Contract Design and Stability in Many-to-Many Matching☆

John William Hatfield

McCombs School of Business, University of Texas at Austin

Scott Duke Kominers*

Society of Fellows, Department of Economics, Center of Mathematical Sciences and Applications, Center for Research on Computation and Society, and Program for Evolutionary Dynamics, Harvard University, Cambridge, MA, United States

Harvard Business School, Boston, MA, United States

National Bureau of Economic Research, Cambridge, MA, United States

Abstract

We develop a model of many-to-many matching with contracts that subsumes as special cases many-to-many matching markets and buyer–seller markets with heterogeneous and indivisible goods. In our setting, substitutable preferences are sufficient to guarantee the existence of stable outcomes; moreover, in contrast to results for the setting of many-to-one matching with contracts, if any agent's preferences are not substitutable, then the existence of a stable outcome cannot be guaranteed.

In many-to-many matching with contracts, a new market design issue arises: The design of the contract language can impact the set of stable outcomes. Bundling contractual primitives encourages substitutability of agents' preferences over contracts and makes stable outcomes more likely to exist; however, bundling also makes the contractual language less expressive. Consequently, in choosing contract language, market designers face a tradeoff between expressiveness and stability.

Key words: Many-to-Many Matching, Stability, Substitutes, Contract Design
JEL: C78, C62, D47, L14

1. Introduction

We develop a model of many-to-many matching in which agents on two opposing sides of a market negotiate over contractual relationships, possibly signing multiple contracts. This setting models several real-world matching markets, such as the United Kingdom Medical Intern match (see Roth and Sotomayor (1990)), the market used to allocate blood from blood banks to hospitals (see Jaume et al. (2012)), and the market for advertising within mobile applications (see Lee et al. (2014)). One important special case of

☆This paper contains parts of two earlier working papers entitled "Contract Design and Stability in Matching Markets" and "Many-to-Many Matching with Contracts." The authors appreciate the helpful comments of Eric Budish, Jeremy Bulow, Peter Cramton, Federico Echenique, Sonia Jaffe, Fuhito Kojima, David Laibson, Paul Milgrom, Alexandru Nichifor, Muriel Niederle, Michael Ostrovsky, David Parkes, Philip Reny, Alvin E. Roth, Tayfun Sönmez, Utku Ünver, Alexander Westkamp, E. Glen Weyl, several referees, the editors (Yan Chen and Thomas Palfrey), seminar audiences at the AEA Meetings, Chicago, Harvard, Northwestern, and Stanford, and especially Drew Fudenberg. Much of this work was completed while Hatfield was at the Stanford University Graduate School of Business and Kominers was at the Becker Friedman Insitute at the University of Chicago. Hatfield appreciates the hospitality of Harvard Business School, which hosted him during parts of this research. Additionally, Kominers gratefully acknowledges the support of a National Science Foundation Graduate Research Fellowship, National Science Foundation Grants CCF-1216095 and SES-1459912, the Harvard Milton Fund, a Yahoo! Key Scientific Challenges Program Fellowship, a Terence M. Considine Fellowship in Law and Economics funded by the John M. Olin Center at Harvard Law School, and an AMS–Simons Travel Grant.

*Arthur Rock Center for Entrepreneurship, Harvard Business School, Soldiers Field, Boston, MA, 02163, United States.
Email addresses: `john.hatfield@utexas.edu` (John William Hatfield), `kominers@fas.harvard.edu` (Scott Duke Kominers)
URL: `http://www.mccombs.utexas.edu/directory/profiles/hatfield-john` (John William Hatfield), `http://www.scottkom.com/` (Scott Duke Kominers)

our model is matching with couples, in which pairs of individuals may choose to act as a single agent that receives (at most) two assignments (see Klaus and Klijn (2005) and Klaus et al. (2007)).[1] Our model includes as special cases many-to-one matching with contracts (Kelso and Crawford (1982); Hatfield and Milgrom (2005)), many-to-many matching (Sotomayor (1999, 2004); Echenique and Oviedo (2006); Konishi and Ünver (2006)), many-to-many matching with wages (Roth (1985); Blair (1988)), and buyer–seller markets with heterogeneous and indivisible goods.[2]

We show that stable outcomes are guaranteed to exist in the setting of many-to-many matching with contracts when preferences are *substitutable* in the sense that no contract becomes desirable when some other contract becomes available.[3] Moreover, substitutability is necessary for the existence of stable outcomes in the maximal domain sense: if any one agent has preferences that are not substitutable, then there exist substitutable preferences for the other agents such that no stable outcome exists. Our maximal domain result is particularly surprising because no analogous result holds in the Hatfield and Milgrom (2005) model of many-to-one matching with contracts (see Hatfield and Kojima (2008, 2010), Hatfield and Kominers (2016), and Hatfield et al. (2015)).

We discuss the structure of the set of stable outcomes, noting extensions of standard lattice structure and rural hospitals results. We then show that when hospital preferences satisfy a more stringent condition than substitutability (so-called *strong substitutability*) and doctor preferences are substitutable, stability is equivalent to a more stringent solution concept: *strong stability*.[4]

Modeling many-to-many matching with contracts raises a subtle conceptual issue: Whereas in many-to-one matching with contracts the entire relationship between two agents must be specified by a single contract, this requirement—which Kominers (2012) calls *unitarity*—is not necessary in many-to-many matching with contracts. Non-unitarity is present in many important applications. For instance, in the United Kingdom Medical Intern match, a student must find both a surgical and a medical position, and hospitals typically hire multiple students. In that context, students are assigned two separate contracts by the match, even if they end up at the same hospital; that is, the United Kingdom Medical Intern match as practiced is non-unitary. In principle, however, one could bundle contractual terms for any application to impose unitarity—i.e., to represent every possible relationship between a doctor–hospital pair as a single contractual relationship—from the designer's perspective. One key contribution of our work is to show that such bundling may not be optimal: for some applications, including the United Kingdom Medical Intern match, market designers may not want to require unitarity and would prefer to leave contracts unbundled.[5]

Allowing multiple contracts between a doctor–hospital pair highlights the importance of *contract language* design, where by contract language we mean the set of possible relationships between a doctor and a hospital that can be expressed as part of a contractual outcome. More generally, however, the choice of a particular contract language is crucial in determining the set of stable outcomes. For example, consider a setting with a doctor d and a hospital h. Contracts can specify one or two of the following terms: the doctor works in

[1] In the United States National Resident Matching Program (NRMP), doctors may apply to as a couple, submitting a preference list over pairs of job assignments, and potentially being assigned to two jobs (see Roth and Peranson (1999)).

[2] Our model is substantively different from the only previous model of many-to-many matching with contracts—that of Klaus and Walzl (2009)—as we allow a given doctor and hospital to sign multiple contracts with each other. This distinction is material to our results, as we discuss in Section 3.4.

[3] We show that this substitutability concept has a natural interpretation in terms of utility theory: preferences over contracts are substitutable if and only if they can be represented by a submodular indirect utility function over sets of offered contracts (see Section 2.1).

[4] Unlike in many-to-one matching, the set of core many-to-many matchings does not generally correspond to the set of stable many-to-many matchings (Blair (1988); see also Echenique and Oviedo (2006) and Konishi and Ünver (2006)). This problem is still extant in the more general setting of many-to-many matching with contracts; hence, we follow Echenique and Oviedo (2006) and Klaus and Walzl (2009) in studying a solution concept alternative to and stronger than stability. Our strengthened stability concept, *strong stability*, is stronger than the similar concept of *setwise stability* studied by Echenique and Oviedo (2006) and Klaus and Walzl (2009).

[5] Unitarity is sometimes problematic for technical reasons, as well (see Section 3.4). Nevertheless, imposing unitarity can sometimes be beneficial because of the additional structure it adds: As Echenique (2012) showed, the many-to-one matching with contracts model with substitutable preferences embeds into the seemingly simpler matching with salaries (and gross substitutes) framework of Kelso and Crawford (1982); this allows efficient proofs of the main existence and lattice results for matching with contracts, by appeal to the analogous results for the Kelso and Crawford (1982) framework. The Echenique (2012) result carries over to many-to-one matching with contracts settings with unilaterally substitutable preferences (although via a different embedding; see Schlegel (2015)) and to unitary many-to-many matching with contracts models (Kominers (2012)).

the morning (m); the doctor works in the afternoon (a). The doctor would most prefer to work in both the morning and the afternoon, but would be willing to work just the afternoon shift; he is unwilling to work only the morning shift. The hospital would hire the doctor for any shift—and for both shifts—but would most prefer that the doctor work only in the morning, and would rather hire the doctor full-time than for just the afternoon. We denote by x^Γ the contract with terms $\Gamma \subseteq \{m, a\}$. When morning and afternoon shifts are contracted separately, the doctor's preferences over contracts are given by

$$P_d : \left\{x^{\{m\}}, x^{\{a\}}\right\} \succ \left\{x^{\{a\}}\right\} \succ \varnothing \succ \left\{x^{\{m\}}\right\},$$

while the hospital's preferences are given by

$$P_h : \left\{x^{\{m\}}\right\} \succ \left\{x^{\{m\}}, x^{\{a\}}\right\} \succ \left\{x^{\{a\}}\right\} \succ \varnothing.$$

There is no stable contracting outcome: for the set $\left\{x^{\{m\}}, x^{\{a\}}\right\}$, the hospital will not be willing to sign $x^{\{a\}}$; for the set $\left\{x^{\{a\}}\right\}$, both parties prefer that the doctor work full time; the set $\left\{x^{\{m\}}\right\}$ is not individually rational for the doctor; and finally both parties agree that $\left\{x^{\{a\}}\right\}$ is better than no relationship at all. This lack of agreement derives from the fact that the preferences of doctor d are not substitutable—there are two contracts ($x^{\{m\}}$ and $x^{\{a\}}$) that exhibit "complementarity" for d, in the sense that d wants one ($x^{\{m\}}$) only if he has the other ($x^{\{a\}}$).

