STABILITY AND STRATEGY-PROOFNESS FOR MATCHING WITH CONSTRAINTS: A NECESSARY AND SUFFICIENT CONDITION

YUICHIRO KAMADA AND FUHITO KOJIMA

Abstract. Distributional constraints are common features in many real matching markets, such as medical residency matching, school admissions, and teacher assignment. We develop a general theory of matching mechanisms under distributional constraints. We identify the necessary and sufficient condition on the constraint structure for the existence of a mechanism that is stable and strategy-proof for the individuals. Our proof exploits a connection between a matching problem under distributional constraints and a matching problem with contracts.
JEL Classification Numbers: C70, D47, D61, D63.
Keywords: matching with constraints, medical residency matching, school choice, stability, strategy-proofness, matching with contracts, hierarchy

Note: This paper subsumes and supersedes part of our unpublished paper "General Theory of Matching under Distributional Constraints," cited as Kamada and Kojima (2015b) in this paper.

Date: December 2, 2016.
Kamada: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, y.cam.24@gmail.com. Kojima: Department of Economics, Stanford University, Stanford, CA 94305, fkojima@stanford.edu. We are grateful to Mustafa Oguz Afacan, Péter Biró, Eric Budish, Sylvain Chassang, Vincent Crawford, Hisao Endo, Clayton Featherstone, Tamás Fleiner, Drew Fudenberg, Tadashi Hashimoto, John William Hatfield, Christian Hellwig, Toshiaki Iizuka, Rob Irving, Ryo Jinnai, Onur Kesten, Scott Duke Kominers, Hideo Konishi, Mihai Manea, David Manlove, Taisuke Matsubae, Aki Matsui, Yusuke Narita, Muriel Niederle, Parag Pathak, Al Roth, Dan Sasaki, Tayfun Sönmez, Satoru Takahashi, William Thomson, Alexis Akira Toda, Kentaro Tomoeda, Utku Ünver, Jun Wako, Alex Westkamp, Yosuke Yasuda, Bumin Yenmez, and participants at numerous seminars and conferences for helpful comments. Doctor Masaaki Nagano answered our questions about medical residency in Japan. Jin Chen, Irene Hsu, Seung Hoon Lee, Stephen Nei, Bobak Pakzad-Hurson, Neil Prasad, Fanqi Shi, Pete Troyan, Wei Wen, and Rui Yu provided excellent research assistance. Kojima gratefully acknowledges financial support from the Sloan Foundation, as well as financial support from the National Research Foundation through its Global Research Network Grant (NRF- 2013S1A2A2035408).

1. Introduction

The theory of two-sided matching has been extensively studied ever since the seminal contribution by Gale and Shapley (1962), and it has been applied to match individuals and institutions in various markets in practice (e.g., students and schools, doctors and hospitals, and workers and firms). However, many real matching markets are subject to distributional constraints, i.e., caps are imposed on the numbers of individuals who can be matched to some subsets of institutions. Traditional theory cannot be applied to such settings because it has assumed away those constraints.

The objective of this paper is to understand the implication of the structure of distributional constraints, by investigating the extent to which a desirable mechanism can be designed under constraints. More specifically, we identify the necessary and sufficient condition on the constraint structure for the existence of a mechanism that is stable and strategy-proof for the individuals.[1] The necessary and sufficient condition is that constraints form a "hierarchy," that is, for any pair of subsets of institutions that are subject to constraints, the two are disjoint or one is a subset of the other.[2,3]

To understand implications of our result, consider matching doctors with hospitals. Suppose –as is the case in many countries– the government desires to keep geographical imbalance of doctors in check, leading to constraints imposed on different regions of the country (Kamada and Kojima, 2015a). Since the family of regions form a partition of the country, a special case of a hierarchy, our result implies that there exists a mechanism which is stable and strategy-proof for doctors. The same positive conclusion holds even if the geographical constraints are imposed on a hierarchy of regions, for instance on states and counties (here, the constraints form a hierarchy because counties are mutually disjoint and subsets of states).

Next, suppose that the government desires to impose caps not only on the number of doctors in different regions, but also on the number of doctors working in different medical specialties (this is also a common public health policy issue; see Section 4 for a concrete policy debate). The constraints do not form a hierarchy in this case because a region and

[1]We formally define a stability concept under constraints in Section 2. Our stability concept reduces to the standard stability concept of Gale and Shapley (1962) if there are no binding constraints.

[2]Such a hierarchical structure is called a "laminar family" in the mathematics literature.

[3]Budish, Che, Kojima, and Milgrom (2013) consider a random object allocation problem with floor and ceiling constraints and show that "bi-hierarchical" constraints are necessary and sufficient for implementability of random assignments. Although the condition they reach is similar to ours, their implementability is unrelated to stability and strategy-proofness that we study here. See also Milgrom (2009).

a specialty partially overlap with each other. Therefore, the necessity part of our result implies that it is impossible to design a mechanism that is stable and strategy-proof for doctors. By contrast, if the constraints are imposed on doctors exercising different medical specialties *in each region*, then the constraints regain a hierarchical structure, so our result implies that a desirable mechanism exists.

Our result also has implications for the design of school choice mechanisms (Abdulkadiroğlu and Sönmez, 2003). Suppose that the school district wants to maintain certain balance of student body at each school in terms of socio-economic class. Different socio-economic classes form a partition (and thus a hierarchy), so a stable and strategy-proof mechanism exists. By contrast if, for instance, the school district desires to maintain balance in both socio-economic class and gender, then the constraints do not form a hierarchy, so a desirable mechanism does not exist.[4]

In addition to its applied value, we believe that our analytical approach is of independent interest. The approach for showing the sufficiency of a hierarchy is to find a connection between our model and the "matching with contracts" model (Hatfield and Milgrom, 2005).[5] More specifically, we define a hypothetical matching problem between doctors and *the "hospital side"* instead of doctors and hospitals; We regard the hospital side as a hypothetical consortium of hospitals that acts as one agent. By imagining that the hospital side (hospital consortium) makes a common employment decision, we can account for interrelated doctor assignments across hospitals, an inevitable feature in markets under distributional constraints. This association necessitates, however, that we distinguish a doctor's matching in different hospitals. We account for this complication by constructing

[4]Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu (2005) show that a stable and strategy-proof mechanism exists under constraints on socio-economic class and give an example that shows nonexistence when constraints are imposed also on gender. There are three main differences between their work and ours. First, we consider a different formulation from theirs. Second, our characterization of the necessary condition for existence is new. Third, we show existence under hierarchical constraints. See also Roth (1991) on gender balance in labor markets, Ergin and Sönmez (2006), Hafalir, Yenmez, and Yildirim (2013), Ehlers, Hafalir, Yenmez, and Yildirim (2014), and Echenique and Yenmez (2015) on diversity in schools, Westkamp (2013) on trait-specific college admission, Abraham, Irving, and Manlove (2007) on project-specific quotas in projects-students matching, and Biró, Fleiner, Irving, and Manlove (2010) on college admission with multiple types of tuitions.

[5]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.

the hospital side's choice defined over *contracts* rather than doctors, where a contract specifies a doctor-hospital pair to be matched. Once this connection is established, with some work we show that results in the matching-with-contract model can be applied to our matching model under distributional constraints.[6] This method shows that there exists a mechanism that is stable and strategy-proofness for doctors. Note that our technique is different from those in school choice with diversity constraints which have schools' aggregated preferences over different types of students *as primitives*. By contrast, aggregated preferences of the hospital side are *not* primitives of our model, and our contribution is to construct an appropriate hypothetical model with an aggregated choice function. We envision that analyzing a hypothetical model of matching with contracts may prove to be a useful approach for tackling complex matching problems one may encounter in the future.[7,8]

A recent paper by Hatfield, Kominers, and Westkamp (2015) characterizes the class of choice functions over abstract contracts such that there exists a mechanism that is stable and strategy-proof for doctors (see a related contribution by Hirata and Kasuya (2015) as well). Our result and theirs are independent of each other because our model and stability concept are different from theirs: Our model is that of matching with constraints and, even though we can associate our stability concept to theirs through our technique, these concepts are still not equivalent (Kamada and Kojima, 2015a).

This paper is part of our research agenda to study matching with constraints. Let us explain the contribution of the present paper relative to our other papers on the agenda. Kamada and Kojima (2015a) consider the case in which constraints form a partition and show that a desirable mechanism exists. In practice, however, non-partitional constraints are prevalent. In order to understand what kinds of applications can be accommodated, the present paper does not presume a partition structure, and instead investigates a conceptual question. More specifically, we find a necessary and sufficient condition for the

[6] Specifically, we invoke results by Hatfield and Milgrom (2005), Hatfield and Kojima (2009, 2010), and Hatfield and Kominers (2009, 2012).

[7] Indeed, after we circulated the first draft of the present paper, this technique was adopted by other studies such as Goto, Iwasaki, Kawasaki, Yasuda, and Yokoo (2014), Goto, Hashimoto, Iwasaki, Kawasaki, Ueda, Yasuda, and Yokoo (2014), and Kojima, Tamura, and Yokoo (2015) in the context of matching with distributional constraints. See Sönmez and Switzer (2013) for a more direct application of matching with contracts model, where a cadet can be matched with a branch under one of two possible contracts. See also Sönmez (2013) and Kominers and Sönmez (2012).

[8] See also Hatfield, Kominers, Nichifor, Ostrovsky, and Westkamp (2013) who connect stability in a trading network with stability in an associated model of many-to-one two-sided matching with transfer due to Kelso and Crawford (1982).

existence of a desirable mechanism in terms of the constraint structure. Kamada and Kojima (2016) investigate how to define the "right" stability concept for matching with constraints, defining what they call strong stability and weak stability. They show strong stability suffers from a non-existence problem whenever constraints are "non-trivial" while weakly stable matchings exist for a wide range of constraints and are efficient. Stability defined in the present paper is weaker than strong stability but stronger than weak stability. In contrast to strong and weak stability, the definition of stability in this paper uses information about policy goals of regions in terms of how to allocate their limited seats.[9] Thus our stability is particularly appropriate when such information is available. Instead of exploring many kinds of stability concepts, the present paper fixes one stability concept (the stability defined in the present paper) and asks the implication of constraint structures for the existence of desirable mechanisms.

The rest of this paper proceeds as follows. In Section 2, we present the model. Section 3 states the main result. Section 4 presents applications of the main result, and Section 5 provides a proof sketch. Section 6 concludes. Appendix provides the formal proof of the result as well as a number of discussions.

