' \succeq_h d\}| \leq \bar{q}_h \right\}.$$

---

[87]See also Kelso and Crawford (1982), who derive comparative statics results in a matching model with wages.

Clearly, both Ch^D and Ch^J satisfy the substitute condition and the law of aggregate demand. Moreover, the matchings corresponding to the results of the cumulative offer processes under Ch^D and Ch^J are identical to the results of the deferred acceptance algorithm and the JRMP mechanism, respectively. Because $\min\{\bar{q}_h, w_h\} \leq (\tilde{\mathrm{Ch}}_r(w(X')))_h \leq q_h$ for all $h \in H_r$ and X', by inspection of the above definitions of the choice rules we obtain $\mathrm{Ch}_r^J(X') \subseteq \mathrm{Ch}_r^F(X') \subseteq \mathrm{Ch}_r^D(X')$ for all X' and r. Thus the desired conclusion follows by Part 1 of Lemma 1.

Part 2: This is a direct corollary of Part 1 and the fact that none of the algorithms considered here matches a doctor to an unacceptable hospital.   □

Proof of Proposition 2. **Part 1:** First, by Part 2 of Lemma 1 and the proof of Theorem 3, the set of contracts that have been offered up to and including the terminal step under the deferred acceptance mechanism is a subset of the one under the flexible deferred acceptance mechanism. Second, by the construction of the flexible deferred acceptance algorithm, and the assumption that hospital h's target capacity is not filled, under the flexible deferred acceptance mechanism h is matched to every doctor who is acceptable to h and who applied to h in some step of the algorithm. These two facts imply the conclusion.

Part 2: First, by Part 2 of Lemma 1 and the proof of Theorem 3, the set of contracts that have been offered up to and including the terminal step under the deferred acceptance mechanism is a subset of the one under the flexible deferred acceptance mechanism. Second, by the construction of the flexible deferred acceptance algorithm, and the assumption that region r's regional cap is not filled, under the flexible deferred acceptance mechanism any hospital h in region r is matched to every doctor who is acceptable and who is among the most preferred q_h doctors who applied to h in some step of the algorithm. These two facts imply the conclusion.   □

Proof of Proposition 3. Let $\mathrm{Ch} = (\mathrm{Ch}_r)_{r \in R}$ and $\mathrm{Ch}' = (\mathrm{Ch}'_r)_{r \in R}$ be the choice rules associated with the flexible deferred acceptance mechanisms (as defined in the proof of Theorem 3) with respect to $(q_r)_{r \in R}$ and $(q'_r)_{r \in R}$, respectively.

Part 1: Because $q'_r \leq q_r$ for each $r \in R$, the definition of these choice rules implies $\mathrm{Ch}'_r(X') \subseteq \mathrm{Ch}_r(X')$ for all X' and r. Hence the desired conclusion follows by Part 1 of Lemma 1.

Part 2: Since $\mathrm{Ch}'_r(X') \subseteq \mathrm{Ch}_r(X')$ for all X' and r as mentioned in the proof of Part 1, Part 2 of Lemma 1 implies that $C(T) \subseteq C'(T')$, where C, T, C', and T' are as defined in Part 2 of the lemma. Note that the sets of contracts allocated to hospitals in r at the conclusions of the cumulative offer processes under Ch and Ch' are given as r's choice

from contracts associated with r in $C(T)$ and $C'(T')$, respectively. Because the choice rules satisfy the law of aggregate demand and the set-inclusion relationship $C(T) \subseteq C'(T')$ holds, for any r such that $q_r = q'_r$, the number of doctors matched in r under a matching produced by the flexible deferred acceptance mechanism under regional caps $(q'_r)_{r \in R}$ is weakly larger than in the matching under $(q_r)_{r \in R}$, completing the proof. $\qquad\square$

Proof of Proposition 4. Let $\mathrm{Ch} = (\mathrm{Ch}_r)_{r \in R}$ and $\mathrm{Ch}' = (\mathrm{Ch}'_r)_{r \in R}$ be the choice rules associated with the JRMP mechanisms (as defined in the proof of Theorem 3) with respect to $(\bar{q}_h)_{h \in H}$ and $(\bar{q}'_h)_{h \in H}$, respectively.

Part 1: Because $\bar{q}'_h \leq \bar{q}_h$ for each $h \in H$, the definition of these choice rules implies $\mathrm{Ch}'_r(X') \subseteq \mathrm{Ch}_r(X')$ for all X' and r. Hence the desired conclusion follows by Part 1 of Lemma 1.

Part 2: Since $\mathrm{Ch}'_r(X') \subseteq \mathrm{Ch}_r(X')$ for all X' and r as mentioned in the proof of Part 1, Part 2 of Lemma 1 implies that $C(T) \subseteq C'(T')$, where C, T, C', and T' are as defined in Part 2 of Lemma 1. Note that the matchings for h at the conclusions of the cumulative offer processes under Ch and Ch' are given as h's most preferred acceptable doctors up to $\bar{q}_h = \bar{q}'_h$ from contracts associated with h in $C(T)$ and $C'(T')$, respectively. Thus the set-inclusion relationship $C(T) \subseteq C'(T')$ implies both of the statements of Part 2.

$\qquad\square$

ONLINE SUPPLEMENTARY APPENDIX

Appendix E. Simulation Methods and Results for the case of the Japanese Residency Matching Program

In this appendix we provide results of simulations using the data on Japanese medical residency match. As we have discussed in the main text of the paper (and in the appendix), there are a number of instances around the world where distributional constraints are imposed. We chose the Japanese case as our principal target for simulation for various reasons: First, all the real data we used in our simulation are available online for free, so one can replicate our simulation results easily. Second, the Japanese medical match is highly centralized, so the conclusions from the simulation results are more meaningful than in some other applications where a (sometimes nontrivial) part of the matching procedure is not explicitly specified. Third, compared to the current practice (the JRMP mechanism), our proposal (the flexible deferred acceptance (FDA) mechanism) has an advantage in terms of efficiency and stability, while its effect on the regional balance of doctors is ambiguous from the theoretical perspective. Thus simulations are useful. Finally, we have been talking to the Japanese government officials about the FDA mechanism, and by quantifying the trade-off that we have just mentioned, we have a better chance of persuading them to use the FDA mechanism. Since the practicality of the theory is one of our main goals, we view simulations as a useful analysis to implement.

As we have reviewed in the paper, the JRMP mechanism was introduced in 2009, and after that, hospitals are asked to gradually decrease their respective capacities, to eventually match the total capacities to the planned regional cap. Since the government publicizes only the reduced capacities, we use the hospital capacities in the data just before 2009. More specifically, we use the data from 2007, because the government specified the regional cap based on that years data. For consistency, we also use other parts of information from the data of the same year. All the data used here can be obtained at the webpage of Japanese Medical Residency Program (http://www.jrmp.jp). The data are in Japanese.

E.1. **Simulation Method.** We obtain the real data of regional caps of all 47 prefectures, the capacity and the region (prefecture) of each hospital.

Using these data, we set the target capacity for each hospital following the description in footnote 13 of the paper. That is, if the sum of the advertised positions in the hospitals region is no more than the regional cap, then the hospitals target capacity is equal to the number of its advertised positions; Otherwise, the target is given by the advertised

number times the fraction of the regional cap over the total number of the advertised positions in the region.[88]

The market size. In the simulation, the numbers of doctors and hospitals are 8,291 and 1,357, respectively, which are the number of doctors and hospitals that actually submitted preference lists in the Japanese residency match in 2007.

Preference lists. We do not have the actual data on submitted or real preferences of the doctors and hospitals, so given the above information we generated preferences and ran the simulation. Fortunately, various public data enabled us to set parameters that mimic the Japanese case, as we explain below.

(1) **Doctors.** We obtain the actual data on the distribution of the length of preference lists of doctors (i.e., the number of hospitals listed in the submitted preference list of each doctor) up to length 8. In the data, the number of doctors who listed k hospitals is not available for $k \geq 9$, while the total number of doctors who listed 9 or more hospitals is available. We also obtain from the data the average number of hospitals listed, which is 3.48.

For the doctors who list 9 or more hospitals in their preference list, we used the truncated exponential distribution such that (i) the number of doctors who list k hospitals is 0.6 times the number of doctors who list $k - 1$ doctors, for $k = 10, 11$, (modulo integer constraints) (ii) the number of doctors at length 9 is adjusted so that the average number of listed hospitals is 3.48, the average from the real data, and (iii) the maximum length is 15. With this specification the number of doctors at length 9 is smaller than the number for length 8.

The actual data describe, for each hospital, the number of doctors who listed that hospital in their preference lists. Using these data, for each hospital h, we define p_h to be the number of doctors who listed it in their preference list divided by the sum of all those numbers across all hospitals (so that it becomes probability, i.e., the numbers sum up to one). Then each doctor with preference list length k independently draws hospitals based on this distribution $(p_h)_{h \in H}$, repeatedly k times without replacement, listing her first pick as the first choice, her second pick as the second choice, and so on.[89]

---

[88]If the resulting number is not an integer, then we round the numbers to one of the adjacent integers in such a way that the sum of the target capacities in the region is equal to the regional cap.

[89]This manner of preference generation is used in a number of matching papers, such as Kojima and Pathak (2009).

(2) **Hospitals.** Each hospital ranks doctors uniform randomly, viewing every doctor acceptable.[90]

Remark 6. There could be alternatives for this method. For example we could have each hospital always rank a doctor with a shorter preference list length than the one with a longer length and those who have the same length are ranked uniformly randomly: this might as well be closer to the real data, because those doctors who rank only a small number of hospitals may be doing so because they are confident that hospitals rank them high. However, without better data on the doctor preference, we did not have better foundation for conducting such biased data generations, hence stuck to the uniformly-random data generation. We hope that the Japanese government disclose anonymous data of preference lists.

E.2. **Simulation Results.** Medical matching has two sides, namely doctors and hospitals. In the context of the Japanese medical match, another important issue is the distributional balance of doctors in different regions. Therefore, in the following we discuss the simulation results pertinent to welfare of doctors and hospitals, and then discuss the distributional consequences of the mechanisms across regions.

E.2.1. *Doctors and Hospitals.*

(1) **The number of matched doctors.** Figure 2 shows that the cost of using the JRMP mechanism is quite significant: almost 600 doctors who are matched in the unconstrained deferred acceptance (DA) mechanism become unmatched in JRMP (1396 versus 805). However, much of the negative effects can be alleviated if we switch to FDA: the number of additional doctors who become unmatched compared to the DA is about 205 (1010 versus 805), which is only about one third of 600, the corresponding number for the JRMP mechanism. Importantly, FDA achieves this improvement while satisfying all the regional caps just like the JRMP does, while DA does not respect regional caps.