By contrast, if the parties are to negotiate over a single contract $x^{\{m,a\}}$ that encodes both the morning and afternoon shifts, the agents' preferences become

$$\begin{aligned} P_d &: \left\{x^{\{m,a\}}\right\} \sim \left\{x^{\{m\}}, x^{\{a\}}\right\} \succ \left\{x^{\{a\}}\right\} \succ \varnothing \succ \left\{x^{\{m\}}\right\}, \\ P_h &: \left\{x^{\{m\}}\right\} \succ \left\{x^{\{m,a\}}\right\} \sim \left\{x^{\{m\}}, x^{\{a\}}\right\} \succ \left\{x^{\{a\}}\right\} \succ \varnothing. \end{aligned}$$

There now exists a unique stable outcome, $\left\{x^{\{m,a\}}\right\}$.

In Section 3, we introduce a theory of contract language and consider how the choice of language affects the substitutability of preferences over contracts and the stability of contract outcomes. Our theory of contract language accommodates not only the setting described above, but also other natural examples such as settings with fixed costs of production (e.g., manufacturing and electricity markets). We show that market designers, when constructing the contract language for a matching market, face a trade-off between expressiveness (i.e., the number of different contractual relationships the language can describe) and stability: the more expressive the language, the less likely it is that preferences are substitutable, and the less likely it is that a stable outcome exists.

The remainder of this paper is organized as follows. In Section 2, we present our basic model and review the standard terminology and solution concepts of matching with contracts. We present our approach to contract language in Section 3, where we also discuss the relationship between language, stability, and substitutability. In Section 4, we study many-to-many matching with contracts, proving the sufficiency and necessity of substitutable preferences for the existence of stable contract outcomes. We conclude in Section 5.

Our discussion of many-to-many matching in Section 4 is essentially self-contained, so that a reader only interested in the discussion of many-to-many matching with contracts may choose to skip Section 3.

2. Model

There are finite sets D and H of *doctors* and *hospitals*; we denote the set of all *agents* by $F \equiv D \cup H$. There is a set X of *contracts* specifying relationships between doctor–hospital pairs. We elaborate upon the structure of the contract set X in Section 3, but for concreteness one may think of the special case in which X takes the form $X = D \times H \times T$, for some finite set T of contractual terms. Each contract $x \in X$ is associated with a doctor $x_D \in D$ and a hospital $x_H \in H$. For a set of contracts $Y \subseteq X$, we let $Y_D \equiv \bigcup_{y \in Y}\{y_D\}$ and $Y_H \equiv \bigcup_{y \in Y}\{y_H\}$. We let $x_F \equiv \{x_D, x_H\}$ be the set of agents associated with contract x, and let

$$Y_f \equiv \{y \in Y : f \in y_F\}$$

be the set of contracts in Y associated with agent $f \in F$.

Each $f \in F$ has a strict preference relation P_f^X over subsets of X_f. For now, we take the preferences P_f^X of the agent f as given, and, when the contract language X is clear from context, we abuse notation by suppressing the superscript and writing P_f for the preference relation of f over sets of contracts in X_f. In Section 3, we elaborate upon the preference relation structure, deriving P_f^X from a preference relation over contractual primitives. We often write $Y \succ_f Z$ to indicate that f prefers Y to Z under P_f.

For any $f \in F$ and offer set $Y \subseteq X$, we let

$$C_f(Y) \equiv \max_{P_f}\{Z \subseteq X : Z \subseteq Y_f\}$$

be the set of contracts f *chooses* from Y.[6,7,8] We let

$$R_f(Y) \equiv Y \smallsetminus C_f(Y)$$

denote the set of contracts f *rejects* from Y.

Let $C_D(Y) \equiv \bigcup_{d \in D} C_d(Y)$ be the set of contracts chosen from Y by doctors. The remaining contracts, rejected by all the doctors, comprise the *rejected set* $R_D(Y) \equiv Y \smallsetminus C_D(Y)$. Similarly, let $C_H(Y) \equiv \bigcup_{h \in H} C_h(Y)$ be the set of contracts chosen from Y by hospitals, and let $R_H(Y) \equiv Y \smallsetminus C_H(Y)$.[9]

An *outcome* is a set of contracts $Y \subseteq X$. Preference relations are naturally extended to outcomes: for two outcomes $Y, Z \subseteq X$, we say that $Y \succ_f Z$ when $Y_f \succ_f Z_f$.

2.1. Substitutability

In matching theory, the key restriction on agents' preferences is substitutability, defined directly from the choice function. Intuitively, contracts x and z are substitutes for $f \in F$ if they are not complements; that is, there are no two contracts $x, z \in X_f$ such that being offered x makes z more desirable for f.

Definition 1. The preferences of $f \in F$ are *substitutable* if, for all $x, z \in X$ and $Y \subseteq X$, if $z \notin C_f(Y \cup \{z\})$, then $z \notin C_f(\{x\} \cup Y \cup \{z\})$.

Substitutability can be rephrased in terms of the rejection function: The preferences of f are substitutable if and only if the rejection function R_f is isotone, i.e., if for any $Y' \subseteq Y \subseteq X$, we have $R_f(Y') \subseteq R_f(Y)$.

An alternative characterization of substitutability can be obtained in terms of submodularity of the indirect utility function. We say that an indirect utility function V over offer sets *represents* preference relation P_f if

$$V(Y) > V(Z) \Leftrightarrow C_f(Y) \succ_f C_f(Z) \text{ for all } Y, Z \subseteq X.$$

That is, under V, an offer set Y provides more utility to f than another offer set Z if f prefers his choice from Y to his choice from Z. In this context, an agent's preferences over contracts are substitutable if an additional offer is more valuable when the agent's original offer set is small.

Proposition 1. *The preferences of $f \in F$ are substitutable if and only if they can be represented by a submodular indirect utility function over offer sets.*

2.2. Stability

Definition 2. An outcome $A \subseteq X$ is *stable* (with respect to X) if it is

[6] Here, agents may choose any subset of the set of contracts offered. We use the term "offer set" instead of "budget set" or "set of alternatives," as agents may choose multiple contracts from an offer set, whereas agents are typically allowed to choose only one option from a budget set. (See the defintion given on the first page of Chapter 1 of Mas-Colell et al. (1995), for instance.)

[7] We use the notation $\max_{P_f}$ to indicate that the maximization is taken with respect to the preferences of agent f.

[8] We have assumed that the choice function is induced by an underlying preference relation in order to facilitate the analysis of contract language. An alternative approach treats the choice functions as primitives; under this convention, an additional *irrelevance of rejected contracts* assumption is required for key results such as those in Section 4 (see Aygün and Sönmez (2013, 2014)). In either case, the assumed structure on preferences implies a revealed preference property that is needed for the analysis (see the presentation by Aygün and Sönmez (2012)).

[9] Note that $R_D(Y) = \bigcup_{d \in D} R_d(Y_d) \neq \bigcup_{d \in D} R_d(Y)$ and similarly $R_H(Y) = \bigcup_{h \in H} R_h(Y_h) \neq \bigcup_{h \in H} R_h(Y)$.

1. *Individually rational*: for all $f \in F$, $C_f(A) = A_f$.
2. *Unblocked*: There does not exist a nonempty *blocking set* $Z \subseteq X$ such that $Z \cap A = \varnothing$ and, for all $f \in Z_F$, $Z_f \subseteq C_f(A \cup Z)$.

This concept generalizes the stability concepts of the one-to-one and many-to-one matching literatures.[10,11] In the one-to-one matching literature, the standard definition of a stable outcome A requires that A be individually rational and that there be no blocking set Z such that $|Z| = 1$; we call this *pairwise stability*. Similarly, in the many-to-one matching literature, where only hospitals are allowed to sign multiple contracts, the standard definition of a stable outcome A requires that A be individually rational and that there be no blocking set Z such that $|Z_H| = 1$. If A is individually rational and there is no blocking set Z such that $|Z_H| = 1$ or $|Z_D| = 1$, we say that A is *many-to-one stable*. It is immediate that any stable outcome is many-to-one stable and that any many-to-one stable outcome is pairwise stable. Our next result shows a partial converse: all three stability concepts are equivalent in the presence of substitutable preferences.

Proposition 2. *If all agents' preferences are substitutable, then stability, many-to-one stability, and pairwise stability are equivalent.*

To show Proposition 2 we show that if A is blocked by Z, then for any $z \in Z$, the set $\{z\}$ blocks A on its own. This fact, in turn, follows from the fact that if $z \in C_f(A \cup Z)$ for each $f \in z_F$, and if the preferences of each $f \in z_F$ are substitutable, then $z \in C_f(A \cup \{z\})$ for each $f \in z_F$.[12]

3. Contract Language

We now develop a theory of the contract set X as a language for expressing sets of underlying primitive contract terms. Throughout this section, we allow the contract set to vary, and discuss the effects of varying contract language on stability and the substitutability of preferences.

3.1. Basic Theory of Language

For each doctor–hospital pair $(d, h) \in D \times H$, there is a set of *contractual primitives* $\pi(d, h)$ that defines the set of possible contractual relationships between d and h. We write

$$\Pi_d \equiv \bigcup_{h \in H} \pi(d, h)$$

for the set of primitives associated to doctor $d \in D$ and

$$\Pi_h \equiv \bigcup_{d \in D} \pi(d, h)$$

for the set of primitives associated to hospital $h \in H$. We require that $\pi(d, h) \cap \pi(d', h') = \varnothing$ for all $(d, h) \neq (d', h')$ so that each primitive uniquely identifies a doctor and hospital. A *primitive outcome* is a collection of primitives

$$\Lambda \subseteq \bigcup_{(d,h) \in D \times H} \pi(d, h).$$

[10] In particular, this definition is equivalent to that of Hatfield and Milgrom (2005) in the context of many-to-one matching with contracts.