2. Model

This section introduces a model of matching under distributional constraints. We describe the model in terms of matching between doctors and hospitals with "regional caps," that is, upper bounds on the number of doctors that can be matched to hospitals in each region. However, the model is applicable to various other situations in and out of the residency matching context. For example, in medical residency applications, a region can represent a geographical region, medical specialty, or a combination of them.[10] Another example is school choice, where a region can represent a socio-economic class of students.

We begin with preliminary definitions for two-sided matching in Section 2.1. Then Section 2.2 introduces our model of matching with constraints.

2.1. **Preliminary Definitions.**

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

[9] This usage of such information will be expressed in the definition of "illegitimate" doctor-hospital pairs.

[10] In real medical matching, a hospital may have multiple residency programs in it. These programs may differ from one another in terms of emphasis on specialties, for example. In such a case, the term "a hospital" should be understood to mean a residency program.

$h \succ_d h'$ 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 hospital h if $d \succ_h \emptyset$.[11] 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 $|D'| > q_h$.

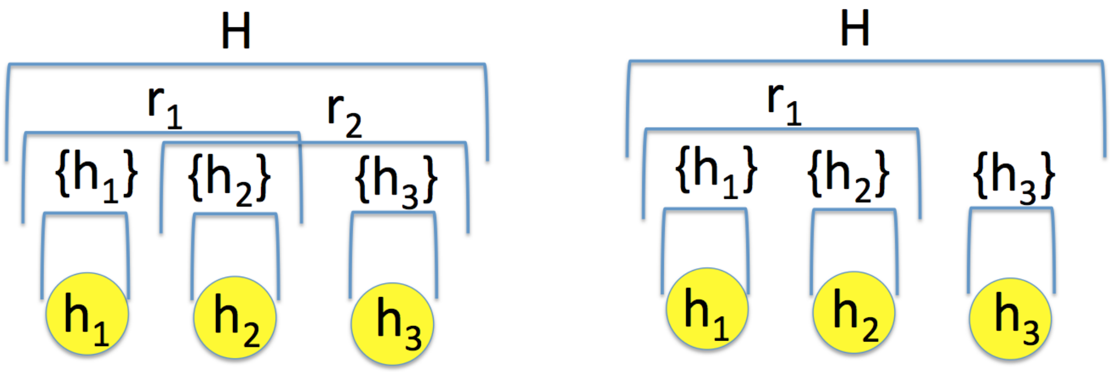
In words, preference relation $\succ_h$ is responsive with a capacity if the ranking of a doctor (or keeping a position vacant) is independent of her colleagues, and any set of doctors exceeding its capacity is unacceptable. We assume that preferences of each hospital h are responsive with capacity q_h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

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 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 $|\mu_h| \leq q_h$. 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.

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



(A) A non-hierarchical set of regions R.

(B) A hierarchical set of regions R'.

FIGURE 1. An example of sets of regions in Example 1.

2.2. **Model with Constraints.**

Region Structure. A collection $R \subseteq 2^H \backslash \{\emptyset\}$ is called a set of **regions**. Assume $\{h\} \in R$ for all $h \in H$ and $H \in R$.

A collection of regions $S \subset R$ is called a **partition of** $r \in R$ if $S \neq \{r\}$, $\bigcup_{r' \in S} r' = r$, and $r_1 \cap r_2 = \emptyset$ for all $r_1, r_2 \in S$ with $r_1 \neq r_2$. A partition S of r is called a **largest partition of** r if there exists no partition $S' \neq S$ of r such that $r' \in S$ implies $r' \subseteq r''$ for some $r'' \in S'$. Note that, for a given region r, there can be more than one largest partition of r. We denote by $\mathcal{LP}(r)$ the collection of largest partitions of r. For $r \in R$ and $S \in \mathcal{LP}(r)$, we refer to each element of S as a **subregion** of r with respect to S.

A set of regions R is a **hierarchy** if $r, r' \in R$ implies $r \subseteq r'$ or $r' \subseteq r$ or $r \cap r' = \emptyset$.

Below is an example of a set of regions.

Example 1. There are hospitals h_1, h_2, and h_3. The regions are

$$R = \{H, r_1, r_2, \{h_1\}, \{h_2\}, \{h_3\}\},$$

where $r_1 = \{h_1, h_2\}$ and $r_2 = \{h_2, h_3\}$. See Figure 1a for a graphical representation. In this example, the largest partitions of H are $S = \{r_1, \{h_3\}\}$ and $S' = \{\{h_1\}, r_2\}$. Regions r_1 and $\{h_3\}$ are subregions of H with respect to S, and $\{h_1\}$ and r_2 are subregions of H with respect to S'. R is not a hierarchy because $r_1 \not\subseteq r_2$, $r_2 \not\subseteq r_1$, and $r_1 \cap r_2 = \{h_2\} \neq \emptyset$. By contrast, $R' := R \backslash \{r_2\}$ is a hierarchy (see Figure 1b for a graphical representation). □

Regional Preferences. When a given region is faced with applications by more doctors than the regional cap, the region has to allocate limited seats among its subregions. We consider the situation in which regions have policy goals in terms of doctor allocations, and formalize such policy goals using the concept of "regional preferences." For example, a situation in which the government has a policy goal regarding doctor allocations across different geographic regions can be captured by the regional preferences of the grand region H. The regional preferences of a geographical region r whose subregions are all singleton sets captures r's policy goal in terms of doctor allocations across different hospitals in r.

For each $r \in R$ that is not a singleton set and $S \in \mathcal{LP}(r)$, a regional preference for r, denoted $\trianglerighteq_{r,S}$, is a weak ordering over $W_{r,S} := \{w = (w_{r'})_{r' \in S} | w_{r'} \in \mathbb{Z}_+ \text{ for every } r' \in S\}$. That is, $\trianglerighteq_{r,S}$ is a binary relation that is complete and transitive (but not necessarily antisymmetric). We write $w \triangleright_{r,S} w'$ if and only if $w \trianglerighteq_{r,S} w'$ holds but $w' \trianglerighteq_{r,S} w$ does not. Vectors such as w and w' are interpreted to be supplies of acceptable doctors to the subregions of region r, but they only specify how many acceptable doctors apply to hospitals in each subregion and no information is given as to who these doctors are. We denote by $\trianglerighteq$ the profile $(\trianglerighteq_{r,S})_{r \in R, S \in \mathcal{LP}(r)}$.

Given $\trianglerighteq_{r,S}$, a function

$$\tilde{\text{Ch}}_{r,S} : W_{r,S} \times \mathbb{Z}_+ \to W_{r,S}$$

is an **associated quasi choice rule** if $\tilde{\text{Ch}}_{r,S}(w;t) \in \arg\max_{\trianglerighteq_{r,S}} \{w' | w' \leq w, \sum_{r' \in S} w'_{r'} \leq t\}$ for any non-negative integer vector $w = (w_{r'})_{r' \in S}$ and non-negative integer t.[12] Intuitively, $\tilde{\text{Ch}}_{r,S}(w,t)$ is a 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 numbers of doctors cannot exceed the quota t. Note that there may be more than one quasi choice rule associated with a given weak ordering $\trianglerighteq_{r,S}$ because the set $\arg\max_{\trianglerighteq_{r,S}} \{w' | w' \leq w, \sum_{r' \in S} w'_{r'} \leq t\}$ may not be a singleton for some w and t.

We assume that the regional preferences $\trianglerighteq_{r,S}$ satisfy $w \triangleright_{r,S} w'$ if $w' \lneq w$. This condition formalizes the idea that region r prefers to fill as many positions in its subregions as possible. This requirement implies that any associated quasi choice rule is **acceptant** in the sense that, for each w and t, if there exists $r' \in S$ such that $[\tilde{\text{Ch}}_{r,S}(w;t)]_{r'} < w_{r'}$, then $\sum_{r'' \in S} [\tilde{\text{Ch}}_{r,S}(w;t)]_{r''} = t$.[13] 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.

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

[13]This condition is a variant of properties used by Alkan (2001) and Kojima and Manea (2010) in the context of choice functions over matchings.

We say that $\trianglerighteq_{r,S}$ is **substitutable** if there exists an associated quasi choice rule $\tilde{\text{Ch}}_{r,S}$ that satisfies

$$w \leq w' \text{ and } t \geq t' \Rightarrow \tilde{\text{Ch}}_{r,S}(w;t) \geq \tilde{\text{Ch}}_{r,S}(w';t') \wedge w.$$

Just as in the standard definition (e.g., Hatfield and Milgrom (2005)), our concept of substitutability requires that weakly more doctors be rejected if more doctors apply to each hospital. In addition, our substitutability requires that weakly more doctors be accepted if the quota increases. We provide a detailed discussion of this concept in Appendix B. Throughout our analysis, we assume that $\trianglerighteq_{r,S}$ is substitutable for any $r \in R$ and $S \in \mathcal{LP}(r)$.

Stability. We assume that each region $r \in R$ is endowed with a nonnegative integer κ_r called a **regional cap**.[14] We denote by $\kappa = (\kappa_r)_{r \in R}$ the profile of regional caps across all regions in R. A matching is **feasible** if $|\mu_r| \leq \kappa_r$ for all $r \in R$, where $\mu_r = \cup_{h \in r} \mu_h$. In other words, feasibility requires that the regional cap for every region is satisfied. For $R' \subseteq R$, we say that μ is **Pareto superior** to μ' for R' if $(|\mu_{r'}|)_{r' \in S} \trianglerighteq_{r,S} (|\mu'_{r'}|)_{r' \in S}$ for all (r, S) where $r \in R'$ and $S \in \mathcal{LP}(r)$, with at least one of the relations holding strictly. Given a matching μ, denote by $\mu^{d \to h}$ the matching such that $\mu^{d \to h}_{d'} = \mu_{d'}$ for all $d' \in D \setminus \{d\}$ and $\mu^{d \to h}_d = h$.

Given these concepts, let us now introduce two key notions to define stability. First, we say that a pair (d, h) is **infeasible** at μ if $\mu^{d \to h}$ is not feasible. Second, we say that a pair (d, h) is **illegitimate** at μ if there exists $r \in R$ with $|\mu_r| = \kappa_r$ such that $\mu^{d \to h}$ is not Pareto superior to μ for $\{r' \in R \mid \mu_d, h \in r' \text{ and } r' \subseteq r\}$.