---

[90]In our simulation code, each hospital actually orders only doctors who find the hospital acceptable. This is without any consequence because none of the algorithms we consider in this paper is affected by whether a doctor who finds a hospital unacceptable is acceptable to the hospital.
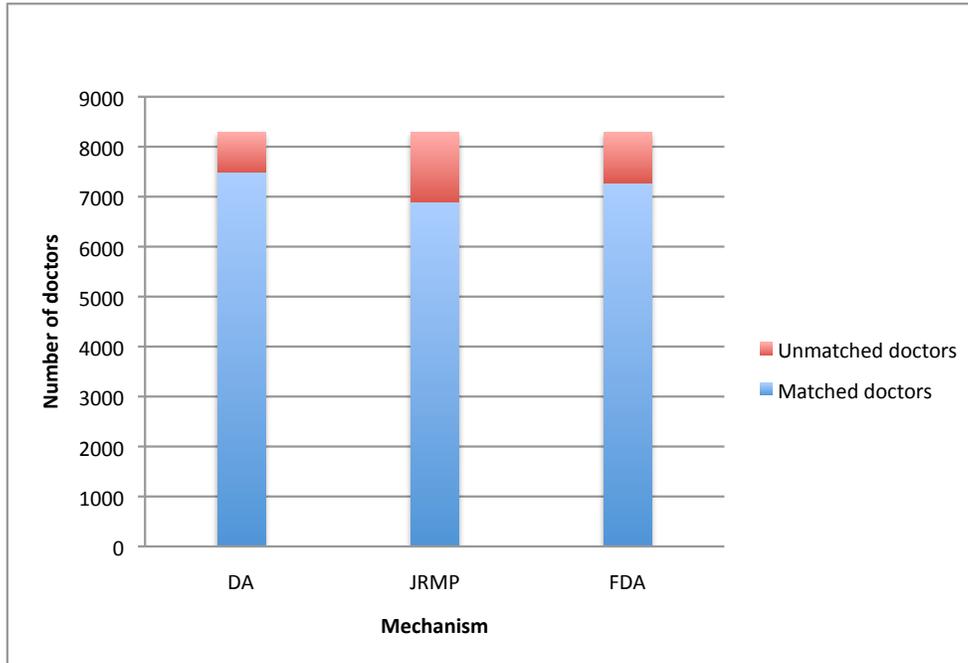
FIGURE 2. The numbers of matched doctors under different mechanisms

(2) **The number of doctors who are made strictly better off from JRMP to FDA, from FDA to DA, and from JRMP to DA.**

From/To	DA	FDA	JRMP
DA	0	0	0
FDA	606 (7.3 %)	0	0
JRMP	1547 (18.7 %)	996 (12.0 %)	0

TABLE 1.   The number of doctors who are made strictly better off from JRMP to FDA, from FDA to DA, and from JRMP to DA.

Table 1 demonstrates the numbers of doctors who become strictly better off by changing the mechanism from the one in the row to the one in the column. For example, 996 doctors become strictly better off by changing the mechanism from the JRMP mechanism to the FDA mechanism. As Theorem 3(1) predicts, there are no doctors who become strictly better off from FDA to JRMP, or from DA to the other two mechanisms, and this prediction is confirmed by the zeros in Table 1.

The theorem predicts that doctors are *weakly* better off by the change from JRMP to the other two mechanisms and from FDA to DA, but it does not pin

down how many doctors become *strictly* better off, and in general it is hard to obtain an analytical result on strict improvement without making additional assumptions on the preference distributions. This is one of the main motivations for our simulations. Since DA is unconstrained with respect to the number of doctors that can be matched to each region, large magnitudes of improvement from the other two mechanisms to DA is expected. Even so, the simulation result shows that the improvement from FDA to DA is moderate (7.3%). The result also shows that the effect of the change from JRMP to FDA–both of which are constrained by the regional cap– is large (12.0%). In view of the fact that the DA gives a (loose) upper bound of what FDA can possibly achieve, the simulation result demonstrates FDAs surprisingly large improvements in doctor welfare upon the current JRMP mechanism.

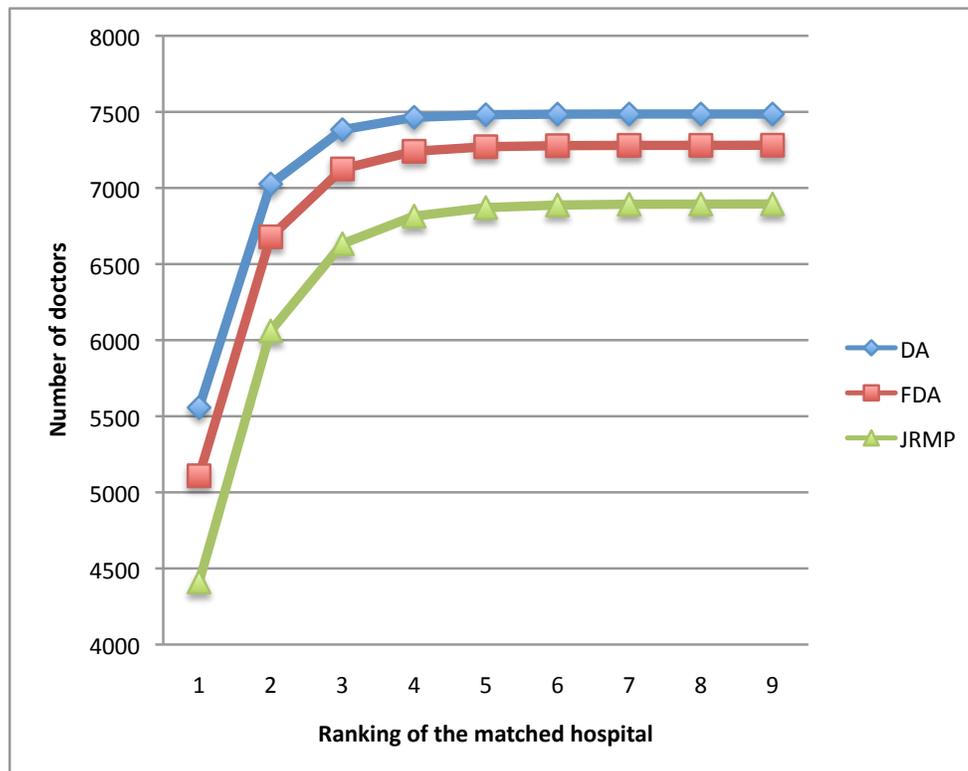(3) **Cumulative number of doctors matched to their k-th or better choices**

FIGURE 3. Cumulative number of doctors matched to their k-th or better choices.

In the above graph, the horizontal axis describes the ranking, and the vertical axis describes the number of doctors. For each of the mechanisms, we plot the

cumulative number of doctors who are matched to their k-th or better choices, for each value k in the horizontal axis.[91]

The graph confirms our prediction in Theorem 3(1). That is, the doctors are better off under DA than under FDA and under FDA than under JRMP. As we discussed in (2), the theorem does not predict magnitudes of the improvement, and this motivates simulations. Although regional caps certainly result in worse outcomes for doctors, more than a half of the loss caused by JRMP compared to the unconstrained DA can be avoided once FDA is used, even though regional caps are satisfied in FDA just as in JRMP. The effect is large: for example, about 500 more doctors are matched to their first choices under FDA compared to JRMP.

(4) **The number of hospitals that are matched to more doctors in DA, FDA, and JRMP.**

From\To	DA	FDA	JRMP
DA	0	138 (10.2%)	222 (16.4%)
FDA	104 (7.7 %)	0	158 (11.6 %)
JRMP	366 (27.0 %)	376 (27.7 %)	0

TABLE 2.   The number of hospitals that are matched to more doctors in DA, FDA, and JRMP.

The above table describes the number of the hospitals that gained more doctors in one mechanism than another. For example, 376 hospitals (27.7 % of the total of 1,357 hospital programs) are matched with more doctors in FDA than in JRMP. Unlike the corresponding table for the doctors, (2), some hospitals receives additional doctors while others lose doctors in any transition between the 3 mechanisms. But overall, the number of hospitals that receive more doctors is larger than the number of hospitals that lose doctors in transitions from JRMP to FDA or DA, which is not surprising given that more doctors are matched in FDA and DA than in JRMP (although this does not necessarily imply that more hospitals are matched under FDA or DA than under JRMP).

E.2.2. *Regions.* It is not surprising that improvement happens for some prefectures from JRMP or FDA to DA. Also, more than 3/4 of the prefectures are assigned more doctors

---

[91]In this figure, we plot only the doctors who are matched with some hospital. This is because it is what JRMP reports in their reports (and that appears to be reasonable statistics which most people care about).

under FDA than under DA. This indicates that, due to the introduction of regional caps, many prefectures became better off in terms of the number of doctors. The more relevant question is the comparison of the improvements from JRMP to FDA and from FDA to JRMP. The exact comparison of the numbers may not make too much sense, but the numbers indicate that the introduction of FDA does not make most regions get worse off. This is one finding that the theory did not tell us.

From\To	DA	FDA	JRMP
DA	0	37 (78.7%)	30 (63.8%)
FDA	6 (12.8 %)	0	25 (53.2 %)
JRMP	16 (34.0 %)	20 (42.6 %)	0

TABLE 3.   The number of regions that are assigned strictly more doctors from JRMP to FDA, from FDA to DA, and from JRMP to DA.

An issue of interest that is suppressed in the above table is the magnitude of improvements and decline in different regions.
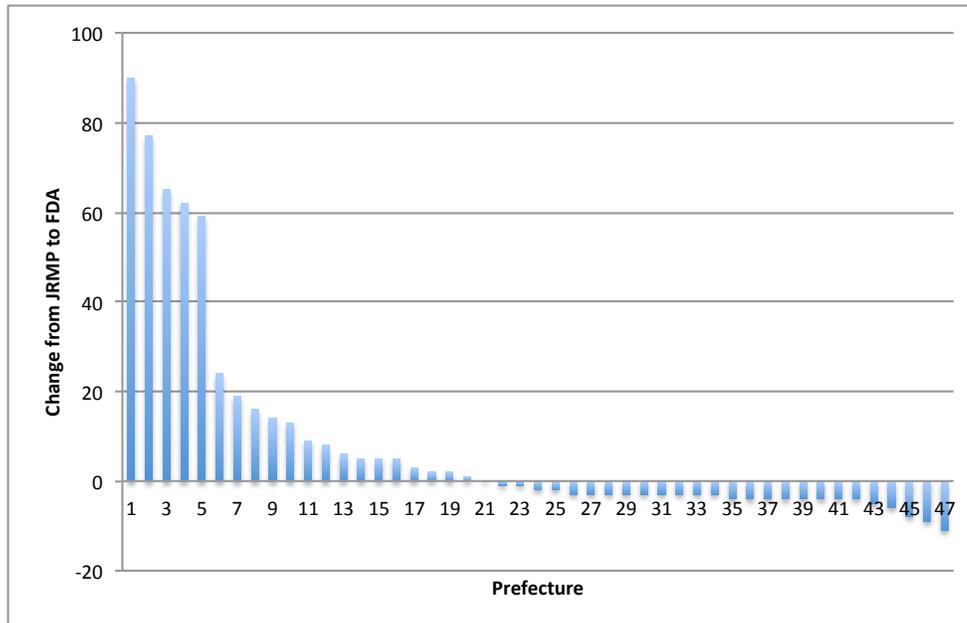

FIGURE 4. The magnitude of improvements and decline in different regions.

To better see such magnitude, in the above figure we plot the change in the number of doctors in each region, in a descending order. The graph shows that the magnitude of improvement from JRMP to FDA is large while that of decline is small. To take some

numbers, the maximum improvement is 90 while the maximum decline is 11. The area above the positive region is larger than the one in the negative region, which is consistent with our overall finding that FDA assigns about 400 more doctors in hospitals than JRMP does.