[11] Our stability concept allows agents associated with a blocking set to disagree as to whether a contract in the original outcome is maintained while deviating, but requires deviating agents to agree on the set of new contracts established in the deviation. Thus, our stability concept is slightly stronger than the *weak setwise stability* concept of Echenique and Oviedo (2006) and Klaus and Walzl (2009), which is nearly the same as our stability concept but requires that agents within the blocking coalition agree as to which contracts are maintained from the pre-deviation outcome. Meanwhile, our stability concept is not directly comparable to the *setwise stability* concept of Sotomayor (1999), Echenique and Oviedo (2006), and Klaus and Walzl (2009) or the *group stability* concept of Konishi and Ünver (2006) (which Klaus and Walzl (2009) call *strong setwise stability*). The *strong stability* concept we examine in Section 4.3 is stronger than setwise stability and weaker than group stability.

[12] Hatfield et al. (2016) provide a generalization of this result to arbitary trading networks.

A *contract* between d and h is a collection of primitives in $\pi(d,h)$. Denoting the power set of $\pi(d,h)$ by $\mathcal{P}(\pi(d,h))$, a contract between d and h is just a nonempty element of $\mathcal{P}(\pi(d,h))$. For example, $\pi(d,h)$ might consist of all the distinct work hours available at hospital h in a given week; a contract between h and d is a subset of $\pi(d,h)$ corresponding to the work hours assigned to d by h.

A *contract language* $X_{(d,h)}$ *for* $(d,h) \in D \times H$ is a set of contracts between d and h, i.e., a subset of $\mathcal{P}(\pi(d,h)) \smallsetminus \{\varnothing\}$. More generally, a *contract language* X is a union of contract languages for each agent pair: $X = \bigcup_{(d,h)\in D\times H} X_{(d,h)}$ with $X_{(d,h)} \subseteq \mathcal{P}(\pi(d,h)) \smallsetminus \{\varnothing\}$ for each $(d,h) \in D \times H$.[13] We say that a primitive outcome Λ is *expressible* in the contract language X if there exists some $Y \subseteq X$ such that $\Lambda = \bigcup_{y \in Y} y$. In this case we say that Y *expresses* Λ.

Each $f \in F$ has a strict preference relation P_f over the set $\mathcal{P}(\Pi_f)$ of sets of primitives involving f. For any contract language X, this preference relation over primitives induces a preference relation, denoted P_f^X, over sets of contracts in X_f (i.e., subsets of $\mathcal{P}(X)$); that is, $Y \subseteq X_f$ is preferred to $Z \subseteq X_f$ under P_f^X if and only if $\cup_{y\in Y} y$ is preferred to $\cup_{z \in Z} z$ under P_f.[14] This induced preference relation is not strict, but its only indifferences arise on sets of contracts $Y, Y' \subseteq X$ that correspond to the same primitive outcome $(\cup_{y\in Y} y = \cup_{y'\in Y'} y')$. When describing preferences in the sequel, despite indifference between these *primitive-equivalent* sets of contracts, we typically assume that the preference relation P_f^X is strict, arbitrarily breaking ties among primitive-equivalent contract sets.[15,16] The choice function associated to P_f^X is denoted by C_f^X. We say that Y is *stable with respect to the contract language* X if there is some tie-breaking rule for which Y is stable under the induced choice functions. As before, when the contract language X is clear from context, we abuse notation by suppressing the superscript and writing P_f for the preference relation of f over contracts in X, and C_f for the associated choice function.

When $\pi(d,h)$ is a singleton for each doctor–hospital pair $(d,h) \in D \times H$, and $X \cong \bigcup_{f \in F} \Pi_f$, we recover the many-to-many matching model considered by Sotomayor (1999), Echenique and Oviedo (2006), and Konishi and Ünver (2006). In this case, each primitive outcome is exactly a (many-to-many) matching between doctors and hospitals.

Although our model can recapture the familiar structure of many-to-many matching, its more general structure exhibits a key distinction from classical matching models: depending upon the structure of the contract language X, some primitive outcomes are not expressible at all, and others may *only* be expressible if doctors $d \in D$ and hospitals $h \in H$ are allowed to sign multiple contracts with each other to describe their mutual obligations. This latter feature stands in sharp contrast to the restriction adopted by Klaus and Walzl (2009) that each doctor–hospital pair sign at most a single contract. As we illustrate in Section 3.4, the ability of doctors to sign multiple contracts with the same hospital has subtle implications for the definition of substitutability.

[13]Here, a contract language specifies the set of contracts allowed by the centralized clearinghouse. There is scope for one further layer between the set of primitives and the contract language: the set of primitives that are *legally contractible*, i.e., the set of primitives that agents can contract on both inside *and outside* the mechanism. In our work, we assume that agents can only deviate to contracts within the contract language; this simplification is certainly not without loss. It would be interesting to understand how the analysis changes if deviations within the full legally contractible set are allowed; at minimum, understanding the structure of agents' preferences over the legally contractible set seems essential for contract language choice (see further discussion in Section 5 as well as Roth and Shorrer (2015)).

[14]Note that here we are explicitly assuming non-unitarity, in the sense that we allow for the possibility that an agent may desire to hold multiple contracts with a single match partner simultaneously. This is in contrast to previous work (e.g., Klaus and Walzl (2009)) where the complete relationship between a doctor–hospital pair for a given outcome was required to be codified in a single contract. Indeed, under unitarity, preferences must be constructed differently, since each agent must deem unacceptable any set of contracts that contains multiple contracts with any other individual agent.

[15]This choice is not entirely without loss of generality—it affects the set of stable outcomes. However, arbitrary tie-breaking is not problematic, as if for a given tie-breaking of indifferences over primitive-equivalent expressions of a primitive outcome Λ, the outcome $Y \subseteq X$ expresses Λ and is stable with respect to X, then for any tie-breaking there is a (possibly distinct) outcome $Y' \subseteq X$ which expresses Λ and is stable with respect to X. For simplicity, when stating induced preferences over contracts, if there are multiple contractual sets that are primitive-equivalent, we only list those contractual sets relevant for the exposition.

[16]On page 7, we discuss an example with two primitive-equivalent sets of contracts—$\{x^{\{w,\$\}}\}$ and $\{x^{\{w\}}, x^{\{\$\}}\}$, which both correspond to the set of primitives $\{w, \$\}$.

3.2. Language and Stability

If a primitive outcome Λ is expressible in the contract language X by an outcome Y that is stable with respect to X, then we say that Λ is *stable with respect to the contract language* X.[17]

It is clear that primitive outcomes may be stable with respect to some contract languages and unstable with respect to others. For example, the empty outcome is stable with respect to an empty contract language, but is generally unstable once contracts with content are allowed. We now formalize and extend the structure behind this observation.

To facilitate comparisons between languages, we introduce a partial order on contract languages.

Definition 3. A contract language X *is finer than* (or *refines*) another contract language X' if $X \supsetneq X'$. In this case, we also say that X' *is coarser than* (or *coarsens*) X and write $X \rhd X'$.[18]

Refinement of a language X' corresponds to an increase in *expressiveness*: if $X \rhd X'$, then each agent may express a richer preference relation over contracts in X than she can over contracts in X'.[19] With the ordering $\rhd$, the set of contract languages forms a lattice, with least upper bound and greatest lower bound operators respectively given by the (setwise) union and intersection operations.

We quickly observe a tradeoff between the expressiveness of a language and the stability of underlying outcomes: finer languages allow more complex preference specification, which leads to (weakly) reduced stability.

Proposition 3. *Suppose that $X \rhd X'$ and that an outcome $Y \subseteq X'$ is stable with respect to X. Then, Y is stable with respect to X'.*

Proposition 3 shows the natural result that coarsening a contract language X preserves the stability of an outcome Y, so long as Y is not eliminated from the language. However, this result applies only to outcomes, not to primitive outcomes. To see this, consider a setting with a single doctor, a single hospital, and two contractual primitives: the doctor working (w) and being compensated (\$). Formally, we write $D = \{d\}$, $H = \{h\}$, and $\pi(d, h) = \{w, \$\}$. We suppose that agents' underlying preferences take the natural form

$$P_d : \{\$\} \succ \{w, \$\} \succ \varnothing, \quad P_h : \{w\} \succ \{w, \$\} \succ \varnothing.$$

Both agents want to contract, but the doctor would most prefer to be paid for nothing, and the hospital would most prefer that the doctor work for free. As before, we denote $x^{\Pi} \equiv \Pi$ for a set of primitives Π. When all contracts are possible—$X = \{x^{\{w\}}, x^{\{\$\}}, x^{\{w,\$\}}\}$—preferences over contracts are

$$\begin{aligned} P_d^X &: \{x^{\{\$\}}\} \succ \{x^{\{w,\$\}}\} \sim \{x^{\{w\}}, x^{\{\$\}}\} \succ \varnothing, \\ P_h^X &: \{x^{\{w\}}\} \succ \{x^{\{w,\$\}}\} \sim \{x^{\{w\}}, x^{\{\$\}}\} \succ \varnothing, \end{aligned}$$

and the unique stable outcome is $\{x^{\{w,\$\}}\}$ (regardless of the tie-breaking rule used for indifferences in agents' preferences (see Footnote 15)). If we coarsen X to $X' = \{\varnothing, x^{\{w\}}, x^{\{\$\}}\}$ by removing the contract $x^{\{w,\$\}}$, agents' preferences reduce to

$$\begin{aligned} P_d^{X'} &: \{x^{\{\$\}}\} \succ \{x^{\{w\}}, x^{\{\$\}}\} \succ \varnothing, \\ P_h^{X'} &: \{x^{\{w\}}\} \succ \{x^{\{w\}}, x^{\{\$\}}\} \succ \varnothing, \end{aligned}$$

under which only $\varnothing$ is stable.[20] Thus, we see that the stability of the primitive outcome $\{w, \$\}$ is not preserved under the coarsening of X to X'.