Definition 1. A matching μ is **stable** if it is feasible, individually rational, and if (d, h) is a blocking pair then $d' \succ_h d$ for all doctors $d' \in \mu_h$ and (d, h) is either infeasible or illegitimate at μ.

The standard definition of stability without regional caps requires individual rationality and the absence of blocking pairs. With regional caps, however, there are cases in which every feasible and individually rational matching admits a blocking pair. For this reason, we allow for the presence of some blocking pairs. To keep the spirit of stability, however, we require only certain kinds of blocking pairs remain. Specifically, we demand that all remaining blocking pairs be either infeasible or illegitimate. Below we provide justification for the choice of these restrictions.

[14] Kamada and Kojima (2015a) use q_r to denote a regional cap, but here we use the notation κ_r to reduce confusion with hospital capacities.

A pair of a doctor d and a hospital h is infeasible if moving d to h while keeping other parts of the matching unchanged leads to a violation of a regional cap. To the extent that regional caps encode what matchings are allowed in the given situation, a demand by a blocking pair that would cause a violation of a regional cap does not have the same normative support as in the case without regional caps. For this reason, our stability concept allows for infeasible blocking pairs to remain.

A doctor-hospital pair is illegitimate if the movement of doctor d to h does not lead to a Pareto superior distribution of doctors for a certain set of regions.We require any region r' in this set satisfy two conditions. First, we require that r' contains both hospitals μ_d (the original hospital for d) and h, as this corresponds to the case in which r' is in charge of controlling distributions of doctors involving these hospitals. Second, the region r' should be currently "constrained." That is, it is a subset of some region r whose regional cap is full in the present matching: In such a case, the region r should ration the distribution of doctors among its subregions, each of which needs to ration the distribution among its subregions, and so forth, which indirectly constrains the number of doctors that can be matched in r'. The requirement that r' is (indirectly) constrained limits the case in which a blocking pair is declared illegitimate. This is in line with our motivation to keep the spirit of stability, i.e., to allow for the presence of blocking pairs only in a conservative manner.[15]

The implicit idea behind the definition is that the government or some authority can interfere and prohibit a blocking pair from being executed if regional caps are an issue. Thus, our preferred interpretation is that stability captures a normative notion that it is desirable to implement a matching that respects participants' preferences to the extent possible. Justification of the normative appeal of stability has been established in the recent matching literature, and Kamada and Kojima (2015a) offer further discussion on this point, so we refer interested readers to that paper for details.

Remark 1. Kamada and Kojima (2016) define two other stability concepts, which they call strong stability and weak stability. Strong stability is more demanding than stability, requiring any blocking pairs lead to infeasibility. Weak stability is less demanding than stability, allowing for some blocking pairs that are neither infeasible nor illegitimate. Kamada and Kojima (2016) establish that a matching is strongly stable if and only if it is stable for all possible regional preference profiles, while a matching is weakly stable if and only if there exists a regional preference profile under which it is stable.

[15] For the case of hierarchies, we provide an alternative interpretation of stability in Remark 5 in the Appendix.

Remark 2. Kamada and Kojima (2016) show that any weakly stable matching is (constrained) efficient, i.e., 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 $i \in D \cup H$.[16] Because stability implies weak stability, a stable matching is efficient for any regional preferences, which provides one normative appeal of our stability concept.

Remark 3. Kamada and Kojima (2016) demonstrate that (i) there does not necessarily exist a strongly stable matching, and (ii) if a mechanism produces a strongly stable matching whenever there exists one, then it is not strategy-proof for doctors. Given these negative findings, the present paper focuses on stability as defined in Definition 1.

Remark 4. Practically, moving a doctor from one hospital to another involves administrative tasks on the part of relevant regions (we give examples of possible organizations dealing with such administration in Section 4), hence disallowing only those blocking pairs that Pareto-improve the relevant regions is, in our view, the most plausible notion in our environment. An alternative notion of illegitimacy may be that a doctor-hospital pair is said to be illegitimate if moving doctor d to h leads to a Pareto inferior distribution of doctors for the set of regions that we consider here. This notion not only is inconsistent with our view just described, but also leads to nonexistence.[17]

__Mechanism.__ Recall that κ denotes the profile of regional caps, and $\trianglerighteq$ denotes the profile of regional preferences. A **mechanism** φ is a function that maps preference profiles to matchings for a given profile $(\kappa, \trianglerighteq)$. The matching under φ at preference profile $\succ$ is denoted $\varphi^{\kappa,\trianglerighteq}(\succ)$, and agent i's match is denoted by $\varphi_i^{\kappa,\trianglerighteq}(\succ)$ for each $i \in D \cup H$.

A mechanism φ is said to be **stable** if, for each $(\kappa, \trianglerighteq)$ and a preference profile $\succ$, the matching $\varphi^{\kappa,\trianglerighteq}(\succ)$ is stable.

A mechanism φ is said to be **strategy-proof for doctors** if there do not exist $(\kappa, \trianglerighteq)$, a preference profile $\succ$, a doctor $d \in D$, and preferences $\succ'_d$ of doctor d such that

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

[16] Since regional caps are a primitive of the environment, we consider a *constrained* efficiency concept.

[17] To see this, consider a two-doctor two-hospital example with one region with cap 1, where doctor $i \in \{1, 2\}$ likes hospital i best and hospital $j \neq i$ second, and hospital i likes doctor $j \neq i$ best and doctor i second. If the region containing hospitals 1 and 2 are indifferent between allocations $(1, 0)$ and $(0, 1)$, then there exists no matching that satisfies the strengthened version of stability. A similar logic is used to show there exists no strongly stable matching in Kamada and Kojima (2016).

That is, no doctor has an incentive to misreport her preferences under the mechanism.[18]

3. Characterization Theorem

Theorem 1. *Fix D with $|D| \geq 2$, H, and a set of regions R. The following statements are equivalent.*

(1) *R is a hierarchy.*

(2) *There exists a mechanism that is stable and strategy-proof for doctors.*

The proof of Theorem 1 is involved, and the proof approach is of independent interest. Thus, we illustrate a proof sketch in the next section, while relegating the formal proof to Appendices C and D.

The theorem identifies the conditions on the markets for which we can find a mechanism that is stable and strategy-proof for doctors. Since our proof for the assertion that statement (1) implies statement (2) is constructive, for those markets in which constraints are a hierarchy, the theorem tells us that we can directly use the mechanism we construct. Also, for those markets in which constraints do not form a hierarchy, the theorem shows that there is no hope of adopting a mechanism that is stable and strategy-proof for doctors. To understand these points in concrete examples, the next section considers practical applications.

4. Applications

This section discusses various applications of Theorem 1. The main goal of this section is to illustrate the implications of Theorem 1 about what types of environments and constraints in practice allow for a desirable mechanism.

4.1. Geographical constraints in medical match.

Consider the medical matching market in Japan. In order to keep geographical imbalance of doctors in check, the Japanese government imposes caps on the numbers of doctors in different prefectures that partition the country. Kamada and Kojima (2015a) propose a mechanism that produces a stable matching and is strategy-proof for doctors. Since partitions are a special case of hierarchy, their result can be derived as a corollary of Theorem 1.[19]

In addition, it may be of interest to consider multiple layers of geographical regions. In our private communication with Japanese government officials, they expressed concerns

[18] Roth (1982) shows that there is no mechanism that produces a stable matching for all possible preference profiles and is strategy-proof for both doctors and hospitals even in a market without regional caps.

[19] See Kamada and Kojima (2015b) for details.

about imbalance of doctors not only across prefectures but also within each prefecture. Theorem 1 provides solutions to this concern by considering hierarchical constraints within each prefecture. For example, according to the Medical Care Act of Japan, each prefecture is divided into smaller geographic units called *Iryo-ken* (medical areas) on which various health care policies are implemented.[20] This is a case of a hierarchy, and hence Theorem 1 implies that a desirable mechanism exists.

4.2. **Geographical and speciality constraints in medical match.** Consider medical match with prefecture-level constraints as in the last subsection, but suppose that the government also desires to impose caps on medical specialities. In the Japanese context, this concern is clearly exemplified by a proposal made to the governmental committee meeting, which is titled "Measures to Address Regional Imbalance and Specialty Imbalance of Doctors" (Ministry of Health, Labour and Welfare, 2008).[21] Suppose for now that the government decides to impose nationwide caps on each specialty. In this situation, a hospital program may be in the "obstetrics" region at the same time as in "Tokyo" region, where neither of the two is a subset of the other. Theorem 1 implies that there exists no mechanism that is stable and strategy-proofness for doctors in this environment. This suggests that the policy maker should either give up one of these desiderata, or rethink which institution/committee to give authority to claim a certain number of doctors and declare their regional preferences. For instance, the policy maker may decide not to give authority to claim caps and preferences to nationwide organizations such as Japan Society of Obstetrics and Gynecology, but to give it to each of the regional ones such as Tokyo Association of Obstetricians & Gynecologists.[22] To the extent that doctors treat patients in person, imposing specialty constraints in each region rather than nationally may be reasonable. Constraints form a hierarchy in such a situation, so Theorem 1 implies that there exists a mechanism with our desired properties.

4.3. **Field and financial constraints in college admission.** In Hungarian college admission, there are upper-bound constraints on the number of students in terms of

[20]Article 30-4 of the Medical Care Act of Japan.

[21]In 2008, Japan took a measure that intended to address (only) regional imbalance, but even in 2007, the sense of crisis about specialty imbalance is shared by the private sector as well: In a sensationally entitled article, "Obstetricians Are in Short Supply! Footsteps of Obstetrics Breakdown," NTT Com Research (2007) reports that many obstetrics hospitals have been closed even in urban areas such as Tokyo.

[22]See http://www.jsog.or.jp for Japan Society of Obstetrics and Gynecology and http://www.taog.gr.jp for Tokyo Association of Obstetricians & Gynecologists.

different fields of study as well as whether the study is subsidized by the government. In this problem, the constraints formed a hierarchy until 2007, but then the constraints were modified such that they do not form a hierarchy any more (Biró, Fleiner, Irving, and Manlove, 2010). Thus, Theorem 1 implies that our mechanism achieves stability and strategy-proofness for students in the old environment, but no mechanism can achieve these properties after the change.