This graph is, however, silent about the distributional consequences across regions. This issue appears to be one of the main concerns in Japan. To study this issue, the figure below plots which regions become better off and which become worse off in the transition from JRMP to FDA.
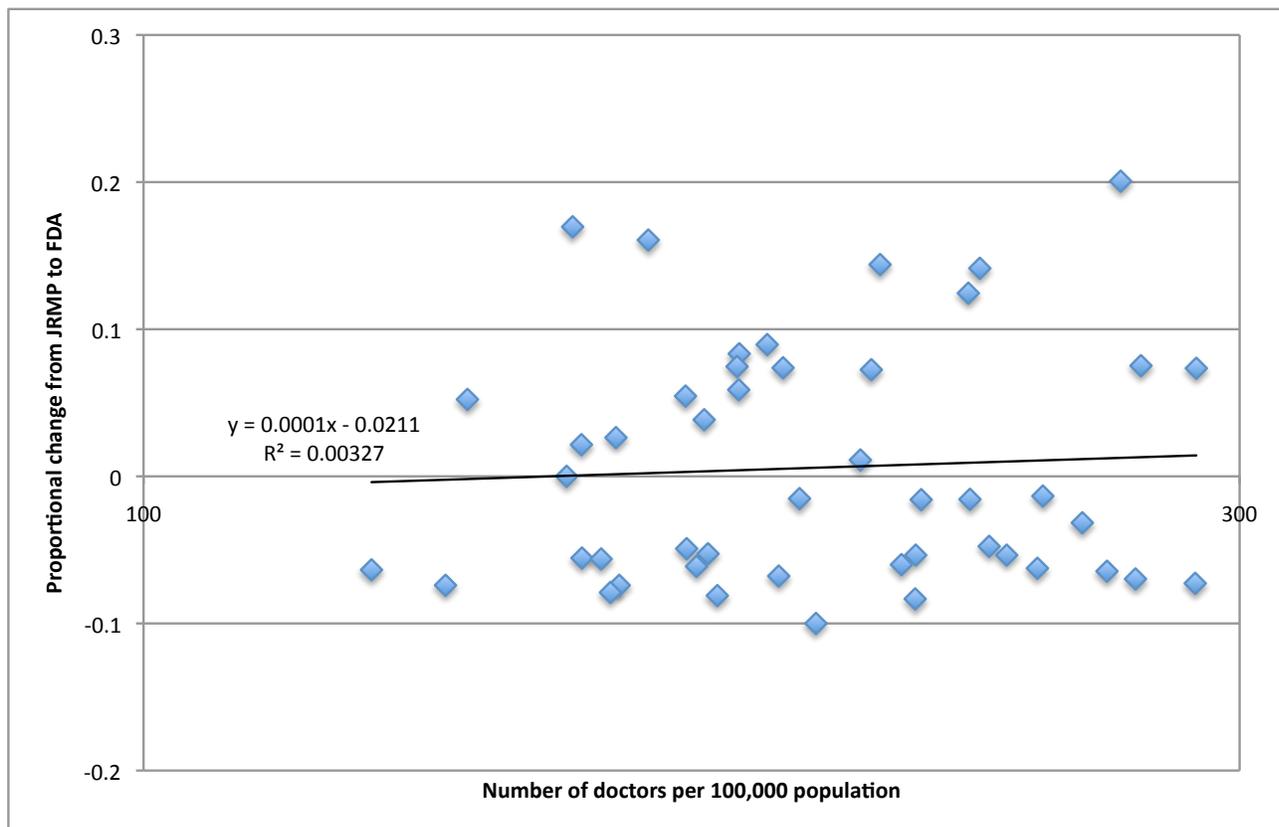

The horizontal axis measures the number of doctors (not limited to residents) per 100,000 population from 2006, which is the latest data before 2007 that is available to us, and the vertical axis describes the proportional increase of doctor assignment in the region caused by the change of the mechanism from JRMP to FDA. The motivation for this figure is to use the number of doctors per capita as a proxy for how popular each prefecture is among doctors, and study whether there is any redistribution of doctors between popular and unpopular areas.

FIGURE 5. Distributional consequence of the change from JRMP to FDA.

This figure suggests that there is virtually no adverse distributional consequence against rural areas. The linear regression suggests only a slight amount of positive relation between popularity of the prefectures and improvement/decline of doctor assignment, and indeed the R^2 value is as low as 0.00327, suggesting that there is virtually no statistical correlation between the improvement/decline of doctor assignment and how popular the area is.

Appendix F. A Generalization for Hierarchies of Regions

This section provides a generalization of the model in Appendix B: we consider the situation where there is a hierarchy of regional caps. For instance, one could consider a hierarchy of regional caps, say one cap for a prefecture and one for each district within the prefecture. Or the policy maker may desire to regulate the total number of doctors practicing in each specialty in each prefecture. We show that a generalization of the flexible deferred acceptance mechanism induces a stable matching appropriately defined.

The set of **regions** R is a subset of $2^H \setminus \{\emptyset\}$ such that $\{h\} \in R$ for all $h \in H$ and $H \in R$ (the region H is called the grand region). Further, we assume that the set of regions R is nested (a hierarchy), that is, $r, r' \in R$ implies $r \subseteq r'$ or $r' \subseteq r$ or $r \cap r' = \emptyset$. H_r denotes the set of hospitals in region r (thus we use H_r and r interchangeably for convenience). For each region r, there is a fixed positive integer q_r which we call the **regional cap** for r. For singleton region $\{h\}$ for each $h \in H$, we let $q_{\{h\}} = q_h$.

For any $r, r' \in R$, region r' is said to be an **immediate subregion** of r if $r' \subsetneq r$ and, for any $r'' \in R$, $r'' \subsetneq r$ implies either $r'' \cap r' = \emptyset$ or $r'' \subseteq r'$. It is straightforward to see that any non-singleton region $r \in R$ is partitioned into its immediate subregions. In the remainder, we simply refer to an immediate subregion as a **subregion**. Denote by $S(r)$ the set of subregions of r.

We say that region r is of **depth** k if $|\{r' \in R | r \subseteq r'\}| = k$. Note that the depth of a "smaller" region is larger. The standard model without regional caps can be interpreted as a model with regions of depths less than or equal to 2 (H and singleton sets), and the model in the main text of this paper has regions of depths less than or equal to 3 (H, "regions," and singleton sets), both with q_H sufficiently large.

Below is an example in which the set of regions forms a hierarchy.

Example 16. There are 6 hospitals, $h_1, h_2, \ldots, h_6$. The regions are

$$R = \{H, r_1, r_2, r_3, r_4, \{h_1\}, \{h_2\}, \{h_3\}, \{h_4\}, \{h_5\}, \{h_6\}\},$$

where $r_1 = \{h_1, h_2\}$, $r_2 = \{h_3, h_4, h_5, h_6\}$, $r_3 = \{h_3, h_4\}$, and $r_4 = \{h_5, h_6\}$. See Figure 6 for a graphical representation. In this example, r_1 and r_2 are the (immediate) subregions
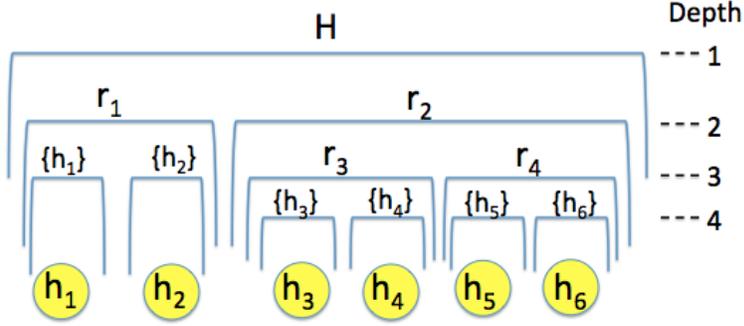

FIGURE 6. A hierarchy of regions in Example 16.

of H, r_3 and r_4 are the (immediate) subregions of r_2, and each singleton region is an (immediate) subregion of r_1 or r_3 or r_4. The depths of regions are as depicted in the figure. For example, the depth of H is 1, that of r_1 is 2, that of $\{h_1\}$ is 3, and that of $\{h_5\}$ is 4. $\square$

Let $\succeq_r$ be a weak ordering over nonnegative-valued integer vectors $W_r := \{w = (w_{r'})_{r' \in S(r)} | w_{r'} \in \mathbb{Z}_+\}$. That is, $\succeq_r$ is a binary relation that is complete and transitive (but not necessarily antisymmetric). We write $w \succ_r w'$ if and only if $w \succeq_r w'$ holds but $w' \succeq_r w$ does not. Vectors such as w and w' will be interpreted to be supplies of acceptable doctors to regions that partition r, but they will only specify how many acceptable doctors apply to each subregion and no information is given as to who these doctors are. Given $\succeq_r$, a function

$$\tilde{\mathrm{Ch}}_r : W_r \times \{0, 1, 2, \ldots, q_r\} \to W_r$$

is an **associated quasi choice rule** if $\tilde{\mathrm{Ch}}_r(w; t) \in \arg\max_{\succeq_r} \{w' | w' \leq w, \sum_{r' \in S(r)} w'_{r'} \leq t\}$ for any non-negative integer vector $w = (w_{r'})_{r' \in S(r)}$ and non-negative integer $t \leq q_r$.[92] Intuitively, $\tilde{\mathrm{Ch}}_r(w, t)$ is the best vector of numbers of doctors allocated to subregions of r given a vector of numbers w under the constraint that the sum of the number of doctors cannot exceed the quota t.

We assume that the regional preferences $\succeq_r$ satisfy $w \succ_r w'$ if $w' \lneq w$. This condition formalizes the idea that region r prefers to fill as many positions in its subregions as

---

[92] For any two vectors $w = (w_{r'})_{r' \in S(r)}$ and $w' = (w'_{r'})_{r' \in S(r)}$, we write $w \leq w'$ if and only if $w_{r'} \leq w'_{r'}$ for all $r' \in S(r)$. We write $w \lneq w'$ if and only if $w \leq w'$ and $w_{r'} < w'_{r'}$ for at least one $r' \in S(r)$. For any $W'_r \subseteq W_r$, $\arg\max_{\succeq_r} W'_r$ is the set of vectors $w \in W'_r$ such that $w \succeq_r w'$ for all $w' \in W'_r$.

possible. This requirement implies that any associated quasi choice rule is **acceptant** in the sense that, for each w and t, if there exists $r' \in S(r)$ such that $[\tilde{\text{Ch}}_r(w; t,)]_{r'} < w_{r'}$, then $\sum_{r' \in S(r)} [\tilde{\text{Ch}}_{r'}(w; t)]_{r'} = t$. This captures the idea that the social planner should not waste caps allocated to the region.[93]

We now define a restriction on preferences that we will maintain throughout our analysis.

Definition 9. The weak ordering $\succeq_r$ is **substitutable** if there exists an associated quasi choice rule $\tilde{\text{Ch}}_r$ that satisfies

$$w \leq w' \text{ and } t \geq t' \Rightarrow \tilde{\text{Ch}}_r(w; t) \geq \tilde{\text{Ch}}_r(w'; t') \wedge w.$$

Remark 7. A number of remarks on the concept of substitutability are in order. First, the condition in the definition of substitutability can be decomposed into two parts, as follows:

(F.1) $$w \leq w' \Rightarrow \tilde{\text{Ch}}_r(w; t) \geq \tilde{\text{Ch}}_r(w'; t) \wedge w, \text{ and}$$

(F.2) $$t \geq t' \Rightarrow \tilde{\text{Ch}}_r(w; t) \geq \tilde{\text{Ch}}_r(w; t').$$

Condition (F.1) imposes a condition on the quasi choice rule for different vectors w and w' with a fixed parameter t while Condition (F.2) places restrictions for different parameters t and t' with a fixed vector w. The former condition is similar to the standard substitutability condition except that it deals with multiunit supplies (that is, coefficients in w can take integers different from 0 or 1).[94] The latter condition may appear less familiar, and it requires that the choice increase (in the standard vector sense) if the allocated quota is increased. Conditions (F.1) and (F.2) are independent from each other. One might suspect that these conditions are related to responsiveness of preferences, but these conditions do no imply responsiveness. In Appendix G we provide examples to distinguish these conditions.