[17] Unfortunately, although agents are indifferent over contract sets which express the same primitive outcomes, not all expressions of a primitive outcome Λ that is stable with respect to X need be stable.

[18] Of course, any (strict) subset of a contract language X coarsens X. Although we could simply denote the refinement relation by the (strict) setwise inclusion relation $\supsetneq$, we use the distinguished notation $\rhd$ to help clarify when we are actively comparing two contract languages.

[19] We need not have $P_f^X \neq P_f^{X'}$ for all $f \in F$, since X might only differ from X' by the addition of contracts disjoint from Π_f.

[20] Note that $\varnothing$ is not blocked by $\{x^{\{w\}}, x^{\{\$\}}\}$, as $C_d^{X'}(\{x^{\{w\}}, x^{\{\$\}}\}) = \{x^{\{\$\}}\} \neq \{x^{\{w\}}, x^{\{\$\}}\}$.

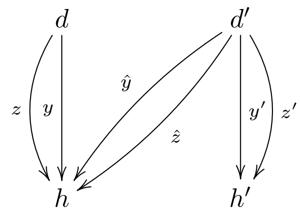

Figure 1: Diagram of contracts under language X.

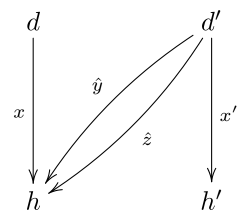

Figure 2: Diagram of contracts under language $\hat{X}$.

A natural assumption when multiple contracts are allowed between a doctor and hospital is that if a doctor chooses to abrogate one of his contracts with a particular hospital, then he must abrogate all contracts with that hospital.[21] Hence, a natural question is whether a primitive outcome Λ stably expressed by Y with respect to a contract language X must also be stable when we consider the language X' that codifies, for each doctor-hospital pair, the entire relationship described by Λ into a single contract; that is, must Λ be stable with respect to a new contract language

$$X' \equiv (X \smallsetminus Y) \cup \left\{ w \in \bigcup_{(d,h) \in D \times H} \{\cup_{z \in Y \cap X_d \cap X_h} z\} : w \neq \varnothing \right\}?$$

Unfortunately, this is not the case, as the following example demonstrates.

Suppose that $D = \{d, d'\}$ and $H = \{h, h'\}$, and let the set of contracts be given by $X = \{y, z, y', z', \hat{y}, \hat{z}\}$, with the associations of doctors and hospitals to contracts as pictured in Figure 1.[22]

Suppose that the (substitutable) preferences of the agents are given by:

$$\begin{aligned}
P_h^X &: \{y, \hat{z}\} \succ \{\hat{y}, \hat{z}\} \succ \{y, z\} \succ \{\hat{y}, z\} \succ \{\hat{z}\} \succ \{y\} \succ \{\hat{y}\} \succ \{z\} \succ \varnothing, \\
P_{h'}^X &: \{y', z'\} \succ \{y'\} \succ \{z'\} \succ \varnothing, \\
P_d^X &: \{y, z\} \succ \{y\} \succ \{z\} \succ \varnothing, \\
P_{d'}^X &: \{\hat{y}, z'\} \succ \{\hat{y}, \hat{z}\} \succ \{y', z'\} \succ \{y', \hat{z}\} \succ \{\hat{y}\} \succ \{z'\} \succ \{\hat{z}\} \succ \{y'\} \succ \varnothing.
\end{aligned}$$

In this case, the only stable outcome is $Y = \{y, z, y', z'\}$. However, if we consider the contractual language X' obtained by binding together y and z, and y' and z'—that is, replacing y and z with the contract $x = y \cup z$ and replacing y' and z' with the contract $x' = y' \cup z'$, thus resulting in the language $X' = \{x, x', \hat{y}, \hat{z}\}$—as

[21] Note that this transformation does not correspond to imposing unitarity. Indeed, imposing unitarity can not be implemented by simply changing the contractual language as unitarity requires not just that there be a single contract that represents the relationship between a doctor–hospital pair, but also that the hospital finds any set with two contracts with the same doctor unacceptable (even if each individual contract, as well as the contract combining the primitives from those two contracts, are acceptable).

[22] For expositional convenience, we present this example in terms of contracts, without explicitly stating the underlying primitives. Nevertheless, it is clear how primitives and preferences over primitives could be chosen so as generate the preferences in the example.

shown in Figure 2, the preferences of the agents now take the form

$$\begin{aligned}
P_h^{\hat{X}} &: \{\hat{y}, \hat{z}\} \succ \{x\} \succ \{\hat{z}\} \succ \{\hat{y}\} \succ \varnothing, \\
P_{h'}^{\hat{X}} &: \{x'\} \succ \varnothing, \\
P_d^{\hat{X}} &: \{x\} \succ \varnothing, \\
P_{d'}^{\hat{X}} &: \{\hat{y}, \hat{z}\} \succ \{x'\} \succ \{\hat{y}\} \succ \{\hat{z}\} \succ \varnothing.
\end{aligned}$$

For the contract language, X', the set $\{x, x'\}$ (which is primitive-equivalent to Y) is *not* stable, as $Z = \{\hat{y}, \hat{z}\}$ constitutes a blocking set.[23] For the language X, the set $Y = \{y, z, y', z'\}$ is stable since any block requires agents to choose all of the blocking contracts, and h will never choose $\hat{y}$ when y is available. However, once y and z are encapsulated into one contract x, hospital h will drop x in order to obtain both $\hat{y}$ and $\hat{z}$. Similarly, d' will never choose $\hat{z}$ when z' is available, but once y' and z' are encapsulated into one contract x', the doctor d' will drop x' in order to obtain both $\hat{y}$ and $\hat{z}$.

3.3. Language and Substitutability

Substitutability is a stringent condition, in practice, and certainly need not be true of agents' underlying preferences over primitives. Nevertheless, clever contract language design can lead to substitutable preferences.

For example, consider a setting with a single doctor, a single hospital, and two contractual primitives: working the morning shift (m) and working the afternoon shift (a). Formally, we write $D = \{d\}$, $H = \{h\}$, and $\pi(d, h) = \{m, a\}$. Suppose that the agents' underlying preferences over primitives are

$$P_d : \{m, a\} \succ \varnothing, \quad P_h : \{m, a\} \succ \varnothing.$$

Both agents want to contract over a full-time job, and neither will contract over a part-time position. If m and a are split into separate "part-time job" contracts $x^{\{m\}}$ and $x^{\{a\}}$, then the agents' preferences are not substitutable—$x^{\{m\}}$ and $x^{\{a\}}$ are complements in this language. This is true even if a single "full-time job" contract $x^{\{m,a\}}$ is available in addition to the part-time contracts. By contrast, if only the full-time contract $x^{\{m,a\}}$ is available, agents' preferences are substitutably expressed as

$$P_d^{\{x^{\{m,a\}}\}} : \{x^{\{m,a\}}\} \succ \varnothing, \quad P_h^{\{x^{\{m,a\}}\}} : \{x^{\{m,a\}}\} \succ \varnothing.$$

Every contract language X has a coarsening X' over which preferences are substitutable. Our next result shows that once such a coarsening X' is found, any further coarsening of X' will induce substitutable preferences as well.

Proposition 4. *Suppose that $X \rhd X'$ and that the preference relation P_f^X of an agent $f \in F$ is substitutable. Then, $P_f^{X'}$ is substitutable, as well.*

Just as Proposition 3 indicates a tradeoff between expressiveness and stability, Proposition 4 indicates a tradeoff between expressiveness and substitutability. Our later results (Theorems 1 and 2) show that substitutability of preferences is sufficient and necessary (in the maximal domain sense) for the existence of stable outcomes; hence, Proposition 4 implies a direct tradeoff between expressiveness and the existence of stable outcomes. However, selecting an effective language seems potentially difficult in practice, as it depends upon parameters which the market designer must assess.

3.4. Allowing Multiple Contracts Between a Doctor–Hospital Pair

It is clear that we could theoretically require that every possible relationship between a doctor and a hospital be codified into a single contract; in the language of Kominers (2012), this would correspond to imposing unitarity. A unitary contract structure is in fact required by Klaus and Walzl (2009). However,

[23] In fact, we can go further, binding $\hat{y}$ and $\hat{z}$ into a single contract $\hat{x}$, resulting in the language $X'' = \{\hat{x}, x', x\}$, which corresponds to a unitary model; again, in that setting, the outcome $\{x, x'\}$ is no longer stable, as it is blocked by $\{\hat{x}\}$.

imposing unitarity may obscure substitutable structure within agents' preferences, as the following example shows.[24]

Consider a hospital h with two positions, a medical position (μ) and a surgical position (σ). We suppose that

- doctor d can take either or both of the positions μ and σ; and
- doctor d' can only take the surgical position σ.