4.4. **Diversity constraints in school choice.** Consider the design of school choice mechanisms (Abdulkadiroğlu and Sönmez, 2003). Maintaining diversity is a major concern in school choice, and many school districts have used policies to achieve this goal, e.g. New York City (NY), Chicago (IL), Seattle (WA), Jefferson County (KY), Louisville (KY), Minneapolis (MN), and White Plains (NY) (Abdulkadiroğlu, Pathak, and Roth, 2005; Hafalir, Yenmez, and Yildirim, 2013; Dur, Pathak, and Sönmez, 2016).

Suppose first that the school district wants to maintain certain balance of student body at each school in terms of socio-economic class such as race.[23] Theorem 1 implies that a stable and strategy-proof mechanism exists in this case as different socio-economic classes form a partition, a special case of hierarchy. However, the constraints do not form a hierarchy if, for example, the policy of the school district is to maintain balance in terms of both socio-economic class and gender. Therefore, Theorem 1 implies that a desirable mechanism does not exist in this case.

5. Proof Sketch of Theorem 1

The proof of Theorem 1 is involved. Thus, we illustrate a sketch of the proof in this section, while presenting the formal proof in Appendix C. We illustrate the basic ideas for the case of hierarchy first and then explain the remaining case.

5.1. **Hierarchy.** For the case of hierarchies, our proof strategy is to connect our matching model with constraints to the "matching with contracts" model (Hatfield and Milgrom, 2005). More specifically, given the original matching model under constraints, we define an "associated model," a hypothetical matching model between doctors and *the "hospital side"* instead of doctors and hospitals; In the associated model, we regard the hospital side as a hypothetical consortium of all hospitals that acts as one agent. By imagining that the hospital side (hospital consortium) makes a coordinated employment decision, we can account for the fact that acceptance of a doctor by a hospital in one region may depend

[23]Affirmative action based on race has been declared illegal in, for instance, *Parents Involved in Community Schools v. Seattle School District No. 1.* Currently, many school districts use other criteria such as income or neighborhood of the family.

on doctor applications to other hospitals in the same region, an inevitable feature in markets under distributional constraints. This association necessitates, however, that we distinguish a doctor's placements in different hospitals. We account for this complication by defining the hospital side's choice function over *contracts* rather than doctors, where a contract specifies a doctor-hospital pair to be matched. We construct such a choice function by using two pieces of information: the preferences of all the hospitals and regional preferences. The idea is that each hospital's preferences are used for choosing doctors *given the number of allocated slots*, while regional preferences are used to regulate slots allocated to different hospitals. In other words, regional preferences trade off multiple hospitals' desires to accept more doctors, when accepting more is in conflict with the regional cap. With the help of this association, we demonstrate that any stable allocation in the associated model with contracts induces a stable matching in the original model with distributional constraints (Proposition 2).

In order to use this association, we show that the key conditions in the associated model—the substitutes condition and the law of aggregate demand—are satisfied (Proposition 1).[24] This enables us to invoke existing results for matching with contracts, namely that an existing algorithm called the "cumulative offer process" finds a stable allocation, and it is (group) strategy-proof for doctors in the associated model (Hatfield and Milgrom, 2005; Hatfield and Kojima, 2009; Hatfield and Kominers, 2012).

The full proof in the Appendix formalizes this idea. In particular, we introduce a new algorithm in the original matching model with constraints, called the flexible deferred acceptance algorithm. That algorithm is a generalization of the deferred acceptance algorithm, in which the acceptance by hospitals in each step is made in a coordinated way that is consistent with the regional constraints. We establish that, with the hierarchical region structure, the outcome of the cumulative offer process in the associated model corresponds to the matching produced by the flexible deferred acceptance algorithm in the original model (Remark 6). This correspondence establishes that the flexible deferred acceptance algorithm finds a stable matching in the original problem and this algorithm is (group) strategy-proof for doctors, proving the statement for the hierarchy case in Theorem 1. For illustration, our proof approach is represented as a chart in Figure 2.

5.2. **Non-Hierarchy.** To illustrate the proof for this part, we present an example of non-hierarchy and show that, in that example, there is no mechanism that is stable and

[24]Substitutability of regional preferences plays a crucial role in the proof of Proposition 1, and we regard it as reasonable in our applications. Substitutability is commonly assumed in various domains in the matching literature.

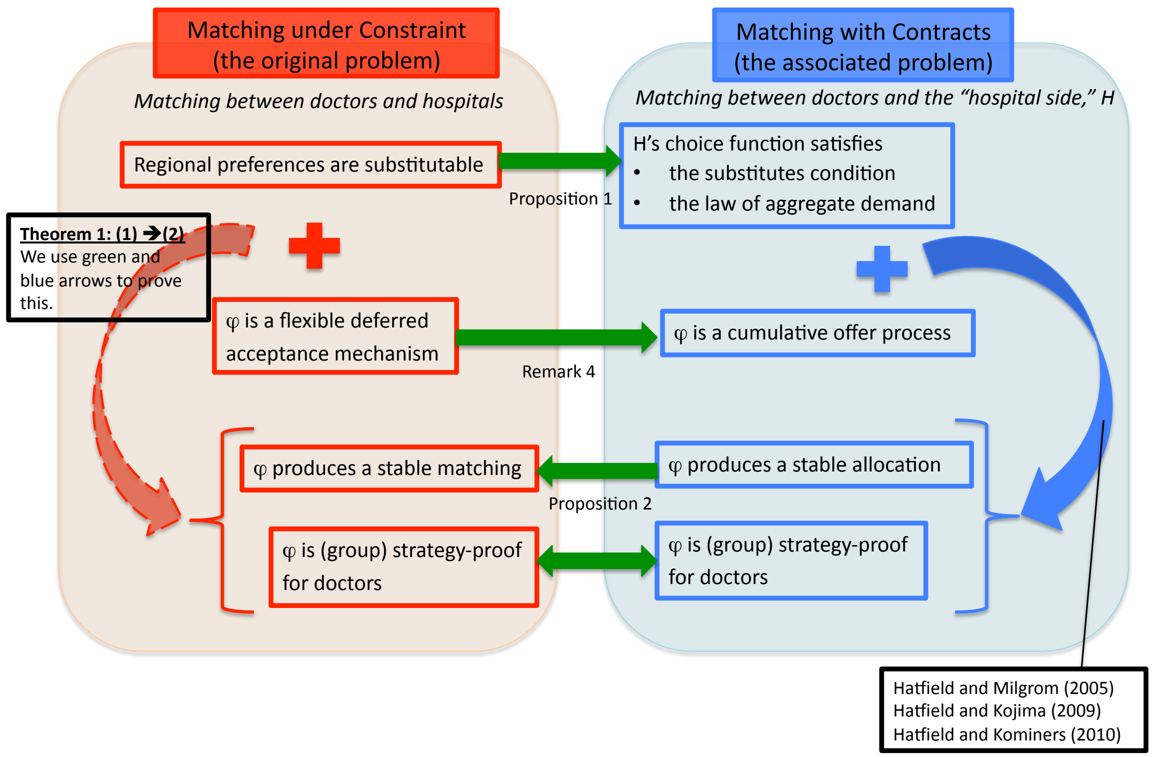


FIGURE 2. Proof sketch for (1) ⇒ (2) of Theorem 1.

strategy-proof for doctors. The general case is somewhat more involved, but the main idea can be seen in this example.

Consider the problem described in Example 1: In that example, there are 3 hospitals, h_1, h_2, and h_3. The regions are $R = \{H, r_1, r_2, \{h_1\}, \{h_2\}, \{h_3\}\}$, where $r_1 = \{h_1, h_2\}$ and $r_2 = \{h_2, h_3\}$. Recall Figure 1a for a graphical representation. Note that $h_1 \in r_1 \setminus r_2$, $h_2 \in r_1 \cap r_2$, and $h_3 \in r_2 \setminus r_1$, so R is not a hierarchy. We will show that there does not exist a mechanism that is stable and strategy-proof for doctors.

Clearly, $S_1 := \{\{h_1\}, \{h_2\}\}$ is the unique largest partition of r_1. Similarly, $S_2 := \{\{h_2\}, \{h_3\}\}$ is the unique largest partition of r_2. Suppose that, under $\rhd_{r_1,S_1}$, region r_1 prefers a vector such that the coordinate corresponding to $\{h_1\}$ is 1 and the other coordinate is 0 to a vector such that the coordinate corresponding to $\{h_2\}$ is 1 and the other coordinate is 0. Also suppose that, under $\rhd_{r_2,S_2}$, region r_2 prefers a vector such that the coordinate corresponding to $\{h_2\}$ is 1 and the other coordinate is 0 to a vector such that the coordinate corresponding to $\{h_3\}$ is 1 and the other coordinate is 0. Finally, let $\kappa_{r_1} = \kappa_{r_2} = 1$ and, for each $\tilde{r} \in R \setminus \{r_1, r_2\}$, $\kappa_{\tilde{r}}$ is sufficiently large so that it never binds.[25]

Suppose that there are two doctors, d_1 and d_2. Finally, assume that preferences of doctors and hospitals are as follows:

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

and the capacity of each hospital is sufficiently large so that it never binds.[26]

By inspection, it is straightforward to see that the following two are the only stable matchings given the above preferences:

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

We consider two cases.

Case 1: Suppose that a mechanism produces μ given the above preference profile. Consider d_1's preferences

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

Under the preference profile $(\succ'_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$ it is straightforward to check that μ' is a unique stable matching. Note that d_1 is matched to h_3 under this new preference profile, which is strictly better under $\succ_{d_1}$ than d_1's match $\emptyset$ under the original preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$. This implies that if a mechanism is stable and produces μ given preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$, then it is not strategy-proof for doctors.

Case 2: Suppose that a mechanism produces μ' given the above preference profile. Consider d_2's preferences

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

Under the preference profile $(\succ_{d_1}, \succ'_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$ it is straightforward to check that μ is a unique stable matching. Note that d_2 is matched to h_2 under this new preference profile, which is strictly better under $\succ_{d_2}$ than d_2's match h_1 under the original preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$. This implies that if a mechanism is stable and produces μ' given preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$, then it is not strategy-proof for doctors.

[25]For example, let $\kappa_{\tilde{r}} = 3$ for each $\tilde{r} \in R \setminus \{r_1, r_2\}$.

[26]For instance, let $q_{h_1} = q_{h_2} = q_{h_3} = 3$.

6. Conclusion

This paper presented a model of matching under distributional constraints. We identified the necessary and sufficient condition on the constraint structure for the existence of a mechanism that is stable and strategy-proof for the individuals. The necessary and sufficient condition is that the constraints form a hierarchy.