Second, Condition (F.1) is equivalent to

(F.3) $$w \leq w' \Rightarrow [\tilde{\text{Ch}}_r(w; t)]_{r'} \geq \min\{[\tilde{\text{Ch}}_r(w'; t)]_{r'}, w_{r'}\} \text{ for every } r' \in S(r).$$

---

[93]This condition is a variant of the concept of acceptance due to Kojima and Manea (2009).

[94]Third, Condition (F.1) is analogous to *persistence* by Alkan and Gale (2003), who define the condition on a choice function in a slightly different context. While our condition is similar to substitutability as defined in standard matching models (see Chapter 6 of Roth and Sotomayor (1990) for instance), there are two differences: (i) it is now defined on a region as opposed to a hospital, and (ii) it is defined over vectors that only specify how many doctors apply to hospitals in the region, and it does not distinguish different doctors.

This condition says that, when the supply of doctors is increased, the number of accepted doctors at a hospital can increase only when the hospital has accepted all acceptable doctors under the original supply profile. Formally, condition (F.3) is equivalent to

$$\text{(F.4)} \qquad w \leq w' \text{ and } [\tilde{\text{Ch}}_r(w;t)]_{r'} < [\tilde{\text{Ch}}_r(w';t)]_{r'} \Rightarrow [\tilde{\text{Ch}}_r(w,t)]_{r'} = w_{r'}.$$

To see that condition (F.3) implies condition (F.4), suppose that $w \leq w'$ and $[\tilde{\text{Ch}}_r(w;t)]_{r'} < [\tilde{\text{Ch}}_r(w';t)]_{r'}$. These assumptions and condition (F.3) imply $[\tilde{\text{Ch}}_r(w;t)]_{r'} \geq w_{r'}$. Since $[\tilde{\text{Ch}}_r(w;t)]_{r'} \leq w_{r'}$ holds by the definition of $\tilde{\text{Ch}}_r$, this implies $[\tilde{\text{Ch}}_r(w;t)]_{r'} = w_{r'}$. To see that condition (F.4) implies condition (F.3), suppose that $w \leq w'$. If $[\tilde{\text{Ch}}_r(w;t)]_{r'} \geq [\tilde{\text{Ch}}_r(w';t)]_{r'}$, the conclusion of (F.3) is trivially satisfied. If $[\tilde{\text{Ch}}_r(w;t)]_{r'} < [\tilde{\text{Ch}}_r(w';t)]_{r'}$, then condition (F.4) implies $[\tilde{\text{Ch}}_r(w;t,)]_{r'} = w_{r'}$, thus the conclusion of (F.3) is satisfied.

Finally, substitutability implies the following natural property that we call "consistency": A quasi choice rule $\tilde{\text{Ch}}_r$ is said to be **consistent** if for any t, $\tilde{\text{Ch}}_r(w;t) \leq w' \leq w \Rightarrow \tilde{\text{Ch}}_r(w';t) = \tilde{\text{Ch}}_r(w;t)$.[95] Consistency requires that, if $\tilde{\text{Ch}}_r(w;t)$ is chosen at w and the supply decreases to $w' \leq w$ but $\tilde{\text{Ch}}_r(w;t)$ is still available under w', then the same choice $\tilde{\text{Ch}}_r(w;t)$ should be made under w' as well. Note that there may be more than one consistent quasi choice rule associated with a given weak ordering $\succeq_r$ because the set $\arg\max_{\succeq_r}\{w'|w' \leq w, \sum_{r' \in S(r)} w'_{r'} \leq t\}$ may not be a singleton for some $\succeq_r$, w, and t. Note also that there always exists a consistent quasi choice rule.[96] We relegate the proof for the fact that substitutability implies consistency to Appendix G. $\qquad\square$

Now we define the notion of stability and the **(generalized) flexible deferred acceptance algorithm** in our context where R is a hierarchy. Let $SC(h, h') \in R$ be the **smallest common region** of hospitals h and h', that is, it is a region $r \in R$ with the property that $h, h' \in H_r$, and there is no $r' \in R$ with $r' \subsetneq r$ such that $h, h' \in H_{r'}$. Given $(\succeq_r)_{r \in R}$, stability is defined as follows.

Definition 10. A matching μ is **stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then there exists $r \in R$ with $h \in H_r$ such that (i) $|\mu_r| = q_r$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and

(iii') either $\mu_d \notin H_r$ or $(w_{r'})_{r' \in S(SC(h,\mu_d))} \succeq_{SC(h,\mu_d)} (w'_{r'})_{r' \in S(SC(h,\mu_d))}$,

where $w_{r'} = \sum_{h' \in r'} |\mu_{h'}|$ for all $r' \in S(r)$ and $w'_{r_h} = w_{r_h} + 1$, $w'_{r_d} = w_{r_d} - 1$ and $w'_{r'} = w_{r'}$ for all other $r' \in S(r)$ where r_h and r_d are subregions of r such that $h \in r'_h$, and $\mu_d \in r_d$.

---

[95]More precisely, it is Condition (F.1) of substitutability that implies consistency.

[96]To see this point consider preferences $\succeq'_r$ such that $w \succ'_r w'$ if $w \succ_r w'$ and $w = w'$ if $w \succeq'_r w'$ and $w' \succeq'_r w$. The quasi choice rule that chooses (the unique element of) $\arg\max_{\succeq'_r}\{w'|w' \leq w, \sum_{r' \in S(r)} w'_{r'} \leq t\}$ for each w is clearly consistent.

Remark 8. Condition (iii') of this definition captures the idea behind stability in Definition 4 in that a region's preferences are invoked when a doctor moves within a region whose regional cap is binding (region r in the definition). However, when r is a strict superset of $SC(h, \mu_d)$, we do not invoke region r's regional preferences, but the preferences of $SC(h, \mu_d)$.[97] The use of preferences of $SC(h, \mu_d)$ reflects the following idea: if the regional cap at r is binding then holding fixed the number of doctors matched in r but not in $SC(h, \mu_d)$, there is essentially a binding cap for $SC(h, \mu_d)$. This motivates our use of the regional preferences of $SC(h, \mu_d)$. The reason for not using preferences of r (or any region between r and $SC(h, \mu_d)$) is that the movement of a doctor within the region $SC(h, \mu_d)$ does not affect the distribution of doctors on which preferences of r (or regions of any smaller depth than $SC(h, \mu_d)$) are defined. $\square$

We proceed to define a quasi choice rule for the "hospital side," denoted $\tilde{\text{Ch}}$: Let $\tilde{q}_H = q_H$. Given $w = (w_h)_{h \in H}$, we define $v_{\{h\}}^w = \min\{w_h, q_h\}$, and inductively define $v_r^w = \min\{\sum_{r' \in S(r)} v_{r'}^w, q_r\}$. Thus, v_r^w is the maximum number that the input w can allocate to its subregions given the feasibility constraints that w and regional caps of subregions of r impose. Note that v_r^w is weakly increasing in w, that is, $w \geq w'$ implies $v_r^w \geq v_r^{w'}$.

We inductively define $\tilde{\text{Ch}}(w)$ following a procedure starting from Step 1, where Step k for general k is as follows:

> Step k: If all the regions of depth k are singletons, then let $\tilde{\text{Ch}}(w) = (\tilde{q}_{\{h\}}^w)_{h \in H}$ and stop the procedure. For each nonsingleton region r of depth k, set $\tilde{q}_{r'}^w = [\tilde{\text{Ch}}_r((v_{r''}^w)_{r'' \in S(r)}; \tilde{q}_r^w)]_{r'}$ for each subregion r' of r. Go to Step $k + 1$.

Assume that $\succeq_r$ is substitutable for every region r. Now we are ready to define a generalized version of the **flexible deferred acceptance algorithm**:

For each region r, fix an associated quasi choice rule $\tilde{\text{Ch}}_r$ for which conditions (F.1) and (F.2) are satisfied (note that the assumption that $\succeq_r$ is substitutable assures the existence of such a quasi choice rule.)

(1) Begin with an empty matching, that is, a matching μ such that $\mu_d = \emptyset$ for all $d \in D$.

(2) Choose a doctor d arbitrarily who is currently not tentatively matched to any hospital and who has not applied to all acceptable hospitals yet. If such a doctor does not exist, then terminate the algorithm.

---

[97] It is important that we allow r to be a strict superset of $SC(h, \mu_d)$. Example 22 in Appendix H points out that, if we further require $r \subseteq SC(h, \mu_d)$ in Definition 10, then there does not need to exist a matching that satisfies this stronger notion of stability.

(3) Let d apply to the most preferred hospital $\bar{h}$ at $\succ_d$ among the hospitals that have not rejected d so far. If d is unacceptable to $\bar{h}$, then reject this doctor and go back to Step 2. Otherwise, define vector $w = (w_h)_{h \in H}$ by

 (a) $w_{\bar{h}}$ is the number of doctors currently held at $\bar{h}$ plus one, and

 (b) w_h is the number of doctors currently held at h if $h \neq \bar{h}$.

(4) Each hospital $h \in H$ considers the new applicant d (if $h = \bar{h}$) *and* doctors who are temporarily held from the previous step together. It holds its $[\tilde{\mathrm{Ch}}(w)]_h$ most preferred applicants among them temporarily and rejects the rest (so doctors held at this step may be rejected in later steps). Go back to Step 2.

We define the **(generalized) flexible deferred acceptance mechanism** to be a mechanism that produces, for each input, the matching given at the termination of the above algorithm.[98]

F.1. **Associated Matching Model with Contracts.** It is useful to relate our model to a (many-to-many) matching model with contracts (Hatfield and Milgrom, 2005). Let there be two types of agents, doctors in D and the "hospital side" (thus there are $|D| + 1$ agents in total). Note that we regard the hospital side, instead of each hospital, as an agent in this model. There is a set of contracts $X = D \times H$.

We assume that, for each doctor d, any set of contracts with cardinality two or more is unacceptable, that is, a doctor can sign at most one contract. For each doctor d, her preferences $\succ_d$ over $(\{d\} \times H) \cup \{\emptyset\}$ are given as follows.[99] We assume $(d, h) \succ_d (d, h')$ in this model if and only if $h \succ_d h'$ in the original model, and $(d, h) \succ_d \emptyset$ in this model if and only if $h \succ_d \emptyset$ in the original model.

For the hospital side, we assume that it has preferences $\succeq$ and its associated choice rule $\mathrm{Ch}(\cdot)$ over all subsets of $D \times H$. For any $X' \subset D \times H$, let $w(X') := (w_h(X'))_{h \in H}$ be the vector such that $w_h(X') = |\{(d, h) \in X' | d \succ_h \emptyset\}|$. For each X', the chosen set of contracts $\mathrm{Ch}(X')$ is defined by

$$\mathrm{Ch}(X') = \bigcup_{h \in H} \left\{ (d, h) \in X' \ \middle| \ |\{d' \in D | (d', h) \in X', d' \succeq_h d\}| \leq [\tilde{\mathrm{Ch}}(w(X'))]_h \right\}.$$

That is, each hospital $h \in H$ chooses its $[\tilde{\mathrm{Ch}}(w(X'))]_h$ most preferred contracts from acceptable contracts in X'.