The hospital would like to assign a doctor to the medical position μ, but would prefer that d', rather than d, take the surgical position σ. These are natural preferences, and intuitively they should be substitutable, as d' "substitutes" for d in the surgical position σ. But if we allow at most one contract per doctor–hospital pair, then the possible assignments of doctor d take the form of three possible contracts between d and h:

- $x^{\{(d,\mu)\}}$, where d takes only the the medical position μ,
- $x^{\{(d,\sigma)\}}$, where d takes only the the surgical position σ, and
- $x^{\{(d,\mu),(d,\sigma)\}}$, where d takes both positions.[25]

Using similar notation, we let $x^{\{(d',\sigma)\}}$ be the contract between d' and h that specifies that d' takes the surgical position σ. Hence, the set of contracts is given by

$$X = \left\{x^{\{(d,\mu)\}}, x^{\{(d,\sigma)\}}, x^{\{(d,\mu),(d,\sigma)\}}, x^{\{(d',\sigma)\}}\right\}.$$

Under the assumption that a hospital can sign at most one contract with a given doctor, as in the framework of Klaus and Walzl (2009),[26] the preferences of h would take the form

$$\left\{x^{\{(d,\mu)\}}, x^{\{(d',\sigma)\}}\right\} \succ \left\{x^{\{(d,\mu),(d,\sigma)\}}\right\} \succ \left\{x^{\{(d,\mu)\}}\right\} \succ \left\{x^{\{(d',\sigma)\}}\right\} \succ \left\{x^{\{(d,\sigma)\}}\right\} \succ \varnothing,$$

which are not substitutable, as

$$R_h\left(\left\{x^{\{(d,\mu),(d,\sigma)\}}, x^{\{(d,\mu)\}}\right\}\right) = \left\{x^{\{(d,\mu)\}}\right\} \not\subseteq \left\{x^{\{(d,\mu),(d,\sigma)\}}\right\} = R_h\left(\left\{x^{\{(d,\mu),(d,\sigma)\}}, x^{\{(d,\mu)\}}, x^{\{(d',\sigma)\}}\right\}\right).$$

However, in our model, where we allow multiple contracts between agent pairs, if we work in the coarser contract language $X' = X \smallsetminus \{x^{\{(d,\mu),(d,\sigma)\}}\}$, the preferences of h can be written in the substitutable form

$$P_h^{X'}: \left\{x^{\{(d,\mu)\}}, x^{\{(d',\sigma)\}}\right\} \succ \left\{x^{\{(d,\mu)\}}, x^{\{(d,\sigma)\}}\right\} \succ \left\{x^{\{(d,\mu)\}}\right\} \succ \left\{x^{\{(d',\sigma)\}}\right\} \succ \left\{x^{\{(d,\sigma)\}}\right\} \succ \varnothing.$$

This rewritten preference relation makes clear the intuitive fact that d' substitutes for d in the surgical position σ. Without the presence of multiple contracts between the doctor–hospital pair $(d,h) \in D \times H$, this intuition is obscured, as is the fact (implied by our existence result, Theorem 1) that stable outcomes exist under P_h so long as the preferences of d and d' are substitutable.

In fact, the structure of the example parallels that of the United Kingdom Medical Intern match, in which students must take on both medical and surgical positions in order to become eligible for full registration with the General Medical Council. Hence, this example shows that allowing multiple contracts between a doctor–hospital pair can uncover the substitutable structure of preferences in a real-world setting.

In our subsequent discussion, we assume the possibility of multiple contracts between doctor–hospital pairs.[27] As the example just presented suggests, the class of substitutable preferences in our framework

---

[24] In this example, to highlight the impact of unitarity, we are changing both the contractual language and the way preferences are constructed, as specified in Footnote 14.

[25] Here, we add the contract $x^{\{(d,\mu),(d,\sigma)\}}$ so that it is still possible the agents can obtain the primitive outcome $\{\mu,\sigma\}$ even after unitarity has been imposed.

[26] Note that the preferences do not take the form used in our framework, as the set $\left\{x^{\{(d,\mu),(d,\sigma)\}}\right\}$ is primitive-equivalent to $\left\{x^{\{(d,\mu)}, x^{(d,\sigma)\}}\right\}$, and so in our model the hospital h should be indifferent between these two sets of contracts.

[27] This is a substantive assumption on the contract set X, but a very weak one.

therefore includes many sets of preferences that are naturally substitutable but are not considered substitutable in unitary many-to-many matching with contracts models.

4. Many-to-Many Matching with Contracts

We show in this section that substitutability is crucial for the existence of stable outcomes: it is both sufficient and necessary (in the maximal domain sense).

To show existence of stable outcomes under substitutable preferences, we follow an approach similar to that of Hatfield and Milgrom (2005): We construct a generalized deferred acceptance operator Φ; we show that fixed points of Φ correspond to stable outcomes; and finally, we use Tarski's fixed point theorem to show the existence of a nonempty lattice of fixed points. However, we introduce here a new generalized deferred acceptance operator which ensures that the correspondence between fixed points and stable outcomes is one-to-one—unlike under the operators of Hatfield and Milgrom (2005), Ostrovsky (2008) and Hatfield and Kominers (2012).[28] We let

$$\begin{aligned}\Phi(X^D, X^H) &\equiv (\Phi_D(X^H), \Phi_H(X^D)), \text{ where}\\ \Phi_D(X^H) &\equiv \{x \in X : x \in C_H(X^H \cup \{x\})\} \text{ and}\\ \Phi_H(X^D) &\equiv \{x \in X : x \in C_D(X^D \cup \{x\})\}.\end{aligned}$$

Under this operator, the sets X^D and X^H represent the sets of contracts "available" to the doctors and hospitals, respectively. After an iteration of the operator Φ, the offer set $\Phi_D(X^H)$ made available to the doctors is the set of contracts that the hospitals would be willing to take given offer set X^H. Analogously, the offer set $\Phi_H(X^D)$ made available to the hospitals is the set of contracts that the doctors would be willing to take given their current offer set X^D.

Now suppose the preferences of all agents are substitutable. If (X^D, X^H) is a fixed point of Φ, then each $x \in X^D \cap X^H \equiv A$ is chosen by x_D from X^D; since the preferences of x_D are substitutable, x_D must then also choose x from $A \subseteq X^D$. Analogously, each $x \in A$ is chosen by x_H from the set X^H; since the preferences of x_H are substitutable, x_H must then also choose x from $A \subseteq X^H$. Hence, A is individually rational. Moreover, if A were blocked, then by Proposition 2 there would be a blocking set of the form $\{z\}$. As z would be chosen by z_D from $A \cup \{z\}$, the contract z would also be chosen from $X^D \cup \{z\}$. Analogous reasoning shows that $z \in \Phi_D(X^H) = X^D$. Hence, we would have $z \in A = X^D \cap X^H$—so $\{z\}$ could not be a blocking set.

Lemma 1. *For any fixed point (X^D, X^H) of Φ, the outcome $X^D \cap X^H$ is a stable outcome. Conversely, for any stable outcome A, there exists a unique fixed point (X^D, X^H) of Φ such that $X^D \cap X^H = A$; moreover, $(X^D, X^H) = \Phi(A, A)$.*

Keeping track of the offer sets X^D and X^H also allows us to determine the "desirability" of a given contract at a fixed point (X^D, X^H): If $x \in X^D \cap X^H = A$, then x is part of the stable outcome. If $x \in X^D \smallsetminus X^H$, then x is desired by x_H but not by x_D. If $x \in X^H \smallsetminus X^D$, then x is desired by x_D but not by x_H. Finally if $x \in X \smallsetminus (X^D \cup X^H)$, then x is desired by neither x_D nor x_H.

When all agents' preferences are substitutable, the operator Φ is *isotone* in the sense that if $X^D \subseteq \tilde{X}^D$ and $X^H \supseteq \tilde{X}^H$, then $\Phi_D(X^H) \subseteq \Phi_D(\tilde{X}^H)$ and $\Phi_H(X^D) \supseteq \Phi_H(\tilde{X}^D)$. Hence, by Tarski's fixed-point theorem, there exists a nonempty lattice of fixed points of Φ. Moreover, this lattice corresponds to a lattice of stable outcomes with the ordering $\succeq_D$, where $A \succeq_D \bar{A}$ if and only if $A \succeq_d \bar{A}$ for all $d \in D$.

Theorem 1. *If all agents' preferences are substitutable, then there exists at least one stable outcome; moreover, the set of stable outcomes forms a lattice with respect to the operator $\succeq_D$.*

The lattice structure identified in Theorem 1 also leads to the standard "opposition of interests" result, that is, for any stable outcomes A and $\bar{A}$, if A is preferred by all the doctors to $\bar{A}$ (i.e., $A \succeq_D \bar{A}$), then all the

[28] The specific operator we use here is inspired by an operator introduced by Azevedo and Hatfield (2015) for a continuum matching setting.

hospitals prefer $\bar{A}$ to A. In particular, doctor-optimal and doctor-pessimal stable outcomes exist, and they are the hospital-pessimal and hospital-optimal stable outcomes, respectively.[29]

In the model of many-to-one matching with contracts, conditions on preferences weaker than substitutability can be found that guarantee the existence of stable outcomes (see Hatfield and Kojima (2010), Hatfield and Kominers (2016), and Hatfield et al. (2015)).[30] Our next result shows that these results for weakened substitutability conditions do *not* carry over to the many-to-many matching with contracts model. In particular, we show that if there are at least two agents of each type and some agent's preferences are not substitutable, then substitutable preferences for the other agents can be constructed such that no stable outcome exists.

Theorem 2. *If the preferences of some agent $f \in F$ are not substitutable, there are at least two other agents of each type, and X contains at least one contract between every doctor–hospital pair, then there exist substitutable preferences for the doctors and hospitals in $F \smallsetminus \{f\}$ such that no (many-to-one) stable outcome exists.*[31]

If the preferences of a hospital h are not substitutable, then there exist contracts $x, z \in X$ and a set of contracts $Y \subseteq X$ such that $z \notin C_h(Y \cup \{z\})$ but $z \in C_h(\{x\} \cup Y \cup \{z\})$. The proof of Theorem 2 proceeds in two cases, depending on whether $x_D \neq z_D$ or $x_D = z_D$. In the first case, we let x_D have a contract x' with a hospital $h' \neq h$, and let z_D have a contract z' with that same hospital h'; we further let x_D prefer $\{x'\}$ to $\{x\}$ (and find $\{x\}$ acceptable but $\{x, x'\}$ unacceptable), while letting z_D prefer $\{z\}$ to $\{z'\}$ (and find $\{z'\}$ acceptable but $\{z, z'\}$ unacceptable). Hospital h' has unit demand, and prefers $\{z'\}$ to $\{x'\}$. Finally, we let all the doctors demand all their contracts in Y, regardless of their other opportunities. Suppose that A is a stable outcome: Then z_D obtains z or z', as h' will always take z' if it is available. If z_D obtains z', then $x' \notin A$, and so $C^h(\{x, z\} \cup Y)$ is a blocking set. But if z_D obtains z, then $x \in A$, and so $\{x'\}$ is a blocking set. Intuitively, hospital h' prefers z', so that whenever z_D accepts the contract z' with h', he blocks x_D from working at h'; then, x_D consents to work for h, who in turn now wishes to take on contract z. However, this opens up the position at h', and now x_D no longer wishes to work at h but instead at h'. The logic for the second case is similar although the technical details differ.