The fact that our condition is both sufficient and necessary gives us a clear guide to future research. First, our sufficiency result implies that, in applications with hierarchical constraints, one can utilize our theory to design a desirable mechanism. We hope that this result will stimulate future works on specific applications. Second, our necessity result suggests that, if the constraints do not form a hierarchy, one needs to weaken either strategy-proofness for doctors or the stability concept as desiderata for a mechanism to be designed.[27] Then, the critical questions will be how such weakening should be done, and what mechanism satisfies the weakened criteria.

Finally, it is worth noting the connection between matching with constraints and matching with contracts that we used in proving the sufficiency result. This technique was subsequently adopted by other studies such as Goto, Iwasaki, Kawasaki, Yasuda, and Yokoo (2014), Goto, Hashimoto, Iwasaki, Kawasaki, Ueda, Yasuda, and Yokoo (2014), and Kojima, Tamura, and Yokoo (2015). We envision that this approach may prove useful for tackling complex matching problems one may encounter in the future.

References

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

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., 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.

[27]In addition to the applications studied in Section 4, the case with floor constraints is worth mentioning. If we translate those constraints to ceiling constraints like ours, the resulting problem does not generally admit a hierarchy. Contributions in the literature studying floor constraints seek different solutions from ours. Fragiadakis and Troyan (2015), for instance, find a mechanism that is strategy-proof for students and has desirable efficiency properties. We plan to further study floor constraints in future work.

Alkan, A. (2001): "On preferences over subsets and the lattice structure of stable matchings," *Review of Economic Design*, 6(1), 99–111.

——— (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.

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.

Blair, C. (1988): "The lattice structure of the set of stable matchings with multiple partners," *Mathematics of operations research*, 13(4), 619–628.

Budish, E., Y.-K. Che, F. Kojima, and P. R. Milgrom (2013): "Designing Random Allocation Mechanisms: Theory and Applications," *American Economic Review*, 103(2), 585–623.

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

Dur, U., P. A. Pathak, and T. Sönmez (2016): "Explicit vs. Statistical Preferential Treatment in Affirmative Action: Theory and Evidence from Chicago's Exam Schools," Discussion paper, National Bureau of Economic Research.

Echenique, F., and B. Yenmez (2015): "How to control controlled school choice," *American Economic Review*, 105(8), 2679–2694.

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

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., and P. Troyan (2015): "Improving Matching under Hard Distributional Constraints," mimeo.

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

Goto, M., N. Hashimoto, A. Iwasaki, Y. Kawasaki, S. Ueda, Y. Yasuda, and M. Yokoo (2014): "Strategy-proof matching with regional minimum quotas," *AAMAS2014*.

Goto, M., A. Iwasaki, Y. Kawasaki, Y. Yasuda, and M. Yokoo (2014): "Improving Fairness and Efficiency in Matching Markets with Regional Caps: Priority-list based Deferred Acceptance Mechanism," mimeo (the latest version is available at http://mpra.ub.uni-muenchen.de/53409/).

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2013): "Effective affirmative action in school choice," *Theoretical Economics*, 8(2), 325–363.

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.

——— (2012): "Matching in Networks with Bilateral Contracts," *American Economic Journal: Microeconomics*, 4(1), 176–208.

——— (2013): "Hidden Substitutes," Discussion paper, Working paper, Harvard Univ.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2013): "Stability and Competitive Equilibrium in Trading Networks," *Journal of Political Economy*, 121, 966–1005.

Hatfield, J. W., S. D. Kominers, and A. Westkamp (2015): "Stability, Strategy-Proofness, and Cumulative Offer Mechanisms," Discussion paper, mimeo.

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

Hirata, D., and Y. Kasuya (2015): "On stable and strategy-proof rules in matching markets with contracts," *Available at SSRN 2654639*.

Kamada, Y., and F. Kojima (2015a): "Efficient Matching under Distributional Constraints: Theory and Applications," *American Economic Review*, 105(1), 67–99.

——— (2015b): "General Theory of Matching under Distributional Constraints," mimeo.

——— (2016): "Stability Concepts in Matching with Distributional Constraints," mimeo.

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

Kojima, F., and M. Manea (2010): "Axioms for Deferred Acceptance," *Econometrica*, 78(2), 633–653.

Kojima, F., A. Tamura, and M. Yokoo (2015): "Designing Matching Mechanisms under Constraints: An Approach from Discrete Convex Analysis," mimeo.

Kominers, S., and T. Sönmez (2012): "Matching with Slot-Specific Priorities: Theory," forthcoming, *Theoretical Economics*.

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

Ministry of Health, Labour and Welfare (2008): "Measures to Address Regional Imbalance and Specialty Imbalance of Doctors (in Japanese)," http://www.mhlw.go.jp/shingi/2008/08/dl/s0823-3e.pdf.

NTT Com Research (2007): "Obstetricians Are in Short Supply! Footsteps of Obstetrics Breakdown," http://research.nttcoms.com/database/data/000687/.

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

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

——— (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.

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

Sönmez, T. (2013): "Bidding for army career specialties: Improving the ROTC branching mechanism," *Journal of Political Economy*, 121(1), 186–219.

Sönmez, T., and T. B. Switzer (2013): "Matching With (Branch-of-Choice) Contracts at the United States Military Academy," *Econometrica*, 81(2), 451–488.

Westkamp, A. (2013): "An Analysis of the German University Admissions System," *Economic Theory*, 53(3), 561–589.

APPENDIX

Appendix A defines the flexible deferred acceptance mechanism under the assumption that the set of regions forms a hierarchy. Appendix B establishes several properties of substitutability that prove useful in subsequent analysis. The proof of Theorem 1 is provided in Appendix C ((1) ⇒ (2)) and Appendix D ((2) ⇒ (1)). Appendix E provides additional discussions.

Appendix A. Flexible Deferred Acceptance Mechanism

Throughout this section, suppose that R is a hierarchy. It is straightforward to see that, for any non-singleton region $r \in R$, $\mathcal{LP}(r)$ is a singleton set. Given this fact, denote by $S(r)$ the unique element of $\mathcal{LP}(r)$, and call each element of $S(r)$ a **subregion** of r. We use simplified notation $\trianglerighteq_r$ for $\trianglerighteq_{r,S(r)}$ and $\tilde{\mathrm{Ch}}_r$ for $\tilde{\mathrm{Ch}}_{r,S(r)}$. We say that $r \in R$ is a **smallest common region** of hospitals h and h' if $h, h' \in r$, and there is no $r' \in R$ with $r' \subsetneq r$ such that $h, h' \in r'$. For any h and h', it is straightforward to see that a smallest common region of h and h' exists and is unique. Given this fact, denote the smallest common region of h and h' by $SC(h, h')$.

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 of Kamada and Kojima (2015a) has regions of depths less than or equal to 3 (H, "regions," and singleton sets), both with κ_H sufficiently large.

We proceed to define a quasi choice rule for the "hospital side," denoted $\tilde{\mathrm{Ch}}$: Let $\tilde{\kappa}_H = \kappa_H$. Given $w = (w_h)_{h \in H}$, we define $v^w_{\{h\}} = \min\{w_h, q_h, \kappa_{\{h\}}\}$ and, for each non-singleton region r, inductively define $v^w_r = \min\{\sum_{r' \in S(r)} v^w_{r'}, \kappa_r\}$. Intuitively, v^w_r 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^w_r is weakly increasing in w, that is, $w \geq w'$ implies $v^w_r \geq v^{w'}_r$.

We inductively define $\tilde{\mathrm{Ch}}(w)$ 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 $\tilde{\mathrm{Ch}}(w) = (\tilde{\kappa}^w_{\{h\}})_{h \in H}$ and stop the procedure. For each non-singleton region r of depth k, set $\tilde{\kappa}^w_{r'} = [\tilde{\mathrm{Ch}}_r((v^w_{r''})_{r'' \in S(r)}; \tilde{\kappa}^w_r)]_{r'}$ for each subregion r' of r. Go to Step $k + 1$.

That is, under $\tilde{\mathrm{Ch}}(w)$, doctors are allocated to subregions of H, and then the doctors allocated to region r are further allocated to subregions of r, and so forth until the bottom of the hierarchy is reached. In doing so, the capacity constraint of each hospital and the feasibility constraint are taken into account. For example, if the capacity is 5 at hospital h, then no more than five doctors at h are allocated to the regions containing h.

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

For each region r, fix an associated quasi choice rule $\tilde{\text{Ch}}_r$ for which the conditions for substitutability are satisfied (note that the assumption that $\trianglerighteq_r$ 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, 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 $[\tilde{\text{Ch}}(w)]_h$ 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 **flexible deferred acceptance mechanism** to be a mechanism that produces, for each input, the matching given at the termination of the above algorithm.[28]

This algorithm is a generalization of the deferred acceptance algorithm of Gale and Shapley (1962) to the model with regional caps. The main differences are found in Steps 3 and 4. Unlike the deferred acceptance algorithm, this algorithm limits the number of doctors (tentatively) matched in each region r at κ_r. This results in rationing of doctors across hospitals in the region, and the rationing rule is governed by regional preferences $\trianglerighteq_r$. Clearly, this mechanism coincides with the standard deferred acceptance algorithm if all the regional caps are large enough and hence non-binding.

[28]Note that this algorithm terminates in a finite number of steps because each doctor makes an application to a particular hospital at most once. In Appendix C we will show that the outcome of the algorithm is independent of the order in which doctors make their applications during the algorithm.

Appendix B. Remarks on Substitutability

The substitutability condition plays an important role in our proofs. This section presents three remarks on substitutability.

First, the condition in the definition of substitutability can be decomposed into two parts, as follows:

$$w \le w' \Rightarrow \tilde{\text{Ch}}_{r,S}(w;t) \ge \tilde{\text{Ch}}_{r,S}(w';t) \wedge w, \text{ and} \tag{B.1}$$

$$t \ge t' \Rightarrow \tilde{\text{Ch}}_{r,S}(w;t) \ge \tilde{\text{Ch}}_{r,S}(w;t'). \tag{B.2}$$

Condition (B.1) imposes a condition on the quasi choice rule for different vectors w and w' with a fixed parameter t while Condition (B.2) places restrictions for different parameters t and t' with a fixed vector w. The former condition requires that, given cap t, 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. This 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).[29] 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 (B.1) and (B.2) are independent of each other. One might suspect that these conditions are related to responsiveness of preferences, but these conditions do not imply responsiveness. In Appendix E.2 we provide examples to distinguish these conditions.