---

[98]Note that this algorithm terminates in a finite number of steps. Note also that, as will be verified via Proposition 12, the outcome of the algorithm is independent of the order in which doctors make their applications during the algorithm.

[99]We abuse notation and use the same notation $\succ_d$ for preferences of doctor d both in the original model and in the associated model with contracts.

Definition 11 (Hatfield and Milgrom (2005))**.** Choice rule $\mathrm{Ch}(\cdot)$ satisfies the **substitutes condition** if there do not exist contracts $x, x' \in X$ and a set of contracts $X' \subseteq X$ such that $x' \notin \mathrm{Ch}(X' \cup \{x'\})$ and $x' \in \mathrm{Ch}(X' \cup \{x, x'\})$.

In other words, contracts are substitutes if adding a contract to the choice set never induces a region to choose a contract it previously rejected. Hatfield and Milgrom (2005) show that there exists a stable allocation (defined in Definition 13) when contracts are substitutes for the hospital side.

Definition 12 (Hatfield and Milgrom (2005))**.** Choice rule $\mathrm{Ch}(\cdot)$ satisfies the **law of aggregate demand** if for all $X' \subseteq X'' \subseteq X$, $|\mathrm{Ch}(X')| \leq |\mathrm{Ch}(X'')|$.

Proposition 10. *Suppose that $\succeq_r$ is substitutable for all $r \in R$.*

(1) *Choice rule $\mathrm{Ch}(\cdot)$ defined above satisfies the substitutes condition*
(2) *Choice rule $\mathrm{Ch}(\cdot)$ defined above satisfies the law of aggregate demand.*

Proof. **Part 1.** Fix $X' \subset X$. Suppose to the contrary, i.e., that there exist X', (d, h) and (d', h') such that $(d', h') \notin \mathrm{Ch}(X' \cup \{(d', h')\})$ and $(d', h') \in \mathrm{Ch}(X' \cup \{(d, h), (d', h')\})$. We will lead to a contradiction.

Let $w' = w(X' \cup \{(d', h')\})$ and $w'' = w(X' \cup \{(d, h), (d', h')\})$. The proof consists of three steps.

Step 1: In this step we observe that $\tilde{q}_{\{h'\}}^{w'} < \tilde{q}_{\{h'\}}^{w''}$. To see this, note that otherwise we would have $\tilde{q}_{\{h'\}}^{w'} \geq \tilde{q}_{\{h'\}}^{w''}$, hence by the definition of Ch we must have $[\mathrm{Ch}(X' \cup \{(d', h')\})]_{h'} \supseteq [\mathrm{Ch}(X' \cup \{(d, h), (d', h')\})]_{h'} \setminus \{(d, h)\}$. This contradicts $(d', h') \notin \mathrm{Ch}(X' \cup \{(d', h')\})$ and $(d', h') \in \mathrm{Ch}(X' \cup \{(d, h), (d', h')\})$.

Step 2: Consider any r such that $h' \in r$. Let $\tilde{q}_r^{w'}$ and $\tilde{q}_r^{w''}$ be as defined in the procedure to compute $\tilde{\mathrm{Ch}}(w')$ and $\tilde{\mathrm{Ch}}(w'')$, respectively. Let $r' \in S(r)$ be the subregion such that $h' \in r'$. Suppose $\tilde{q}_{r'}^{w'} < \tilde{q}_{r'}^{w''}$. We will show that $\tilde{q}_r^{w'} < \tilde{q}_r^{w''}$. To see this, suppose the contrary, i.e., that $\tilde{q}_r^{w'} \geq \tilde{q}_r^{w''}$. Let $v' := (v_{r''}^{w'})_{r'' \in S(r)}$ and $v'' := (v_{r''}^{w''})_{r'' \in S(r)}$. Since $w' \leq w''$ and $v_{r''}^w$ is weakly increasing in w for any region r'', it follows that $v' \leq v''$. This and substitutability of $\succeq_r$ imply

$$[\tilde{\mathrm{Ch}}_r(v'; \tilde{q}_r^{w'})]_{r'} \geq \min\{[\tilde{\mathrm{Ch}}_r(v''; \tilde{q}_r^{w''})]_{r'}, v_{r'}'\}.$$

Since we assume $\tilde{q}_{r'}^{w'} < \tilde{q}_{r'}^{w''}$, or equivalently

$$[\tilde{\mathrm{Ch}}_r(v'; \tilde{q}_r^{w'})]_{r'} < [\tilde{\mathrm{Ch}}_r(v''; \tilde{q}_r^{w''})]_{r'},$$

this means $[\tilde{\mathrm{Ch}}_r(v'; \tilde{q}_r^{w'})]_{r'} \geq v_{r'}'$. But then by $[\tilde{\mathrm{Ch}}_r(v'; \tilde{q}_r^{w'})]_{r'} \leq v_{r'}'$ (from the definition of $\tilde{\mathrm{Ch}}$) we have $[\tilde{\mathrm{Ch}}_r(v'; \tilde{q}_r^{w'})]_{r'} = v_{r'}'$. But this contradicts the assumption that $(d', h') \notin$

$\mathrm{Ch}(X' \cup \{(d', h')\})$, while d' is acceptable to h' (because $(d', h') \in \mathrm{Ch}(X' \cup \{(d, h), (d', h')\})$). Thus we must have that $\tilde{q}_r^{w'} < \tilde{q}_r^{w''}$.

Step 3: Step 1 and an iterative use of Step 2 imply that $\tilde{q}_H^{w'} < \tilde{q}_H^{w''}$. But we specified $\tilde{q}_H^w$ for any w to be equal to q_H, so this is a contradiction.

Part 2. To show that Ch satisfies the law of aggregate demand, let $X' \subseteq X$ and (d, h) be a contract such that $d \succ_h \emptyset$. We shall show that $|\mathrm{Ch}(X')| \leq |\mathrm{Ch}(X' \cup \{(d, h)\})|$. To show this, denote $w = w(X')$ and $w' = w(X' \cup \{(d, h)\})$. By definition of $w(\cdot)$, we have that $w'_h = w_h + 1$ and $w'_{h'} = w_{h'}$ for all $h' \neq h$. Consider the following cases.

(1) Suppose $\sum_{r' \in S(r)} v_{r'}^w \geq q_r$ for some $r \in R$ such that $h \in r$. Then we have:

Claim 2. $v_{r'}^{w'} = v_{r'}^w$ unless $r' \subsetneq r$.

Proof. Let r' be a region that does not satisfy $r' \subsetneq r$. First, note that if $r' \cap r = \emptyset$, then the conclusion holds by the definitions of $v_{r'}^w$ and $v_{r'}^{w'}$ because $w'_{h'} = w_{h'}$ for any $h' \notin r$. Second, consider r' such that $r \subseteq r'$ (since R is hierarchical, these cases exhaust all possibilities). Since $v_r^w = \min\{\sum_{r' \in S(r)} v_{r'}^w, q_r\}$, the assumption $\sum_{r' \in S(r)} v_{r'}^w \geq q_r$ implies $v_r(w) = q_r$. By the same argument, we also obtain $v_r(w') = q_r$. Thus, for any r' such that $r \subseteq r'$, we inductively obtain $v_{r'}^{w'} = v_{r'}^w$. $\square$

The relation $v_{r'}^{w'} = v_{r'}^w$ for all $r' \subsetneq r$ implies that, together with the construction of $\tilde{\mathrm{Ch}}$,

(F.5) $$[\tilde{\mathrm{Ch}}(w')]_{h'} = [\tilde{\mathrm{Ch}}(w)]_{h'} \text{ for any } h' \notin r.$$

To consider hospitals in r, first observe that r satisfies $\sum_{r' \in S(r)} v_{r'}^w \geq q_r$ by assumption, so $v_r^w = \min\{\sum_{r' \in S(r)} v_{r'}^w, q_r\} = q_r$, and similarly $v_r^{w'} = q_r$, so $v_r^w = v_r^{w'}$. Therefore, by construction of $\tilde{\mathrm{Ch}}$, we also have $v_{r'}^w = v_{r'}^{w'}$ for any region r' such that $r \subseteq r'$. This implies $\tilde{q}_r^w = \tilde{q}_r^{w'}$, where $\tilde{q}_r^w$ and $\tilde{q}_r^{w'}$ are the assigned regional caps on r under weight vectors w and w', respectively, in the algorithm to construct $\tilde{\mathrm{Ch}}$.

Now note the following: For any $r' \in R$, since $v_{r'}^w$ is defined as $\min\{\sum_{r'' \in S(r')} v_{r''}^w, q_{r'}\}$ and all regional preferences are acceptant, the entire assigned regional cap $\tilde{q}_{r'}^w$ is allocated to some subregion of r', that is, $\bar{q}_{r'}^w = \sum_{r'' \in S(r')} \bar{q}_{r''}^w$. Similarly we also have $\bar{q}_{r'}^{w'} = \sum_{r'' \in S(r')} \bar{q}_{r''}^{w'}$. This is the case for not only for $r' = r$ but also for all subregions of r, their further subregions, and so forth. Going forward until this

reasoning reaches the singleton sets, we obtain relation

(F.6) $$\sum_{h'\in r}[\tilde{\text{Ch}}(w')]_{h'} = \sum_{h'\in r}[\tilde{\text{Ch}}(w)]_{h'}.$$

By (F.5) and (F.6), we conclude that

$$|\text{Ch}(X')| = \sum_{h'\in H}[\tilde{\text{Ch}}(w)]_{h'} = \sum_{h'\in H}[\tilde{\text{Ch}}(w')]_{h'} = |\text{Ch}(X'\cup\{(d,h)\})|,$$

completing the proof for this case.

(2) Suppose $\sum_{r'\in S(r)} v_{r'}^{w} < q_r$ for all $r \in R$ such that $h \in r$. Then the regional cap for r is not binding for any r such that $h \in r$, so we have

(F.7) $$[\tilde{\text{Ch}}(w')]_h = [\tilde{\text{Ch}}(w)]_h + 1.$$

In addition, the following claim holds.

Claim 3. $[\tilde{\text{Ch}}(w')]_{h'} = [\tilde{\text{Ch}}(w)]_{h'}$, *for all* $h' \neq h$.

Proof. First, note that $v_r^{w'} = v_r^{w} + 1$ for all r such that $h \in r$ because the regional cap for r is not binding for any such r. Then, consider the largest region H. By assumption, q_H has not been reached under w, that is, $\sum_{r'\in S(H)} v_{r'}^{w} < q_H$. Thus, since $\tilde{\text{Ch}}_H$ is acceptant, the entire vector $(v_{r'}(w))_{r'\in S(H)}$ is accepted by $\tilde{\text{Ch}}_H$, that is, $\tilde{q}_{r'}^{w} = v_{r'}^{w}$. Hence, for any $r' \in S(H)$ such that $h \notin r'$, both its assigned regional cap and all v's in their regions are identical under w and w', that is, $\tilde{q}_{r'}^{w} = \tilde{q}_{r'}^{w'}$ and $w_{h'}' = w_{h'}$ for all $h' \in r'$. So, for any hospital $h' \in r'$, the claim holds.