The viability of outcomes with multiple contracts between the same doctor–hospital pair is crucial to the proof of Theorem 2, as the proof requires that doctors in Y_D (other than x_D and z_D) be willing to accept any and all contracts offered to them. Since in principle X can contain multiple contracts with each doctor, the doctors in Y_D must in general be willing to accept multiple contracts. Thus, the distinction of our model from that of Klaus and Walzl (2009)—that a doctor may sign multiple contracts with a given hospital—is directly relevant. In fact, as Klaus and Walzl (2009) and Yenmez (2014) have demonstrated, conditions weaker than substitutability are sufficient to ensure existence of equilibria in unitary matching models.[32]

Weakening the solution concept beyond many-to-one stability may assuage the difficulty presented in Theorem 2, but may be otherwise unsatisfactory. For example, it is well-understood that pairwise stability is an inappropriate solution concept in many-to-many matching with contracts, as there are many pairwise stable outcomes that we would not expect to be stable in practice (see, e.g., Echenique and Oviedo (2006) and Hatfield and Kominers (2012)).[33]

[29] Analogous opposition of interests results have been identified in most matching settings, including those of Roth (1984b), Blair (1988), Hatfield and Milgrom (2005), and Echenique and Oviedo (2006).

[30] This exception holds only in models with contracts. In particular, if there is a unique contract between each doctor–hospital pair, then substitutability is required (in the maximal domain sense) for the existence of stable outcomes (Hatfield and Kojima (2008)). Moreover, Hatfield and Kojima (2008) show that in the Kelso and Crawford (1982) model of many-to-one matching with wages substitutability is required (in the maximal domain sense) for the existence of stable outcomes, implying that substitutability is also required for the existence of stable outcomes in models of many-to-many matching with wages. Substitutability is also required for the existence of stable outcomes in the more general setting of trading networks with transferable utility (see Hatfield et al. (2013)).

[31] Our maximal domain result is stronger than the specialization of the analogous result of Hatfield and Kominers (2012) to our setting.

[32] In particular, there is no direct analogue of Theorem 2 for unitary matching models. How imposing unitarity affects the set of stable primitive outcomes is unclear.

[33] For a simple example, consider the following preferences:

$$P_h : \{x, z\} \succ \varnothing, \qquad P_{x_D} : \{x'\} \succ \{x\} \succ \varnothing,$$
$$P_{h'} : \{z'\} \succ \{x'\} \succ \varnothing, \qquad P_{z_D} : \{z\} \succ \{z'\} \succ \varnothing;$$

4.1. Couples Matching

There is a great deal of interest in the question of when stable matches are guaranteed to exist in the presence of couples (see, e.g., Klaus et al. (2007), Klaus and Klijn (2007), Kojima (2015)). The answer to this question is of practical importance for real world applications such as the NRMP (Roth and Peranson (1999)). Many previous studies of matching with couples have, for simplicity, assumed that the hospitals have singleton preferences while couples may desire two positions. However, for applications such as the NRMP, hospitals typically desire to fill multiple positions.

Theorem 2 shows that the previous literature understates the difficulty of finding stable couples matchings, as if hospitals are given more realistic, substitutable preferences, the class of substitutable preferences is the most general class of preferences for couples of doctors under which a stable match is guaranteed to exist. Furthermore, substitutability is an extremely restrictive (and unrealistic) condition on the preferences of couples: it requires that the members of the couple do not find any two jobs "complementary", e.g., if the wife receives a job offer in New York, it does not make jobs in New York more desirable for the husband. Hence, substitutability effectively requires that the couple behaves as two separate doctors, with each member of the couple taking the best option available to him or her regardless of the set of positions available to the other member of the couple.[34]

4.2. The Structure of the Set of Stable Outcomes

The preferences of f satisfy the *law of aggregate demand* if for all $X'' \subseteq X' \subseteq X$, $|C_f(X'')| \leq |C_f(X')|$.[35] Under this condition, we obtain an analogue of the rural hospitals theorem of Roth (1984a): each agent signs the same number of contracts at every stable outcome (see Appendix B).[36] However, depending on how contractual primitives are assembled into contracts (as discussed in Section 3), the implications of this result can be unclear. Consider the example in which $D = \{d\}$, $H = \{h\}$ and contracts denote work shifts of different lengths: $X = \{x^{\{20\}}, x^{\{40\}}\}$ where $x^{\{t\}}$ encodes a t-hour work shift for doctor d at hospital h. In this case, even if the total number of contracts signed by h is invariant across stable outcomes, the total number of hours worked at h may nevertheless change.[37] In the terminology of contract language, this problem occurs because the contracts represent different numbers of primitive work-units (in this case, twenty-hour shifts). When all contracts are denoted in a fixed unit, however, the rural hospitals result has the natural interpretation that each agent receives the same amount of work at every stable outcome.

The law of aggregate demand is also the key condition for two other additional results in the many-to-one matching literature: one-sided (group) strategy-proofness and weak Pareto optimality (Hatfield and Milgrom (2005); Kojima (2007); Hatfield and Kojima (2010)). The standard one-sided (group) strategy-proofness result states that when doctors have unit demand, the mechanism that chooses the doctor-optimal stable outcome is strategy-proof for the doctors. The standard weak Pareto optimality result for doctors states that, again when doctors have unit demand, there does not exist an individually rational matching that all doctors strictly prefer to the doctor-optimal stable match. Unfortunately, these results do not carry over to the context of many-to-many matching—even without contracts. Indeed, the proof of Theorem 5.10 of Roth and Sotomayor (1990) provides an example in which the unique stable outcome is not weakly Pareto optimal for the hospitals, and the proof of Theorem 5.14 of Roth and Sotomayor (1990) provides an example

for these preferences, $\{z'\}$ is a pairwise stable outcome, as any block involving h must include contracts with both x_D and z_D. Nevertheless, we would not expect such an outcome to be stable in practice, as a deviation to $\{x, z\}$ seems quite likely. And indeed, there are no (many-to-one) stable outcomes for the preferences just described.

[34]Our results expand upon an earlier insight of Cantala (2004), who showed that when couples' preferences have a specific form of non-substitutability—in particular, a preference for colocation—stable matchings do not exist in general.

[35]This condition was introduced by Hatfield and Milgrom (2005). Alkan and Gale (2003) introduced a related condition called "size monotonocity."

[36]Our proof of this rural hospitals theorem requires just lattice structure and the law of aggregate demand.

[37]To see this, suppose that the agents' preferences are given by

$$P_h : \{x^{\{20\}}\} \succ \{x^{\{40\}}\} \succ \varnothing,$$
$$P_d : \{x^{\{40\}}\} \succ \{x^{\{20\}}\} \succ \varnothing.$$

Then $\{x^{\{20\}}\}$ and $\{x^{\{40\}}\}$ are both stable but correspond to distinct numbers of total work-hours.

in which one hospital has an incentive to misstate its preferences to a mechanism choosing the hospital-optimal stable outcome.[38]

The stable outcome correspondence in our context *is* Nash implementable whenever it is nonempty and there are at least three agents (see Appendix C). Informally, this means that all stable outcomes can be achieved non-cooperatively, through strategic interactions in equilibrium.[39]

4.3. Strong Substitutability and Strong Stability

Echenique and Oviedo (2006) introduced a condition, strong substitutability, that is more restrictive than substitutability; Klaus and Walzl (2009) extended this condition to the setting of many-to-many matching with contracts.

Definition 4. The preferences of $f \in F$ are *strongly substitutable* if for all $X'', X' \subseteq X$ such that $C_f(X'') \succ_f C_f(X')$, we have $(X' \cap C_f(X'')) \subseteq C_f(X')$.

Intuitively, strong substitutability means that if an agent f chooses contract x from a set of contracts X'', and if $x \in X'$ and X'' is a "better" offer set for f than X' is, then f still chooses x from X'.

There is also a more restrictive stability concept for many-to-many matching problems.

Definition 5. An outcome A is *strongly stable* if it is

1. Individually rational and
2. *Strongly unblocked*: There does not exist a nonempty set $Z \subseteq X$ such that $Z \cap A = \varnothing$ and, for all $f \in Z_F$, there exists an individually rational Y^f such that $Z_f \subseteq Y^f \subseteq Z \cup A$ and $Y^f \succ_f A$.

The key difference between stability and strong stability is that strong stability does not require deviations to be self-enforcing—they need only be individually rational. Strong stability is stronger than the *setwise stability* condition of Sotomayor (1999), Echenique and Oviedo (2006), and Klaus and Walzl (2009), a similar concept that imposes the additional requirement that the deviating agents agree on which contracts in the original outcome A to drop (i.e., for all $y \in A$, $y \in Y^{y_D}$ if and only if $y \in Y^{y_H}$). Strong stability is weaker than the *group stability* concept of Konishi and Ünver (2006), which does not require that the deviation sets Y^f be individually rational.

For many-to-one (and one-to-one) matching, an outcome A is stable if and only if it is strongly stable, and both of these conditions are equivalent to A being in the core. However, this is no longer true in the many-to-many matching context.[40]

We now show that if preferences on one side of the market are strongly substitutable, and if those on the other side are substitutable, then any stable outcome is strongly stable. This result generalizes the analogous results of Echenique and Oviedo (2006) and Klaus and Walzl (2009).