Second, Condition (B.1) is equivalent to

$$w \le w' \Rightarrow [\tilde{\text{Ch}}_{r,S}(w;t)]_{r'} \ge \min\{[\tilde{\text{Ch}}_{r,S}(w';t)]_{r'}, w_{r'}\} \text{ for every } r' \in S. \tag{B.3}$$

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 (B.3) is equivalent to

$$w \le w' \text{ and } [\tilde{\text{Ch}}_{r,S}(w;t)]_{r'} < [\tilde{\text{Ch}}_{r,S}(w';t)]_{r'} \Rightarrow [\tilde{\text{Ch}}_{r,S}(w,t)]_{r'} = w_{r'}. \tag{B.4}$$

To see that condition (B.3) implies condition (B.4), suppose that $w \le w'$ and $[\tilde{\text{Ch}}_{r,S}(w;t)]_{r'} < [\tilde{\text{Ch}}_{r,S}(w';t)]_{r'}$. These assumptions and condition (B.3) imply $[\tilde{\text{Ch}}_{r,S}(w;t)]_{r'} \ge w_{r'}$. Since

[29]Condition (B.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.

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

Finally, in the proof of Theorem 1, (1) $\Rightarrow$ (2), we use the fact that substitutability implies the following natural property called "consistency": A quasi choice rule $\tilde{\text{Ch}}_{r,S}$ is said to be **consistent** if for any t, $\tilde{\text{Ch}}_{r,S}(w;t) \leq w' \leq w \Rightarrow \tilde{\text{Ch}}_{r,S}(w';t) = \tilde{\text{Ch}}_{r,S}(w;t)$.[30] Consistency requires that, if $\tilde{\text{Ch}}_{r,S}(w;t)$ is chosen at w and the supply decreases to $w' \leq w$ but $\tilde{\text{Ch}}_{r,S}(w;t)$ is still available under w', then the same choice $\tilde{\text{Ch}}_{r,S}(w;t)$ 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 $\trianglerighteq_{r,S}$ because the set $\arg\max_{\trianglerighteq_{r,S}}\{w'|w' \leq w, \sum_{r'\in S} w'_{r'} \leq t\}$ may not be a singleton for some $\trianglerighteq_{r,S}$, w, and t. Note also that there always exists a consistent quasi choice rule associated with a given weak ordering $\trianglerighteq_{r,S}$.[31]

Claim 1. *Condition (B.1) implies consistency.*[32]

Proof. Fix $\trianglerighteq_{r,S}$ and its associated quasi choice rule $\tilde{\text{Ch}}_{r,S}$, and suppose that for some t, $\tilde{\text{Ch}}_{r,S}(w';t) \leq w \leq w'$. Suppose also that condition (B.1) holds. We will prove $\tilde{\text{Ch}}_{r,S}(w;t) = \tilde{\text{Ch}}_{r,S}(w';t)$. Condition (B.1) implies $w \leq w' \Rightarrow \tilde{\text{Ch}}_{r,S}(w;t) \geq \tilde{\text{Ch}}_{r,S}(w';t) \wedge w$. Since $\tilde{\text{Ch}}_{r,S}(w';t) \leq w$ implies $\tilde{\text{Ch}}_{r,S}(w';t) \wedge w = \tilde{\text{Ch}}_{r,S}(w';t)$, this means that $\tilde{\text{Ch}}_{r,S}(w';t) \leq \tilde{\text{Ch}}_{r,S}(w;t) \leq w'$. If $\tilde{\text{Ch}}_{r,S}(w;t) \neq \tilde{\text{Ch}}_{r,S}(w';t)$ then by the assumption that $\tilde{\text{Ch}}_{r,S}$ is acceptant, we must have $\tilde{\text{Ch}}_{r,S}(w;t) \triangleright_{r,S} \tilde{\text{Ch}}_{r,S}(w';t)$. But then $\tilde{\text{Ch}}_{r,S}(w';t)$ cannot be an element of $\arg\max_{\trianglerighteq_{r,S}}\{w''|w'' \leq w', \sum_{r'\in S} w''_{r'} \leq t\}$ because $\tilde{\text{Ch}}_{r,S}(w;t) \in \{w''|w'' \leq w', \sum_{r'\in S} w''_{r'} \leq t\}$. Hence we have $\tilde{\text{Ch}}_{r,S}(w';t) = \tilde{\text{Ch}}_{r,S}(w;t)$. □

Appendix C. Proof of Theorem 1, (1) $\Rightarrow$ (2)

With the definition of the flexible deferred acceptance mechanism, we are now ready to present the following statement.

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

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

[32]Fleiner (2003) and Aygün and Sönmez (2012) prove analogous results although they do not work on substitutability defined over the space of integer vectors. Conditions that are analogous to our consistency concept are used by Blair (1988), Alkan (2002), and Alkan and Gale (2003) in different contexts.

Suppose that R is a hierarchy and $\trianglerighteq_r$ is substitutable for every $r \in R$. Then the flexible deferred acceptance mechanism produces a stable matching for any input and is group strategy-proof for doctors.

This statement suffices to show Theorem 1, (1) $\Rightarrow$ (2). Therefore, the remainder of this section establishes the above statement.

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.[33] 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 and its associated choice rule $\text{Ch}(\cdot)$ 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 $\text{Ch}(X')$ is defined by

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

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

Definition 2 (Hatfield and Milgrom (2005))**.** Choice rule $\text{Ch}(\cdot)$ satisfies the **substitutes condition** if there do not exist contracts $x, x' \in X$ and a set of contracts $X' \subseteq X$ such that $x' \notin \text{Ch}(X' \cup \{x'\})$ and $x' \in \text{Ch}(X' \cup \{x, x'\})$.

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 4) when contracts are substitutes for the hospital side.

[33] 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 3 (Hatfield and Milgrom (2005))**.** Choice rule $\text{Ch}(\cdot)$ satisfies the **law of aggregate demand** if for all $X' \subseteq X'' \subseteq X$, $|\text{Ch}(X')| \le |\text{Ch}(X'')|$.[34]

Proposition 1. *Suppose that $\trianglerighteq_r$ is substitutable for all $r \in R$.*

(1) *Choice rule* $\text{Ch}(\cdot)$ *defined above satisfies the substitutes condition.*[35]
(2) *Choice rule* $\text{Ch}(\cdot)$ *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 $(d',h') \notin \text{Ch}(X' \cup \{(d',h')\})$ and $(d',h') \in \text{Ch}(X' \cup \{(d,h),(d',h')\})$. 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 $\tilde{\kappa}^{w'}_{\{h'\}} < \tilde{\kappa}^{w''}_{\{h'\}}$. To see this, note that otherwise we would have $\tilde{\kappa}^{w'}_{\{h'\}} \ge \tilde{\kappa}^{w''}_{\{h'\}}$, hence by the definition of Ch we must have $[\text{Ch}(X' \cup \{(d',h')\})]_{h'} \supseteq [\text{Ch}(X' \cup \{(d,h),(d',h')\})]_{h'} \setminus \{(d,h)\}$. This contradicts $(d',h') \notin \text{Ch}(X' \cup \{(d',h')\})$ and $(d',h') \in \text{Ch}(X' \cup \{(d,h),(d',h')\})$.

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

$$[\tilde{\text{Ch}}_r(v'; \tilde{\kappa}^{w'}_r)]_{r'} \ge \min\{[\tilde{\text{Ch}}_r(v''; \tilde{\kappa}^{w''}_r)]_{r'}, v'_{r'}\}.$$

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

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

this means $[\tilde{\text{Ch}}_r(v'; \tilde{\kappa}^{w'}_r)]_{r'} \ge v'_{r'}$. But then by $[\tilde{\text{Ch}}_r(v'; \tilde{\kappa}^{w'}_r)]_{r'} \le v'_{r'}$ (from the definition of $\tilde{\text{Ch}}$) we have $[\tilde{\text{Ch}}_r(v'; \tilde{\kappa}^{w'}_r)]_{r'} = v'_{r'}$. This contradicts the assumption that $(d',h') \notin \text{Ch}(X' \cup \{(d',h')\})$, while d' is acceptable to h' (because $(d',h') \in \text{Ch}(X' \cup \{(d,h),(d',h')\})$). Thus we must have that $\tilde{\kappa}^{w'}_r < \tilde{\kappa}^{w''}_r$.

[34] Analogous conditions called cardinal monotonicity and size monotonicity are introduced by Alkan (2002) and Alkan and Gale (2003) for matching models without contracts.

[35] Note that choice rule $\text{Ch}(\cdot)$ allows for the possibility that multiple contracts are signed between the same pair of a region and a doctor. Without this possibility, the choice rule may violate the substitutes condition (Sönmez and Switzer, 2013; Sönmez, 2013). Hatfield and Kominers (2013) explore this issue further.

Step 3: Step 1 and an iterative use of Step 2 imply that $\tilde{\kappa}_H^{w'} < \tilde{\kappa}_H^{w''}$. But we specified $\tilde{\kappa}_H^w$ for any w to be equal to κ_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 $|\text{Ch}(X')| \leq |\text{Ch}(X' \cup \{(d, h)\})|$. 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 $\sum_{r' \in S(r)} v_{r'}^w \geq \kappa_r$ 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 $v_r^w = \min\{\sum_{r' \in S(r)} v_{r'}^w, \kappa_r\}$, the assumption $\sum_{r' \in S(r)} v_{r'}^w \geq \kappa_r$ implies $v_r(w) = \kappa_r$. By the same argument, we also obtain $v_r(w') = \kappa_r$. Thus, for any r' such that $r \subseteq r'$, we inductively obtain $v_{r'}^{w'} = v_{r'}^w$. □

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

$$[\tilde{\text{Ch}}(w')]_{h'} = [\tilde{\text{Ch}}(w)]_{h'} \text{ for any } h' \notin r. \tag{C.1}$$

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

Now note the following: For any $r' \in R$, since $v_{r'}^w$ is defined as $\min\{\sum_{r'' \in S(r')} v_{r''}^w, \kappa_{r'}\}$ and all regional preferences are acceptant, the entire assigned regional cap $\tilde{\kappa}_{r'}^w$ is allocated to some subregion of r', that is, $\tilde{\kappa}_{r'}^w = \sum_{r'' \in S(r')} \tilde{\kappa}_{r''}^w$. Similarly we also have $\tilde{\kappa}_{r'}^{w'} = \sum_{r'' \in S(r')} \tilde{\kappa}_{r''}^{w'}$. This is the case 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 the relation

$$\sum_{h' \in r} [\tilde{\text{Ch}}(w')]_{h'} = \sum_{h' \in r} [\tilde{\text{Ch}}(w)]_{h'}. \tag{C.2}$$

By (C.1) and (C.2), we conclude that

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

completing the proof for this case.