Now, consider $r \in S(H)$ such that $h \in r$. By the above argument, the assigned regional cap has increased by one in w' compared to w. But since r's regional cap q_r has not been binding under w, all the v's in the subregions of r are accepted in both w and w'. This means that (1) for each subregion r' of r such that $h \notin r'$, it gets the same assigned regional cap and v's, so the conclusion of the claim holds for these regions, and (2) for the subregion r' of r such that $h \in r'$, its assigned regional cap is increased by one in w' compared to w, and its regional cap $q_{r'}$ has not been binding. And (2) guarantees that we can follow the same argument inductively, so the conclusion holds for all $h \neq h'$. $\square$

By equation (F.7) and Claim 3, we obtain

$$|\text{Ch}(X'\cup\{(d,h)\})| = \sum_{h'\in H}[\tilde{\text{Ch}}(w')]_{h'} = \sum_{h'\in H}[\tilde{\text{Ch}}(w)]_{h'} + 1 = |\text{Ch}(X')| + 1,$$

so we obtain $|\text{Ch}(X'\cup\{(d,h)\})| > |\text{Ch}(X')|$, completing the proof.

$\square$

A subset X' of $X = D \times H$ is said to be **individually rational** if (1) for any $d \in D$, $|\{(d, h) \in X' | h \in H\}| \leq 1$, and if $(d, h) \in X'$ then $h \succ_d \emptyset$, and (2) $\text{Ch}(X') = X'$.

Definition 13. A set of contracts $X' \subseteq X$ is a **stable allocation** if

(1) it is individually rational, and

(2) there exists no hospital $h \in H$ and a doctor $d \in D$ such that $(d, h) \succ_d x$ and $(d, h) \in \text{Ch}(X' \cup \{(d, h)\})$, where x is the contract that d receives at X' if any and $\emptyset$ otherwise.

When condition (2) is violated by some (d, h), we say that (d, h) is a **block** of X'.

Given any individually rational set of contracts X', define a **corresponding matching** $\mu(X')$ in the original model by setting $\mu_d(X') = h$ if and only if $(d, h) \in X'$ and $\mu_d(X') = \emptyset$ if and only if no contract associated with d is in X'. For any individually rational X', $\mu(X')$ is well-defined because each doctor receives at most one contract at such X'.

Proposition 11. *Suppose that $\succeq_r$ is substitutable for all $r \in R$. If X' is a stable allocation in the associated model with contracts, then the corresponding matching $\mu(X')$ is a stable matching in the original model.*

Proof. Suppose that X' is a stable allocation in the associated model with contracts and denote $\mu := \mu(X')$. Individual rationality of μ is obvious from the construction of μ. Suppose that (d, h) is a blocking pair of μ. By the definition of stability, it suffices to show that there exists a region r that includes h such that the following conditions (F.8), (F.9), and $\mu_d \notin H_r$ hold, or (F.8), (F.9), (F.10), and $h, \mu_d \in r$ hold:

(F.8) $$|\mu_{H_r}| = q_r,$$

(F.9) $$d' \succ_h d \text{ for all } d' \in \mu_h,$$

(F.10) $$(w_{r''})_{r'' \in S(SC(h,\mu_d))} \succeq_{SC(h,\mu_d)} (w'_{r''})_{r'' \in S(SC(h,\mu_d))},$$

where for any region r' we write $w_{r''} = \sum_{h' \in r''} |\mu_{h'}|$ for all $r'' \in S(r')$ and $w'_{r_h} = w_{r_h} + 1$, $w'_{r_d} = w_{r_d} - 1$ and $w'_{r''} = w_{r''}$ for all other $r'' \in S(r')$ where $r_h, r_d \in S(r)$, $h \in r_h$, and $\mu_d \in r_d$. Let $w = (w_h)_{h \in H}$.

For each region r that includes h, let $w''_{r'} = w_{r'} + 1$ for r' such that $h \in r'$ and $w''_{r''} = w_{r''}$ for all other $r'' \in S(r)$. Let $w'' = (w''_h)_{h \in H}$.

Claim 4. *Condition (F.9) holds, and there exists r that includes h such that Condition (F.8) holds.*

Proof. First note that the assumption that $h \succ_d \mu_d$ implies that $(d, h) \succ_d x$ where x denotes the (possibly empty) contract that d signs under X'.

(1) Assume by contradiction that condition (F.9) is violated, that is, $d \succ_h d'$ for some $d' \in \mu_h$. First, note that $[\tilde{\mathrm{Ch}}(w'')]_h \geq [\tilde{\mathrm{Ch}}(w)]_h$. That is, weakly more contracts involving h are signed at $X' \cup (d, h)$ than at X'. This is because for any r and $r' \in S(r)$ such that $h \in r'$,

(F.11) $\qquad [\tilde{\mathrm{Ch}}_r((v_{r''}^{w''})_{r'' \in S(r)}; \tilde{q}_r)]_{r'} \geq [\tilde{\mathrm{Ch}}_r((v_{r''}^{w})_{r'' \in S(r)}; \tilde{q}_r')]_{r'}$ if $\tilde{q}_r \geq \tilde{q}_r'$.

To see this, first note that $[\tilde{\mathrm{Ch}}_r((v_{r''}^{w})_{r'' \in S(r)}; \tilde{q}_r)]_{r'} \geq [\tilde{\mathrm{Ch}}_r((v_{r''}^{w})_{r'' \in S(r)}; \tilde{q}_r')]_{r'}$ by substitutability of $\succeq_r$. Also, by consistency of $\tilde{\mathrm{Ch}}_r$ and $v_{r''}^{w''} \geq v_{r''}^{w}$ for every region r'', the inequality

$$[\tilde{\mathrm{Ch}}_r((v_{r''}^{w''})_{r'' \in S(r)}; \tilde{q}_r)]_{r'} \geq [\tilde{\mathrm{Ch}}_r((v_{r''}^{w})_{r'' \in S(r)}; \tilde{q}_r)]_{r'}$$

follows,[100] showing condition (F.11). An iterative use of condition (F.11) gives us the desired result that $[\tilde{\mathrm{Ch}}(w'')]_h \geq [\tilde{\mathrm{Ch}}(w)]_h$. This property, together with the assumptions that $d \succ_h d'$ and that $(d', h) \in X'$ imply that $(d, h) \in \mathrm{Ch}(X' \cup (d, h))$.[101] Thus, together with the above-mentioned property that $(d, h) \succ_d x$, (d, h) is a block of X' in the associated model of matching with contracts, contradicting the assumption that X' is a stable allocation.

(2) Assume by contradiction that condition (F.8) is violated, so that $|\mu_{H_r}| \neq q_r$ for every r that includes h. Then, for such r, since $|\mu_{H_r}| \leq q_r$ by the construction of μ and the assumption that X' is individually rational, it follows that $|\mu_{H_r}| < q_r$. Then $(d, h) \in \mathrm{Ch}(X' \cup (d, h))$ because,

(a) $d \succ_h \emptyset$ by assumption,

(b) since $\sum_{r' \in S(r)} w_{r'} = \sum_{h \in H_r} |\mu_h| = |\mu_{H_r}| < q_r$, it follows that $\sum_{r' \in S(r)} w_{r'}'' = \sum_{r' \in S(r)} w_{r'} + 1 \leq q_r$. This property and the fact that $\tilde{\mathrm{Ch}}_r$ is acceptant and the definition of the function $v_{r'}$ for regions r' imply that $\tilde{\mathrm{Ch}}(w'') = w''$. In

---

[100]To show this claim, let $v = (v_{r''}^{w})_{r'' \in S(r)}$ and $v'' = (v_{r''}^{w''})_{r'' \in S(r)}$ for notational simplicity and assume for contradiction that $[\tilde{\mathrm{Ch}}_r(v''; \tilde{q}_r)]_{r'} < [\tilde{\mathrm{Ch}}_r(v; \tilde{q}_r)]_{r'}$. Then, $[\tilde{\mathrm{Ch}}_r(v''; \tilde{q}_r)]_{r'} < [\tilde{\mathrm{Ch}}_r(v; \tilde{q}_r)]_{r'} \leq v_{r'}$. Moreover, since $v_{r''}'' = v_{r''}$ for every $r'' \neq r'$ by the construction of v'', it follows that $[\tilde{\mathrm{Ch}}_r(v'')]_{r''} \leq v_{r''}'' = v_{r''}$. Combining these inequalities, we have that $\tilde{\mathrm{Ch}}_r(v'') \leq v$. Also we have $v \leq v''$ by the definition of v'', so it follows that $\tilde{\mathrm{Ch}}_r(v'') \leq v \leq v''$. Thus, by consistency of $\tilde{\mathrm{Ch}}_r$, we obtain $\tilde{\mathrm{Ch}}_r(v'') = \tilde{\mathrm{Ch}}_r(v)$, a contradiction to the assumption $[\tilde{\mathrm{Ch}}_r(v'')]_{r'} < [\tilde{\mathrm{Ch}}_r(v)]_{r'}$.

[101]The proof of this claim is as follows. $\mathrm{Ch}(X')$ induces hospital h to select its $[\tilde{\mathrm{Ch}}(w)]_h$ most preferred contracts while $\mathrm{Ch}(X' \cup (d, h))$ induces h to select a weakly larger number $[\mathrm{Ch}(w'')]_h$ of its most preferred contracts. Since (d', h) is selected as one of the $[\tilde{\mathrm{Ch}}(w)]_h$ most preferred contracts for h at X' and $d \succ_h d'$, we conclude that (d, h) must be one of the $[\mathrm{Ch}(w'')]_h$ ($\geq [\tilde{\mathrm{Ch}}(w)]_h$) most preferred contracts at $X' \cup (d, h)$, thus selected at $X' \cup (d, h)$.

particular, this implies that every contract $(d', h) \in X' \cup (d, h)$ such that $d' \succ_h \emptyset$ is chosen at $\mathrm{Ch}(X' \cup (d, h))$.

Thus, together with the above-mentioned property that $(d, h) \succ_d x$, (d, h) is a block of X' in the associated model of matching with contract, contradicting the assumption that X' is a stable allocation.

$\square$

To finish the proof of the proposition suppose for contradiction that there is no r that includes h such that (F.8), (F.9), and $\mu_d \notin H_r$ hold, and that condition (F.10) fails. That is, we suppose $(w'_{r''})_{r'' \in S(SC(h,\mu_d))} \succ_{SC(h,\mu_d)} (w_{r''})_{r'' \in S(SC(h,\mu_d))}$. Then it must be the case that $[\tilde{\mathrm{Ch}}_r((v^{w''}_{r''})_{r'' \in S(SC(h,\mu_d))}; \tilde{q}^{w''}_{SC(h,\mu_d)})]_{r'} = w''_{r'} = w_{r'} + 1 = |\mu_h| + 1$, where $h \in r'$ and $\tilde{q}^{w''}_{SC(h,\mu_d)}$ is as defined in the procedure to compute $\tilde{\mathrm{Ch}}(w'')$.[102] Note that for all r' such that $h \in r'$ and $r' \subsetneq SC(h, \mu_d)$, it follows that $\mu_d \notin H_{r'}$. Also note that (F.9) is satisfied by Claim 4. Therefore we have $|\mu_{r'}| < q_{r'}$ for all $r' \subsetneq SC(h, \mu_d)$ that includes h by assumption and hence $|\mu_{r'}| + 1 \leq q_{r'}$ for all such r'. Moreover we have $d \succ_h \emptyset$, thus

$$(d, h) \in \mathrm{Ch}(X' \cup (d, h)).$$

This relationship, together with the assumption that $h \succ_d \mu_d$, and hence $(d, h) \succ_d x$, is a contradiction to the assumption that X' is stable in the associated model with contracts.