Theorem 3. *If all agents' preferences are substitutable, and if furthermore the preferences of all agents of one type (doctors or hospitals) are strongly substitutable, then an outcome is stable if and only if it is strongly stable.*

Theorem 3 implies that if all agents' preferences are substitutable, and if the preferences of all agents of one type are strongly substitutable, then

- strongly stable outcomes exist, and

[38] However, when all agents' preferences satisfy the law of aggregate demand, the hospital-optimal stable mechanism is (group) strategy-proof and the hospital-optimal stable outcome is weakly Pareto optimal *for the set of hospitals that have unit demand* (Hatfield and Kominers (2012)).

[39] This extends the analogous results of Kara and Sönmez (1996, 1997) and Haake and Klaus (2009a,b) for less-general matching settings. The requirement of three agents is as sharp as possible, since Kara and Sönmez (1996) have already proven that the stable matching correspondence is not Nash implementable in the setting of one-to-one matching when there are fewer than three agents.

[40] Blair (1988) provides an example of a match that is stable but not strongly stable according to our definitions. This example is in a many-to-many matching (without contracts) context, and hence allows only one possible relationship between each pair of agents; therefore, the distinction between stability and strong stability does not hinge on the availability of multiple contracts between pairs of agents.

- the strongly stable outcome correspondence is Nash implementable.

Unfortunately, in contrast to our results for substitutable preferences and stable outcomes, strongly substitutable preferences are not necessary (in the maximal domain sense) for the existence of strongly stable outcomes. To see this, consider the setting where $D = \{i, j, k\}$, $H = \{h\}$, and $X = \{x, y, z\}$ where $x_D = i$, $y_D = j$, $z_D = k$, and $x_H = y_H = z_H = h$. Let the preferences of hospital h be given by

$$P_h : \{x, y\} \succ \{x, z\} \succ \{x\} \succ \{y\} \succ \{z\}.$$

It is not possible to give substitutable preferences for the doctors and strongly substitutable preferences for hospitals other than h such that no strongly stable match exists.

However, without strongly substitutable preferences, the existence of strongly stable outcomes is not guaranteed. Consider the setting where $D = \{i, j\}$, $H = \{h, h'\}$, and $X = \{x, \hat{x}, x', y\}$ where $x_D = \hat{x}_D = x'_D = i$, $y_D = j$, $x_H = \hat{x}_H = y_H = h$, and $x'_H = h'$. Let the preferences of the agents be given by

$$\begin{aligned} P_h &: \{\hat{x}, y\} \succ \{\hat{x}, x\} \succ \{\hat{x}\} \succ \{y\} \succ \{x\}, \\ P_{h'} &: \{x'\}, \\ P_i &: \{x', x\} \succ \{\hat{x}, x\} \succ \{x\} \succ \{x'\} \succ \{\hat{x}\}, \\ P_j &: \{y\}. \end{aligned}$$

Here, only the preferences of h and i are not strongly substitutable, but the only stable outcome—$\{x', y\}$—is not strongly unblocked.[41]

5. Conclusion

Many-to-many matching with contracts is a general framework that can be used to describe buyer–seller markets with heterogeneous goods, labor market equilibria between firms and workers, the allocation of consulting work between firms and consultants, and a variety of other important economic settings. In our framework, substitutable preferences are sufficient and necessary (in the maximal domain sense) for the existence of stable outcomes.

In related work (Hatfield and Kominers, 2016), we apply the results obtained here in the context of *many-to-one matching with contracts*: We identify a class of preferences that are not substitutable in the context of many-to-one-matching with contracts, but are projections of substitutable many-to-many matching with contracts preferences. Hence, the present results for many-to-many matching with contracts imply the existence of a new weakened substitutability condition sufficient to guarantee the existence of stable outcomes in the context of many-to-one matching with contracts; this is relevant to a broad array of applications including military cadet–branch matching (Sönmez and Switzer (2013); Sönmez (2013)), the design of affirmative action mechanisms (Kominers and Sönmez (2013, forthcoming)), and teacher allocation ((Hatfield and Kominers, 2016)).

Our results imply that careful selection of the contract language is essential for functioning matching markets. Contract design can determine which—and even more importantly, if—stable relationships can be found. Moreover, when the language is chosen effectively, many key results of matching theory apply.

Throughout, we have assumed that the market designer has complete control of the scope of possible contract language, but no power to prevent "blocks" that arise when parties deviate by recontracting within the provided language. In this setting, the stable outcomes essential for applications of matching[42] are obtained only up to blocking deviations using contracts within the available language. This approach is admittedly limited, as in practice agents who circumvent centralized clearinghouses contract outside of (and typically before) the matching mechanism. Clearly, there is no *a priori* reason why those agents should deal within the match's contract language—they could contract over any primitives that are legally contractible. Nevertheless, we believe that a centralized matching mechanism is likely to see continued participation

[41] To see this, take $Z = \{\hat{x}, x\}$ in the definition of strong unblockedness.

[42] Roth (1984a, 1991), and Roth and Xing (1994) provided empirical evidence that the stability of the outcome recommended by a centralized match is essential to the long-run success of the matching system.

if its contractual language is both expressive and guarantees the existence of stable outcomes (see also Roth and Shorrer (2015)).

If market participants have substitutable preferences over the set of primitives that are legally contractible, then the clearinghouse is well-advised to admit all feasible contracts, as using a coarser language may not sacrifice stability but could reduce efficiency. However, when both participants' preferences over primitives *and* participants' preferences over legally contractible primitives are nonsubstitutable, there is scope for nontrivial contract language design within the market clearinghouse; in such cases, optimal selection of the contract language depends upon application-specific parameters which the market designer must assess.[43,44] Hence, our work leaves substantial room for market design.

Our work also suggests a number of avenues for future research: For the problem of matching couples to hospitals with multiple positions, we now know that a stable match is only theoretically guaranteed if both hospitals' and couples' preferences are substitutable. However, although one would not expect couples' preferences to be substitutable for practical applications such as the NRMP, stable couples matches appear to exist in practice (see Roth (2008)). Since it is now clear that substitutability is a necessary condition for stability, the infrequency of instabilities in the NRMP is puzzling.[45] Of course, these issues are magnified when more complicated complementarities in preferences are present, as in the case of combinatorial package auctions (see, e.g., Ausubel and Milgrom (2002), Milgrom (2004), Kwasnica et al. (2005), and Brunner et al. (2010)). Finally, although we have identified and examined tradeoffs in the design of contract languages, it is neither clear when languages induce substitutable preferences (and hence induce stability), nor how a putative language should be judged in practice. We leave these questions for future research.

[43]Additionally, the choice of contract language may matter even when participants' preferences over legally contractible primitives are substitutable, as some contract languages may not uncover that substitutable structure (see Section 3.4 and Hatfield and Kominers (2016)).

[44]We thank a referee for pointing out these observations.

[45]Recent work by Kojima et al. (2013) and Ashlagi et al. (2014) has argued that large-market effects may explain this phenomenon.

References

Alkan, A., Gale, D., 2003. Stable schedule matching under revealed preference. Journal of Economic Theory 112, 289–306.

Ashlagi, I., Braverman, M., Hassidim, A., 2014. Matching with couples revisited. Operations Research 62, 713–772.

Ausubel, L.M., Milgrom, P., 2002. Ascending auctions with package bidding. Frontiers of Theoretical Economics 1, 1–42.

Aygün, O., Sönmez, T., 2012. Matching with contracts: The critical role of irrelevance of rejected contracts. Boston College Working Paper.

Aygün, O., Sönmez, T., 2013. Matching with contracts: Comment. American Economic Review 103, 2050–2051.

Aygün, O., Sönmez, T., 2014. The importance of irrelevance of rejected contracts in matching under weakened substitutes conditions. Boston College Working Paper.

Azevedo, E.M., Hatfield, J.W., 2015. Existence of equilibrium in large matching markets with complementarities.

Blair, C., 1988. The lattice structure of the set of stable matchings with multiple partners. Mathematics of Operations Research 13, 619–628.

Brunner, C., Goeree, J.K., Holt, C.A., Ledyard, J.O., 2010. An experimental test of flexible combinatorial spectrum auction formats. American Economic Journal: Microeconomics 2, 39–57.

Cantala, D., 2004. Matching markets: The particular case of couples. Economics Bulletin 3, 1–11.

Chambers, C.P., Echenique, F., 2009. Supermodularity and preferences. Journal of Economic Theory 144, 1004–1014.

Echenique, F., 2012. Contracts vs. salaries in matching. American Economic Review 102, 594–601.

Echenique, F., Oviedo, J., 2006. A theory of stability in many-to-many matching markets. Theoretical Economics 1, 233–273.

Haake, C.J., Klaus, B., 2009a. Monotonicity and Nash implementation in matching markets with contracts. Economic Theory 41, 393–410.

Haake, C.J., Klaus, B., 2009b. Stability and Nash implementation in matching markets with couples. Theory and Decision 69, 537–554.

Hatfield, J.W., Kojima, F., 2008. Matching with contracts: Comment. American Economic Review 98, 1189–1194.

Hatfield, J.W., Kojima, F., 2010. Substitutes and stability for matching with contracts. Journal of Economic Theory 145, 1704–1723.

Hatfield, J.W., Kominers, S.D., 2012. Matching in networks with bilateral contracts. American Economic Journal: Microeconomics 4, 176–208.

Hatfield, J.W., Kominers, S.D., 2016. Hidden substitutes. Mimeo, Harvard University.

Hatfield, J.W., Kominers, S.D., Nichifor, A., Ostrovsky, M., Westkamp, A., 2013. Stability and competitive equilibrium in trading networks. Journal of Political Economy 121, 966–1005.

Hatfield, J.W., Kominers, S.D., Nichifor, A., Ostrovsky, M., Westkamp, A., 2016. Chain stability in trading networks. Mimeo, Stanford University.

Hatfield, J.W., Kominers, S.D., Westkamp, A., 2015. Stability, strategy-proofness, and cumulative offer mechanisms. Mimeo, Harvard University.