(2) Suppose $\sum_{r' \in S(r)} v^w_{r'} < \kappa_r$ 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

$$[\tilde{\mathrm{Ch}}(w')]_h = [\tilde{\mathrm{Ch}}(w)]_h + 1. \tag{C.3}$$

In addition, the following claim holds.

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

Proof. First, note that $v^{w'}_r = v^w_r + 1$ 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, κ_H has not been reached under w, that is, $\sum_{r' \in S(H)} v^w_{r'} < \kappa_H$. Thus, since $\tilde{\mathrm{Ch}}_H$ is acceptant, the entire vector $(v_{r'}(w))_{r' \in S(H)}$ is accepted by $\tilde{\mathrm{Ch}}_H$, that is, $\tilde{\kappa}^w_{r'} = v^w_{r'}$. 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, $\tilde{\kappa}^w_{r'} = \tilde{\kappa}^{w'}_{r'}$ 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 κ_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 $\kappa_{r'}$ has not been binding. And (2) guarantees that we can follow the same argument inductively, so the conclusion holds for all $h \neq h'$. □

By equation (C.3) and Claim 3, we obtain

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

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

□

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) $\mathrm{Ch}(X') = X'$.

Definition 4. 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 $(d,h) \in \mathrm{Ch}(X' \cup \{(d,h)\})$, 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 2. *Suppose that $\trianglerighteq_r$ 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. First, the following observation is straightforward.

Observation 1. *Suppose that R is a hierarchy. Then a matching μ is stable if and only if it is feasible, individually rational, and if (d,h) is a blocking pair then there exists $r \in R$ with $h \in r$ such that (i) $|\mu_r| = \kappa_r$, (ii) $d' \succ_h d$ for all doctors $d' \in \mu_h$, and (iii) either $\mu_d \notin r$ or $(w_{r'})_{r' \in S(SC(h,\mu_d))} \trianglerighteq_{SC(h,\mu_d)} (w'_{r'})_{r' \in S(SC(h,\mu_d))}$, 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 $\mu_d \in r_d$.*

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 above observation, it suffices to show that there exists a region r that includes h such that the following conditions (C.4), (C.5), and $\mu_d \notin r$ hold, or (C.4), (C.5), (C.6), and $h, \mu_d \in r$ hold:

$$|\mu_r| = \kappa_r, \tag{C.4}$$

$$d' \succ_h d \text{ for all } d' \in \mu_h, \tag{C.5}$$

$$(w_{r''})_{r'' \in S(SC(h,\mu_d))} \trianglerighteq_{SC(h,\mu_d)} (w'_{r''})_{r'' \in S(SC(h,\mu_d))}, \tag{C.6}$$

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 (C.5) holds, and there exists r that includes h such that Condition (C.4) 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 (C.5) is violated, that is, $d \succ_h d'$ for some $d' \in \mu_h$. First, note that $[\tilde{\text{Ch}}(w'')]_h \geq [\tilde{\text{Ch}}(w)]_h$. 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 $h \in r'$,

$$\text{(C.7)} \qquad [\tilde{\text{Ch}}_r((v^{w''}_{r''})_{r'' \in S(r)}; \tilde{\kappa}_r)]_{r'} \geq [\tilde{\text{Ch}}_r((v^{w}_{r''})_{r'' \in S(r)}; \tilde{\kappa}'_r)]_{r'} \text{ if } \tilde{\kappa}_r \geq \tilde{\kappa}'_r.$$

To see this, first note that $[\tilde{\text{Ch}}_r((v^{w}_{r''})_{r'' \in S(r)}; \tilde{\kappa}_r)]_{r'} \geq [\tilde{\text{Ch}}_r((v^{w}_{r''})_{r'' \in S(r)}; \tilde{\kappa}'_r)]_{r'}$ by substitutability of $\trianglerighteq_r$. Also, by consistency of $\tilde{\text{Ch}}_r$ and $v^{w''}_{r''} \geq v^{w}_{r''}$ for every region r'', the inequality

$$[\tilde{\text{Ch}}_r((v^{w''}_{r''})_{r'' \in S(r)}; \tilde{\kappa}_r)]_{r'} \geq [\tilde{\text{Ch}}_r((v^{w}_{r''})_{r'' \in S(r)}; \tilde{\kappa}_r)]_{r'}$$

follows,[36] showing condition (C.7). An iterative use of condition (C.7) gives us the desired result that $[\tilde{\text{Ch}}(w'')]_h \geq [\tilde{\text{Ch}}(w)]_h$. This property, together with the assumptions that $d \succ_h d'$ and that $(d',h) \in X'$ imply that $(d,h) \in \text{Ch}(X' \cup (d,h))$.[37] 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 (C.4) is violated, so that $|\mu_r| \neq \kappa_r$ for every r that includes h. Then, for such r, since $|\mu_r| \leq \kappa_r$ by the construction of

[36]To show this claim, let $v = (v^{w}_{r''})_{r'' \in S(r)}$ and $v'' = (v^{w''}_{r''})_{r'' \in S(r)}$ for notational simplicity and assume for contradiction that $[\tilde{\text{Ch}}_r(v''; \tilde{\kappa}_r)]_{r'} < [\tilde{\text{Ch}}_r(v; \tilde{\kappa}_r)]_{r'}$. Then, $[\tilde{\text{Ch}}_r(v''; \tilde{\kappa}_r)]_{r'} < [\tilde{\text{Ch}}_r(v; \tilde{\kappa}_r)]_{r'} \leq v_{r'}$. Moreover, since $v''_{r''} = v_{r''}$ for every $r'' \neq r'$ by the construction of v'', it follows that $[\tilde{\text{Ch}}_r(v'')]_{r''} \leq v''_{r''} = v_{r''}$. Combining these inequalities, we have that $\tilde{\text{Ch}}_r(v'') \leq v$. Also we have $v \leq v''$ by the definition of v'', so it follows that $\tilde{\text{Ch}}_r(v'') \leq v \leq v''$. Thus, by consistency of $\tilde{\text{Ch}}_r$, we obtain $\tilde{\text{Ch}}_r(v'') = \tilde{\text{Ch}}_r(v)$, a contradiction to the assumption $[\tilde{\text{Ch}}_r(v'')]_{r'} < [\tilde{\text{Ch}}_r(v)]_{r'}$.

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

μ and the assumption that X' is individually rational, it follows that $|\mu_r| < \kappa_r$. Then $(d,h) \in \text{Ch}(X' \cup (d,h))$ because,

(a) $d \succ_h \emptyset$ by assumption,

(b) since $\sum_{r' \in S(r)} w_{r'} = \sum_{h \in r} |\mu_h| = |\mu_r| < \kappa_r$, it follows that $\sum_{r' \in S(r)} w''_{r'} = \sum_{r' \in S(r)} w_{r'} + 1 \leq \kappa_r$. This property and the fact that $\tilde{\text{Ch}}_r$ is acceptant and the definition of the function $v_{r'}$ for regions r' imply that $\tilde{\text{Ch}}(w'') = w''$. In particular, this implies that every contract $(d',h) \in X' \cup (d,h)$ such that $d' \succ_h \emptyset$ is chosen at $\text{Ch}(X' \cup (d,h))$.

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.

□

To finish the proof of the proposition suppose for contradiction that there is no r that includes h such that (C.4), (C.5), and $\mu_d \notin r$ hold, and that condition (C.6) fails. That is, we suppose $(w'_{r''})_{r'' \in S(SC(h,\mu_d))} \rhd_{SC(h,\mu_d)} (w_{r''})_{r'' \in S(SC(h,\mu_d))}$. Then it must be the case that $[\tilde{\text{Ch}}_r((v^{w''}_{r''})_{r'' \in S(SC(h,\mu_d))}; \tilde{\kappa}^{w''}_{SC(h,\mu_d)})]_{r'} = w''_{r'} = w_{r'} + 1 = |\mu_h| + 1$, where $h \in r'$ and $\tilde{\kappa}^{w''}_{SC(h,\mu_d)}$ is as defined in the procedure to compute $\tilde{\text{Ch}}(w'')$.[38] Note that for all r' such that $h \in r'$ and $r' \subsetneq SC(h,\mu_d)$, it follows that $\mu_d \notin r'$. Also note that (C.5) is satisfied by Claim 4. Therefore we have $|\mu_{r'}| < \kappa_{r'}$ for all $r' \subsetneq SC(h,\mu_d)$ that includes h by assumption and

[38] To show this claim, assume for contradiction that $[\tilde{\text{Ch}}_{SC(h,\mu_d)}((v^{w''}_{r''})_{r'' \in S(SC(h,\mu_d))}; \tilde{\kappa}^{w''}_{SC(h,\mu_d)})]_{r'} \leq w_{r'}$ 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

$$\tilde{\text{Ch}}_{SC(h,\mu_d)}(v''; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}) \leq (w_{r''})_{r'' \in SC(h,\mu_d)} \leq (w''_{r''})_{r'' \in SC(h,\mu_d)}.$$

But $\tilde{\text{Ch}}_{SC(h,\mu_d)}(v; \tilde{\kappa}^{w}_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$ 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

$$\tilde{\text{Ch}}_{SC(h,\mu_d)}(v''; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}) \leq v \leq v''.$$

Thus by consistency of $\tilde{\text{Ch}}_{SC(h,\mu_d)}$, we obtain

$$\tilde{\text{Ch}}_{SC(h,\mu_d)}(v''; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}) = \tilde{\text{Ch}}_{SC(h,\mu_d)}(v; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}).$$

But again by $\tilde{\text{Ch}}_{SC(h,\mu_d)}(v; \tilde{\kappa}^{w}_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$, by substitutability we obtain $\tilde{\text{Ch}}_{SC(h,\mu_d)}(v; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$, thus $\tilde{\text{Ch}}_{SC(h,\mu_d)}(v''; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}) = (w_{r''})_{r'' \in SC(h,\mu_d)}$. This is a contradiction because $(w'_{r''})_{r'' \in SC(h,\mu_d)} \leq (w''_{r''})_{r'' \in SC(h,\mu_d)} = v''$ and $(w'_{r''})_{r'' \in SC(h,\mu_d)} \rhd_{SC(h,\mu_d)} (w_{r''})_{r'' \in SC(h,\mu_d)}$ while $\tilde{\text{Ch}}_{SC(h,\mu_d)}(v''; \tilde{\kappa}^{w''}_{SC(h,\mu_d)}) \in \arg\max_{\trianglerighteq_{SC(h,\mu_d)}} \{(w'''_{r''})_{r'' \in SC(h,\mu_d)} | (w'''_{r''})_{r'' \in SC(h,\mu_d)} \leq v'', \sum_{r'' \in S(SC(h,\mu_d))} w'''_{r''} \leq \tilde{\kappa}^{w''}_{SC(h,\mu_d)}\}$.

hence $|\mu_{r'}| + 1 \le \kappa_{r'}$ for all such r'. Moreover we have $d \succ_h \emptyset$, thus

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

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.