$\square$

---

[102]To show this claim, assume for contradiction that $[\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}((v^{w''}_{r''})_{r'' \in S(SC(h,\mu_d))}; \tilde{q}^{w''}_{SC(h,\mu_d)})]_{r'} \leq w_{r'}$ where $h \in r'$. Let $v := (v^w_{r''})_{r'' \in S(SC(h,\mu_d))}$ and $v'' := (v^{w''}_{r''})_{r'' \in S(SC(h,\mu_d))}$. Since $w''_{r''} = w_{r''}$ for any $r'' \neq r'$ by the definition of w'', it follows that

$$\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v''; \tilde{q}^{w''}_{SC(h,\mu_d)}) \leq (w_{r''})_{r'' \in SC(h,\mu_d)} \leq (w''_{r''})_{r'' \in SC(h,\mu_d)}.$$

But $\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v; \tilde{q}^w_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$ because X' is a stable allocation in the associated model of matching with contracts, which in particular implies $v = (w_{r''})_{r'' \in SC(h,\mu_d)}$. Since $v \leq v''$, this means that

$$\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v''; \tilde{q}^{w''}_{SC(h,\mu_d)}) \leq v \leq v''.$$

Thus by consistency of $\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}$, we obtain

$$\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v''; \tilde{q}^{w''}_{SC(h,\mu_d)}) = \tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v; \tilde{q}^{w''}_{SC(h,\mu_d)}).$$

But again by $\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v; \tilde{q}^w_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$, by substitutability we obtain $\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v; \tilde{q}^{w''}_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$, thus $\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v''; \tilde{q}^{w''}_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$. This is a contradiction because $(w'_{r''})_{r'' \in SC(h,\mu_d)} \leq (w''_{r''})_{r'' \in SC(h,\mu_d)} = v''$ and $(w'_{r''})_{r'' \in SC(h,\mu_d)} \succ_{SC(h,\mu_d)} (w_{r''})_{r'' \in SC(h,\mu_d)}$ while $\tilde{\mathrm{Ch}}_{SC(h,\mu_d)}(v''; \tilde{q}^{w''}_{SC(h,\mu_d)}) \in \arg\max_{\succeq_{SC(h,\mu_d)}} \{(w'''_{r''})_{r'' \in SC(h,\mu_d)} | (w'''_{r''})_{r'' \in SC(h,\mu_d)} \leq v'', \sum_{r'' \in S(SC(h,\mu_d))} w'''_{r''} \leq \tilde{q}^{w''}_{SC(h,\mu_d)}\}$.

A **doctor-optimal stable allocation** in the matching model with contracts is a stable allocation that every doctor weakly prefers to every other stable allocation (Hatfield and Milgrom, 2005). We will show that the flexible deferred acceptance mechanism is "isomorphic" to the doctor-optimal stable mechanism in the associated matching model with contracts.

Proposition 12. *Suppose that $\succeq_r$ is substitutable for every $r \in R$. Then the doctor-optimal stable allocation in the associated matching model with contracts, X', exists. In the original model, the flexible deferred acceptance mechanism produces matching $\mu(X')$ in a finite number of steps.*

Proof. First observe that the doctor-optimal stable allocation in matching with contracts can be found by the cumulative offer process in a finite number of steps (Hatfield and Milgrom, 2005; Hatfield and Kojima, 2010). Then, we observe that each step of the flexible deferred acceptance algorithm corresponds to a step of the cumulative offer process, that is, at each step, if d proposes to h in the flexible deferred acceptance algorithm, then at the same step of the cumulative offer process, contract (d, h) is proposed. Moreover, the set of doctors accepted for hospitals at a step of the flexible deferred acceptance algorithm corresponds to the set of contracts held at the corresponding step of the cumulative offer process. □

Theorem 6. *Suppose that $\succeq_r$ is substitutable for every $r \in R$. Then the flexible deferred acceptance algorithm stops in a finite number of steps. The mechanism produces a stable matching for any input and is group strategy-proof for doctors.*

Proof. Propositions 11 and 12 imply that the flexible deferred acceptance algorithm finds a stable matching in a finite number of steps. Also, Propositions 10 and 12 imply that the flexible deferred acceptance mechanism is (group) strategy-proof for doctors, as the substitutes condition and the law of aggregate demand imply that any mechanism that selects the doctor-optimal stable allocation is (group) strategy-proof (Hatfield and Milgrom, 2005; Hatfield and Kojima, 2009; Hatfield and Kominers, 2010). □

Appendix G. Discussion on Substitutability

In this section we aim to deepen our understanding of substitutability conditions. First we study the relationship between substitutability and consistency, and then we show that conditions (F.1) and (F.2) are independent.

Claim 5. *Condition (F.1) implies consistency.*[103]

Proof. Fix $\succeq$ and its associated quasi choice rule $\tilde{\mathrm{Ch}}_r$, and suppose that for some t, $\tilde{\mathrm{Ch}}_r(w';t) \leq w \leq w'$. Suppose also that condition (F.1) holds. We will prove $\tilde{\mathrm{Ch}}_r(w;t) = \tilde{\mathrm{Ch}}_r(w';t)$. Condition (F.1) implies $w \leq w' \Rightarrow \tilde{\mathrm{Ch}}_r(w;t) \geq \tilde{\mathrm{Ch}}_r(w';t) \wedge w$. Since $\tilde{\mathrm{Ch}}_r(w';t) \leq w$ implies $\tilde{\mathrm{Ch}}_r(w';t) \wedge w = \tilde{\mathrm{Ch}}_r(w';t)$, this means that $\tilde{\mathrm{Ch}}_r(w';t) \leq \tilde{\mathrm{Ch}}_r(w;t) \leq w'$. If $\tilde{\mathrm{Ch}}_r(w;t) \neq \tilde{\mathrm{Ch}}_r(w';t)$ then by the assumption that $\tilde{\mathrm{Ch}}_r$ is acceptant, we must have $\tilde{\mathrm{Ch}}_r(w;t) \succ_r \tilde{\mathrm{Ch}}_r(w';t)$. But then $\tilde{\mathrm{Ch}}_r(w';t)$ cannot be an element of $\arg\max_{\succeq_r}\{w''|w'' \leq w', \sum_{r'\in S(r)} w''_{r'} \leq t\}$ because $\tilde{\mathrm{Ch}}_r(w;t) \in \{w''|w'' \leq w', \sum_{r'\in S(r)} w''_{r'} \leq t\}$. Hence we have $\tilde{\mathrm{Ch}}_r(w';t) = \tilde{\mathrm{Ch}}_r(w;t)$. □

Example 17 (Regional preferences that violate (F.1) while satisfying (F.2))**.** There is a grand region r in which two hospitals reside. The capacity of each hospital is 2. Region r's preferences are as follows.

$$\succ_r\colon (2,2),(2,1),(1,2),(2,0),(0,2),(1,1),(1,0),(0,1),(0,0).$$

One can check by inspection that condition (F.2) and consistency are satisfied. To show that (F.1) is not satisfied, observe first that there is a unique associated choice rule (since preferences are strict), and denote it by $\tilde{\mathrm{Ch}}_r$. The above preferences imply that $\tilde{\mathrm{Ch}}_r((1,2);2) = (0,2)$ and $\tilde{\mathrm{Ch}}_r((2,2);2) = (2,0)$. But this is a contradiction to (F.1) because $(1,2) \leq (2,2)$ but $\tilde{\mathrm{Ch}}_r((1,2);2) \geq \tilde{\mathrm{Ch}}_r((2,2);2) \wedge (1,2)$ does not hold (the left hand side is $(0,2)$ while the right hand side is $(1,0)$). □

Example 18 (Regional preferences that violate (F.2) while satisfying (F.1))**.** There is a grand region r in which three hospitals reside. The capacity of each hospital is 1. Region r's preferences are as follows.

$$\succ_r\colon (1,1,1),(1,1,0),(1,0,1),(0,1,1),(0,0,1),(0,1,0),(1,0,0),(0,0,0).$$

One can check by inspection that condition (F.1) (and hence consistency by Claim 5) are satisfied. To show that (F.2) is not satisfied, observe first that there is a unique associated choice rule (since preferences are strict), and denote it by $\tilde{\mathrm{Ch}}_r$. The above preferences imply that $\tilde{\mathrm{Ch}}_r((1,1,1);1) = (0,0,1)$ and $\tilde{\mathrm{Ch}}_r((1,1,1);2) = (1,1,0)$. But this is a contradiction to (F.2) because $1 \leq 2$ but $\tilde{\mathrm{Ch}}_r((1,1,1);1) \leq \tilde{\mathrm{Ch}}_r((1,1,1);2)$ does not hold (the left hand side is $(0,0,1)$ while the right hand side is $(1,1,0)$). □

---

[103]Aygün and Sönmez (2012) independently prove analogous results although they do not work on substitutability defined over the space of integer vectors.

Appendix H. Additional Examples

In this section we present five additional examples. The first three examples present various comparative statics, and the last two examples find the limits to which our theory can be extended.

The first two examples strengthen the examples on comparative statics regarding regional preferences in the main text by showing that they hold under stronger assumptions on hospital preferences.

Example 19 (Ordering a hospital earlier may make it worse off even under homogenous hospital preferences). Let there be hospitals h_1 and h_2 in region r_1, and h_3 and h_4 in region r_2. Suppose that $(q_{h_1}, q_{h_2}, q_{h_3}, q_{h_4}) = (2, 2, 2, 2)$ and $(\bar{q}_{h_1}, \bar{q}_{h_2}, \bar{q}_{h_3}, \bar{q}_{h_4}) = (1, 0, 0, 0)$. The regional cap of r_1 is 2 and that for r_2 is 1. Preferences are

$$\succ_{h_i}: d_1, d_2, d_3, d_4 \quad \text{for all } i = 1, \ldots, 4,$$

$$\succ_{d_1}: h_4, h_1, \quad \succ_{d_2}: h_1, \quad \succ_{d_3}: h_2, \quad \succ_{d_4}: h_1, h_3.$$

We assume that h_3 is ordered earlier than h_4.

(1) Assume that h_1 is ordered earlier than h_2. In that case, in the flexible deferred acceptance mechanism, d_1 applies to h_4, d_2 and d_4 apply to h_1, and d_3 applies to h_2. d_1, d_2, and d_4 are accepted while d_3 is rejected. The matching finalizes with:

$$\mu = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 & \emptyset \\ d_2, d_4 & \emptyset & \emptyset & d_1 & d_3 \end{pmatrix}.$$

(2) Assume that h_1 is ordered after h_2. In that case, in the flexible deferred acceptance mechanism, d_1 applies to h_4, d_2 and d_4 apply to h_1, and d_3 applies to h_2. d_1, d_2, and d_3 are accepted while d_4 is rejected. d_4 applies to h_3 next, and d_1 is rejected. d_1 then applies to h_1, which now rejects d_2. The matching finalizes with:

$$\mu' = \begin{pmatrix} h_1 & h_2 & h_3 & h_4 & \emptyset \\ d_1 & d_3 & d_4 & \emptyset & d_2 \end{pmatrix}.$$

First, notice that hospital h_2 is better off in case (2). Thus being ordered earlier helps h_2 in this example. However, if h_1 prefers $\{d_1\}$ to $\{d_2, d_4\}$ (which is consistent with the assumption that hospital preferences are responsive with capacities), then h_1 is also made better off in case (2). Therefore, the effect of a picking order on hospitals' welfare is not monotone. □

Example 20 (Target monotonicity may fail even under homogenous hospital preferences). Consider a market that is identical to the one in Example 10, except that the target of h_1

is now decreased to 0, with the order such that h_1 chooses before h_2. Then h_1 is matched to $\{d_1\}$ under the flexible deferred acceptance mechanism. Therefore, if h_1 prefers $\{d_1\}$ to $\{d_2, d_4\}$, then h_1 is made better off when its target capacity is smaller. □

In these examples, it is hospitals that have homogeneous preferences. However, these examples can be modified so that doctors have homogeneous preferences. To do so, modify preferences to

$$\succ_{h_1}: d_1, d_2, d_4, \quad \succ_{h_2}: d_3, \quad \succ_{h_3}: d_4, \quad \succ_{h_4}: d_1,$$
$$\succ_{d_i}: h_4, h_1, h_3, h_2 \quad \text{for all } i = 1, \ldots, 4.$$

That is, hospital h finds doctor d acceptable if and only if d finds h acceptable in the previous examples, while all doctors find all hospitals acceptable and the ranking between two hospitals are consistent with the rankings between two acceptable hospitals in the previous examples. By construction, the matchings produced by the flexible deferred acceptance algorithm in this market are identical to those in the previous examples.

The next example studies comparative statics. Consider splitting a region into a number of smaller regions that partition the original region, and dividing the original regional cap among the new smaller regions. One might suspect that doing so makes doctors weakly worse off because the new set of constraints based on smaller regions may appear more stringent. The following example shows that this conjecture is incorrect. In fact, splitting regions can make some doctors and hospitals strictly better off, while making other doctors and hospitals strictly worse off.

Example 21 (Splitting regions has ambiguous welfare effects)**.** Let there be three hospitals, h_i for $i = 1, 2, 3$ in the grand region r with regional cap of 1. The capacity of each hospital is 1. There are three doctors in the market, d_i for $i = 1, 2, 3$. Suppose that the regional preferences are such that $(1, 0, 0) \succ_r (0, 1, 0) \succ_r (0, 0, 1)$.

We examine the effect of splitting region r into two smaller regions, $r' = \{h_1, h_3\}$ and $r'' = \{h_2\}$. The splitting needs some rule of allocating the regional cap to the smaller regions, which in this example corresponds to allocating the cap 1 of r either to r' or to r'' (while allocating the regional cap of zero to the other region).[104] In what follows we show that in either case, there exists a preference profile such that the welfare effect of splitting is ambiguous (i.e., under such a preference profile it is not the case that every agent

---

[104]It is only for simplicity that we use an example in which the regional cap of the grand region is one, and thus one of the smaller regions has a regional cap of zero. Our conclusion does not depend on this (perhaps unrealistic) assumption: The same point can be made in examples in which a region with regional cap larger than one is split.

of one side of the market becomes weakly better/worse off) under the flexible deferred acceptance mechanism.

Suppose first that the cap 1 of r is allocated to r'. Then, suppose

$$\succ_{d_i}: h_i, \qquad \succ_{h_i}: d_i$$

for $i = 2, 3$, and d_1 and h_1 regard no one as acceptable. The flexible deferred acceptance mechanism produces a matching μ such that $\mu_{d_2} = h_2$ before splitting, while it produces a matching μ' such that $\mu'_{d_3} = h_3$ after splitting (no other doctors are matched in either matching). Thus, splitting the region r makes d_2 and h_2 strictly worse off, while making d_3 and h_3 strictly better off.

Suppose second that the cap 1 of r is allocated to r''. Then, suppose

$$\succ_{d_i}: h_i, \qquad \succ_{h_i}: d_i$$

for $i = 1, 2$, and d_3 and h_3 regard no one as acceptable. The flexible deferred acceptance mechanism produces a matching μ such that $\mu_{d_1} = h_1$ before splitting, while it produces a matching μ' such that $\mu'_{d_2} = h_2$ after splitting (no other doctors are matched in either matching). Thus, splitting the region r makes d_1 and h_1 strictly worse off, while making d_2 and h_2 strictly better off. □

Note that an analogous example can be easily constructed to show that the effect of splitting on the welfare of the hospitals outside the split region is also ambiguous. Finally, also note that the conclusion holds regardless of how we define regional preferences after splitting the grand region r.

The next example shows that there exists no matching that satisfies a certain strengthening of the stability concept (see footnote 97).

Example 22 (Stable matchings do not necessarily exist under a stronger definition). Suppose that in the definition of stability (Definition 10), we further require that $r \subseteq SC(h, \mu_d)$. We demonstrate that there does not necessarily exist a stable matching under this notion.

There is a grand region r in which two subregions r' and r'' exist. Two hospitals h_1 and h_2 reside in r', and one hospital h_3 resides in r''. The capacity of each hospital is 1. The regional caps are 1 for r, 2 for r', and 1 for r''. Regional preferences are as follows.

$$\succ_r : (1,0), (0,1), (0,0),$$
$$\succ_{r'} : (0,2), (1,1), (2,0), (0,1), (1,0), (0,0).$$

There are two doctors d_1 and d_2. Preferences are as follows:

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

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

and preferences of h_3 are arbitrary.

To show that there exists no stable matching under the stronger definition, first note that the matching in which all doctors are unmatched is clearly unstable because, for example, pair (d_1, h_1) is a valid blocking pair. Also note that no matching under which both of the two doctors are matched is stable because the regional cap for the grand region r is one. Thus we are left with the cases in which only one doctor is matched to a hospital.

(1) Consider a matching μ such that $\mu_{d_1} = h_1$. Pair (d_2, h_1) is a blocking pair and, because $d_2 \succ_{h_1} d_1$, this is a legitimate blocking pair, showing that μ is unstable.

(2) Consider a matching μ such that $\mu_{d_1} = h_2$. First, note that pair (d_1, h_1) is a blocking pair. Moreover, since $SC(h_1, \mu_d) = r'$, we only need to check whether the cap $q_{r'}$ of region r' and the cap $q_{\{h_1\}}$ of the region $\{h_1\}$ are binding. Because $|\mu_{r'}| = 1 < 2 = q_{r'}$ and $|\mu_{h_1}| = 0 < 1 = q_{h_1}$, the regional caps are not binding. Hence the conditions in the stability concept are not satisfied, showing that showing that μ is unstable.

(3) Consider a matching μ such that $\mu_{d_1} = h_3$. Since h_3 is unacceptable to d_1, μ is unstable.

Any matching in which d_2 is matched to a hospital can be shown to be unstable in a symmetric manner. Hence, there does not exist any stable matching under the stronger definition. □

The final example of this section shows that a stable matching does not necessarily exist if the set of regions violates the assumption of a hierarchical structure.

Example 23 (Non-Hierarchical Regions)**.** Suppose that there are three hospitals, h_1, h_2, and h_3. Suppose that regions are not hierarchical, and

$$R = \{\{h_1\}, \{h_2\}, \{h_3\}, \{h_1, h_2\}, \{h_2, h_3\}, \{h_3, h_1\}, \{h_1, h_2, h_3\}\}.$$

Each region's regional cap is 1. There are two doctors, d_1 and d_2, and preferences are as follows:

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

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

Regional preferences for binary regions are that $\{h_1, h_2\}$ prefers a doctor to be in h_1 rather than h_2, $\{h_2, h_3\}$ prefers a doctor to be in h_2 rather than h_3, and $\{h_3, h_1\}$ prefers a doctor to be in h_3 rather than h_1.

Given the above specification, we show that there is no stable matching. First it is straightforward to see that there is no stable matching in which zero or two doctors are matched. So consider the case in which one doctor is matched. By the definition of stability, no hospital is matched to its second-choice doctor in any stable matching. This leaves us with only three possibilities: $\mu_{d_2} = h_1$, $\mu_{d_2} = h_2$, and $\mu_{d_1} = h_3$.

In the first case, (d_2, h_3) is a blocking pair, and from regional preferences of $\{h_3, h_1\}$, the existence of such a blocking pair violates stability. In the second case, (d_2, h_1) is a blocking pair, and from regional preferences of $\{h_1, h_2\}$, the existence of such a blocking pair violates stability. Finally, in the third case, (d_1, h_2) is a blocking pair, and from regional preferences of $\{h_2, h_3\}$, the existence of such a blocking pair violates stability. Hence there is no stable matching. $\square$

Appendix I. Semi-Strong Stability

In the main text, we pointed out that a strongly stable matching may not exist. Then we weakened the requirement and introduced the stability concept. A natural question is whether a concept stronger than stability can be imposed. To investigate this issue, we define the following notion.

Definition 14. A matching μ is **semi-strongly stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then (i) $|\mu_{r(h)}| = q_{r(h)}$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and (iii") either $\mu_d \notin H_{r(h)}$ or $|\mu_h| - \bar{q}_h \geq 0 \geq |\mu_{\mu_d}| - \bar{q}_{\mu_d}$.

The second part of condition (iii") says that a blocking pair (d, h) is not deemed as a legitimate deviation if doctor d is currently assigned in the region $r(h)$, the number of doctors matched with hospital μ_d is no more than its target, and that of hospital h is no less than its target. That is, a blocking pair that moves the distribution of doctors unambiguously away from the target capacity is not deemed to be a legitimate deviation. Note that some blocking pairs that are regarded as illegitimate deviations under stability are considered legitimate under this concept. For example, if hospital h_1 has the target capacity of 1 and $|\mu_{h_1}| = 10$, hospital h_2 has the target capacity of 5 and $|\mu_{h_2}| = 7$, and these two hospitals are in the same region, then a movement of a doctor from h_2 to a vacant position of h_1 is considered a legitimate deviation in semi-strong stability but not in stability.

Although semi-strong stability may seem to be an appropriate weakening of strong stability, unfortunately it has the same deficiency as strong stability: a semi-strongly stable matching does not necessary exist, and there exists no mechanism that is strategy-proof for doctors and selects a semi-strongly stable matching whenever there exists one.

The following example shows that a semi-strongly stable matching may not exist.

Example 24 (Semi-strongly stable matching may not exist)**.** There is one region r with regional cap $q_r = 1$, in which three hospitals, h_1, h_2 and h_3, reside. Each hospital h has a capacity of $q_h = 1$. Suppose that there are two doctors, d_1 and d_2. The target capacities of hospitals are $(\bar{q}_{h_1}, \bar{q}_{h_2}, \bar{q}_{h_3}) = (0, 0, 1)$. We assume the following preference:

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

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

Matching μ such that $\mu_{h_1} = \{d_1\}$ and $\mu_{h_2} = \mu_{h_3} = \emptyset$ is stable. Similarly μ' such that $\mu'_{h_1} = \mu'_{h_3} = \emptyset$ and $\mu'_{h_2} = \{d_2\}$ is also stable. It is easy to see that these are the only stable matchings. However, neither μ nor μ' is semi-strongly stable. To see that μ is not semi-strongly stable, note that a pair (d_1, h_2) constitutes a blocking pair and $\mu_{d_1} = h_1 \in H_{r(h_2)}$, and $|\mu_{h_1}| > \bar{q}_{h_1}$. Similarly μ' is not semi-strongly stable. Therefore, a semi-strongly stable matching does not exist in this market. $\square$

Note that Example 24 is similar to Example 2. In an analogous manner, we can easily modify Example 3 to construct an example in which there is no mechanism that is strategy-proof for doctors and finds a semi-strongly stable matching whenever there exists one.