□

Remark 5. The definition of stability in this paper is based on Pareto improvement for multiple regions. For the case of hierarchies, we provide an alternative interpretation here. The idea of Condition (iii) in Observation 1 is to invoke a region's preferences 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 $SC(h, \mu_d)$. 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. □

Remark 6. Each step of the flexible deferred acceptance algorithm corresponds to a step of the cumulative offer process (Hatfield and Milgrom, 2005), that is, at each step, if doctor d proposes to hospital 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. Therefore, if X' is the allocation that is produced by the cumulative offer process, then $\mu(X')$ is the matching produced by the flexible deferred acceptance algorithm.

Proof of Theorem 1, (1) ⇒ (2). By Proposition 1, the choice rule $\text{Ch}(\cdot)$ satisfies the substitutes condition and the law of aggregate demand in the associated model of matching with contracts. By Hatfield and Milgrom (2005), Hatfield and Kojima (2009), and Hatfield and Kominers (2012), the cumulative offer process with choice rules satisfying these conditions produces a stable allocation and is (group) strategy-proof.[39] The former fact,

[39]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.

together with Remark 6 and Proposition 2, implies that the outcome of the flexible deferred acceptance algorithm is a stable matching in the original model. The latter fact and Remark 6 imply that the flexible deferred acceptance mechanism is (group) strategy-proof for doctors. □

Appendix D. Proof of Theorem 1, (2) ⇒ (1)

Fix R, and suppose that it is not a hierarchy. Then, there exist $r, r' \in R$ and $h_1, h_2, h_3 \in H$ such that $h_1 \in r \setminus r'$, $h_2 \in r \cap r'$, and $h_3 \in r' \setminus r$. We will show that there does not exist a mechanism that is stable and strategy-proof for doctors.

To see this, first pick $r_1 \in R$ such that (i) $\{h_1, h_2\} \subseteq r_1 \subseteq r$ and (ii) there is no $\tilde{r} \in R$ with the property that $\{h_1, h_2\} \subseteq \tilde{r} \subsetneq r_1$. Similarly, pick $r_2 \in R$ such that (i) $\{h_2, h_3\} \subseteq r_2 \subseteq r'$ and (ii) there is no $\tilde{r} \in R$ with the property that $\{h_2, h_3\} \subseteq \tilde{r} \subsetneq r_2$.

By the construction of r_1, there exist $S_1 \subseteq R$ and $\hat{r}_1, \hat{r}_2 \in R$ such that (i) S_1 is a largest partition of r_1, (ii) $\hat{r}_1 \in S_1$ and $h_1 \in \hat{r}_1$, and (iii) $\hat{r}_2 \in S_1$ and $h_2 \in \hat{r}_2$. Similarly, by the construction of r_2, there exist $S_2 \subseteq R$ and $\tilde{r}_2, \tilde{r}_3 \in R$ such that (i) S_2 is a largest partition of r_2, (ii) $\tilde{r}_2 \in S_2$ and $h_2 \in \tilde{r}_2$, and (iii) $\tilde{r}_3 \in S_2$ and $h_3 \in \tilde{r}_3$.

Let $\hat{w}^1$ be a vector of nonnegative integers over the set S_1 such that the coordinate corresponding to $\hat{r}_1$ is 1 and other coordinates are 0. Also, let $\hat{w}^2$ be a vector of nonnegative integers over the set S_1 such that the coordinate corresponding to $\hat{r}_2$ is 1 and other coordinates are 0. Suppose that $\hat{w}^1 \rhd_{r_1,S_1} \hat{w}^2$. Similarly, let $\tilde{w}^2$ be a vector of nonnegative integers over the set S_2 such that the coordinate corresponding to $\tilde{r}_2$ is 1 and other coordinates are 0. Also, let $\tilde{w}^3$ be a vector of nonnegative integers over the set S_2 such that the coordinate corresponding to $\tilde{r}_3$ is 1 and other coordinates are 0. Suppose that $\tilde{w}^2 \rhd_{r_2,S_2} \tilde{w}^3$.

Let $\kappa_{r_1} = \kappa_{r_2} = 1$, and $\kappa_{\bar{r}} = |D| + 1$ for all $\bar{r} \in R \setminus \{r_1, r_2\}$. Fix two doctors d_1 and d_2 in D. Finally, assume that preferences of doctors and hospitals are as follows:

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

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

the capacity of h_1, h_2, and h_3 are sufficiently large so that they never bind,[40] and all doctors in $D \setminus \{d_1, d_2\}$ regard all hospitals unacceptable.[41]

Since the choice rules in our context satisfy the substitutes condition and the law of aggregate demand, the conclusions go through.

[40]For instance, let $q_{h_1} = q_{h_2} = q_{h_3} = |D| + 1$.

[41]Preferences for hospitals in $H \setminus \{h_1, h_2, h_3\}$ can be arbitrary.

By inspection, it is straightforward to see that the following two are the only stable matchings given the above preferences:

$$\mu = \begin{pmatrix} h_1 & h_2 & h_3 & \text{other hospitals} & \emptyset & \emptyset \\ \emptyset & d_2 & \emptyset & \cdots\emptyset\cdots & d_1 & \text{other doctors} \end{pmatrix},$$

$$\mu' = \begin{pmatrix} h_1 & h_2 & h_3 & \text{other hospitals} & \emptyset \\ d_2 & \emptyset & d_1 & \cdots\emptyset\cdots & \text{other doctors} \end{pmatrix}.$$

We consider two cases.

<u>Case 1:</u> Suppose that a mechanism produces μ given the above preference profile. Consider d_1's preferences

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

Under the preference profile $(\succ'_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$ it is straightforward to check that μ' is a unique stable matching. Note that d_1 is matched to h_3 under this new preference profile, which is strictly better under $\succ_{d_1}$ than d_1's match $\emptyset$ under the original preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$. This implies that if a mechanism is stable and produces μ given preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$, then it is not strategy-proof for doctors.

<u>Case 2:</u> Suppose that a mechanism produces μ' given the above preference profile. Consider d_2's preferences

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

Under the preference profile $(\succ_{d_1}, \succ'_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$ it is straightforward to check that μ is a unique stable matching. Note that d_2 is matched to h_2 under this new preference profile, which is strictly better under $\succ_{d_2}$ than d_2's match h_1 under the original preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$. This implies that if a mechanism is stable and produces μ' given preference profile $(\succ_{d_1}, \succ_{d_2}, \succ_{h_1} \succ_{h_2}, \succ_{h_3})$, then it is not strategy-proof for doctors.

Appendix E. Additional Discussions

E.1. **Group strategy-proofness.** The statement of Theorem 1 holds when we strengthen the incentive compatibility requirement. A mechanism φ is said to be **group strategy-proof for doctors** if there are no $(\kappa, \trianglerighteq)$, 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

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

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. Clearly, this property is stronger than strategy-proofness for doctors. The proof in the Appendix shows that the statement of Theorem 1 holds when we replace strategy-proofness for doctors with group strategy-proofness for doctors.

E.2. **Further Discussion on Substitutability.** The following examples show that conditions (B.1) and (B.2) of substitutability are independent.

Example 2 (Regional preferences that violate (B.1) while satisfying (B.2))**.** There is a region r in which two hospitals h_1 and h_2 reside. $S = \{\{h_1\}, \{h_2\}\}$ is the unique largest partition of r. The capacity of each hospital is 2. Region r's preferences are as follows.

$$\rhd_{r,S} : (2,2), (2,1), (1,2), (2,0), (0,2), (1,1), (1,0), (0,1), (0,0).$$

One can check by inspection that condition (B.2) and consistency are satisfied. To show that (B.1) is not satisfied, observe first that there is a unique associated choice rule (since preferences are strict), and denote it by $\tilde{\text{Ch}}_{r,S}$. The above preferences imply that $\tilde{\text{Ch}}_{r,S}((1,2);2) = (0,2)$ and $\tilde{\text{Ch}}_{r,S}((2,2);2) = (2,0)$. But this is a contradiction to (B.1) because $(1,2) \leq (2,2)$ but $\tilde{\text{Ch}}_{r,S}((1,2);2) \geq \tilde{\text{Ch}}_{r,S}((2,2);2) \wedge (1,2)$ does not hold (the left hand side is $(0,2)$ while the right hand side is $(1,0)$). □

Example 3 (Regional preferences that violate (B.2) while satisfying (B.1))**.** There is a region r in which three hospitals h_1, h_2, and h_3 reside. $S = \{\{h_1\}, \{h_2\}, \{h_3\}\}$ is the unique largest partition of r. The capacity of each hospital is 1. Region r's preferences are as follows.

$$\rhd_{r,S} : (1,1,1), (1,1,0), (1,0,1), (0,1,1), (0,0,1), (0,1,0), (1,0,0), (0,0,0).$$

One can check by inspection that condition (B.1) (and hence consistency by Claim 1) are satisfied. To show that (B.2) is not satisfied, observe first that there is a unique associated choice rule (since preferences are strict), and denote it by $\tilde{\text{Ch}}_{r,S}$. The above preferences imply that $\tilde{\text{Ch}}_{r,S}((1,1,1);1) = (0,0,1)$ and $\tilde{\text{Ch}}_{r,S}((1,1,1);2) = (1,1,0)$. But this is a contradiction to (B.2) because $1 \leq 2$ but $\tilde{\text{Ch}}_{r,S}((1,1,1);1) \leq \tilde{\text{Ch}}_{r,S}((1,1,1);2)$ does not hold (the left hand side is $(0,0,1)$ while the right hand side is $(1,1,0)$). □