Hatfield, J.W., Milgrom, P., 2005. Matching with contracts. American Economic Review 95, 913–935.

Jaume, D., Massó, J., Neme, A., 2012. The multiple-partners assignment game with heterogeneous sells and multi-unit demands: Competitive equilibria. Mathematical Methods of Operations Research 76, 161–187.

Kara, T., Sönmez, T., 1996. Nash implementation of matching rules. Journal of Economic Theory 68, 425–439.

Kara, T., Sönmez, T., 1997. Implementation of college admission rules. Economic Theory 9, 197–218.

Kelso, A.S., Crawford, V.P., 1982. Job matching, coalition formation, and gross substitutes. Econometrica 50, 1483–1504.

Klaus, B., Klijn, F., 2005. Stable matchings and preferences of couples. Journal of Economic Theory 121, 75–106.

Klaus, B., Klijn, F., 2007. Paths to stability for matching markets with couples. Games and Economic Behavior 58, 154–171.

Klaus, B., Klijn, F., Massó, J., 2007. Some things couples always wanted to know about stable matchings (but were afraid to ask). Review of Economic Design 11, 175–184.

Klaus, B., Walzl, M., 2009. Stable many-to-many matchings with contracts. Journal of Mathematical Economics 45, 422–434.

Kojima, F., 2007. The law of aggregate demand and welfare in the two-sided matching market. Economics Letters 99, 581–584.

Kojima, F., 2015. Finding all stable matchings with couples. Journal of Dynamics and Games 2, 321–330.

Kojima, F., Pathak, P.A., Roth, A.E., 2013. Matching with couples: Stability and incentives in large markets. Quarterly Journal of Economics 128, 1585–1632.

Kominers, S.D., 2012. On the correspondence of contracts to salaries in (many-to-many) matching. Games and Economic Behavior 75, 984–989.

Kominers, S.D., Sönmez, T., 2013. Designing for diversity in matching. Mimeo, Boston College.

Kominers, S.D., Sönmez, T., forthcoming. Matching with slot-specific priorities: Theory. Theoretical Economics .

Konishi, H., Ünver, M.U., 2006. Credible group-stability in many-to-many matching problems. Journal of Economic Theory 129, 57–80.

Kwasnica, A.M., Ledyard, J.O., Porter, D., DeMartini, C., 2005. A new and improved design for multiobject iterative auctions. Management Science 51, 419–434.

Lee, G.M., Lee, J., Whinston, A.B., 2014. Matching mobile applications for cross-promotion. Mimeo, University of Texas at Austin.

Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press.

Maskin, E.S., 1999. Nash equilibrium and welfare optimality. Review of Economic Studies 66, 23–38.

Milgrom, P.R., 2004. Putting Auction Theory to Work. Cambridge University Press.

Ostrovsky, M., 2008. Stability in supply chain networks. American Economic Review 98, 897–923.

Roth, A.E., 1984a. The evolution of the labor market for medical interns and residents: A case study in game theory. Journal of Political Economy 92, 991–1016.

Roth, A.E., 1984b. Stability and polarization of interests in job matching. Econometrica 52, 47–57.

Roth, A.E., 1985. Conflict and coincidence of interest in job matching: some new results and open questions. Mathematics of Operations Research 10, 379–389.

Roth, A.E., 1991. A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the united kingdom. American Economic Review 81, 415–440.

Roth, A.E., 2008. Deferred acceptance algorithms: History, theory, practice, and open questions. International Journal of Game Theory 36, 537–569.

Roth, A.E., Peranson, E., 1999. The redesign of the matching market for american physicians: Some engineering aspects of economic design. American Economic Review 89, 748–780.

Roth, A.E., Sotomayor, M.A.O., 1990. Two-sided matching: A study in game-theoretic modeling and analysis. Cambridge University Press.

Roth, A.E., Xing, X., 1994. Jumping the gun: Imperfections and institutions related to the timing of market transactions. American Economic Review 84, 992–1044.

Roth, B.N., Shorrer, R.I., 2015. Mechanism design in the presence of a pre-existing game. Mimeo, Harvard University.

Schlegel, J.C., 2015. Contracts versus salaries in matching: A general result. Journal of Economic Theory 159, 552–573.

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

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

Sotomayor, M.A.O., 1999. Three remarks on the many-to-many stable matching problem. Mathematical Social Sciences 38, 55–70.

Sotomayor, M.A.O., 2004. Implementation in the many-to-many matching market. Games and Economic Behavior 46, 199–212.

Yenmez, M.B., 2014. College admissions. GSIA Working Paper #2014-E24.

A. Proofs Omitted from the Main Text

Proof of Proposition 1

Suppose that the preferences of f are not substitutable. Then there exist contracts $x, z \in X$ and $Y \subseteq X$ such that

$$z \notin C_f(Y \cup \{z\}) \text{ and } z \in C_f(\{x\} \cup Y \cup \{z\}).$$

Now consider any indirect utility function V which represents these preferences. Clearly, $V(Y) = V(Y \cup \{z\})$ and $V(\{x\} \cup Y \cup \{z\}) > V(\{x\} \cup Y)$, and so

$$V(Y \cup \{z\}) - V(Y) = 0 < V(\{x\} \cup Y \cup \{z\}) - V(\{x\} \cup Y);$$

hence, V is not submodular.

Suppose that the preferences of f are substitutable. Suppose there are N sets of contracts that are individually rational for f, and that the preferences of f are given by

$$P_f : Z^N \succ Z^{N-1} \succ \cdots \succ Z^2 \succ Z^1 \succ \varnothing.$$

Let $V(Z^n) = 1 - 2^{-n}$.[46] Now consider any $Z \subseteq Y \subseteq X$ and $x \in X$. If $x \in Y$, then $C_f(Y) = C_f(\{x\} \cup Y)$ and we are done; if $x \notin C_f(\{x\} \cup Y)$, then the same conclusion holds. Now, if $x \notin Y$ and $x \in C_f(\{x\} \cup Y)$, then, as the preferences of f are substitutable, $x \in C_f(\{x\} \cup Z)$. Let $Z^n = C_f(Z)$ and $Z^{n'} = C_f(Y)$ where $n \leq n'$ as $Z \subseteq Y$. Hence $V(\{x\} \cup Z) - V(Z) \geq 2^{-n-1} \geq 2^{-n'-1} \geq V(\{x\} \cup Y) - V(Y)$ and so

$$V(\{x\} \cup Z) - V(Z) \geq V(\{x\} \cup Y) - V(Y),$$

so that V is submodular.

Proof of Proposition 2

We prove a lemma which directly implies the result.

Lemma 2. *Suppose that Z is a blocking set for Y, and that the preferences of all agents are substitutable. Then for any $z \in Z$, the set $\{z\}$ is a blocking set for Y.*

Proof. If Z is a blocking set for Y, then

$$Z \subseteq C_H(Y \cup Z) \tag{1}$$

by definition.

We fix any $z \in Z$. By (1), we have $z \in C_{z_H}(Y \cup Z)$; as the preferences of z_H are substitutable, we must then have $z \in C_{z_H}(Y \cup \{z\})$. Similarly, we find that $z \in C_{z_D}(Y \cup \{z\})$. It follows that $\{z\}$ is a blocking set for Y. □

Lemma 2 implies the result, as it shows that if there exists some Z blocking Y, then there is some Z' with $|Z'| = 1$ which blocks Y, as well—indeed, taking $Z' = \{z\}$ for any $z \in Z$ suffices. Thus, any Y which is not blocked in the sense of Definition 2 cannot be pairwise stable (and hence cannot be many-to-one stable).

Proof of Proposition 3

Suppose that Y is blocked in X' by some set of contracts $Z' \subseteq X'$.[47] Since $X \rhd X'$, we have $Z' \subseteq X' \subseteq X$. But $Z' \nsubseteq Y$ and by construction we must have $Z'_f \subseteq C_f^X(Z' \cup Y)$ for each $f \in F$, contradicting the stability of Y with respect to X.

[46] A similar method is used by Chambers and Echenique (2009) to prove that for any increasing quasisupermodular function, there exists a monotonic transformation such that the transformed function is supermodular.

[47] It is clear that Y is individually rational, so Y can only be unstable if it is blocked.

Proof of Proposition 4

If $P_f^{X'}$ is not substitutable for some $f \in F$, then there exist $z, x \in X'$ and $Y \subseteq X'$ such that $z \notin C_f^{X'}(Y \cup \{z\})$ but $z \in C_f^{X'}(Y \cup \{z, x\})$. But $P_f^{X'}$ is just the restriction of P_f^X to sets of contracts wholly contained in X', so in particular $z, x \in X'$ and $Y \subseteq X'$ comprise a counterexample to the substitutability of P_f^X.

Proof of Lemma 1

Throughout the proof, we use the fact that all choice functions we consider satisfy the following irrelevance of rejected contracts condition.[48]

Definition 6. A choice function C^f satisfies the *irrelevance of rejected contracts* condition if for all $Y \subseteq X$ and $z \in X \smallsetminus Y$, if $z \notin C^f(Y \cup \{z\})$, then $C^f(Y \cup \{z\}) = C^f(Y)$.

If (X^D, X^H) is a fixed point of Φ, then for any $x \in X^D \cap X^H \subseteq X^D$, we have $X^D = \Phi_H(X^D) = \{x \in X : x \in C_D(X^D \cup \{x\})\}$, so that

$$x \in C_D(X^D \cup \{x\}) = C_D(X^D). \tag{2}$$

As each doctor has substitutable preferences, (2) implies that $x \in C_D(X^D \cap X^H)$. By analogous reasoning, we see that $x \in C_H(X^D \cap X^H)$. Hence, we see that $X^D \cap X^H$ is individually rational.

Now, we suppose that $X^D \cap X^H$ is blocked. Then, by Proposition 2, there exists a blocking set $\{z\} \not\subseteq X^D \cap X^H$. Then, either $z \notin X^D$ or $z \notin X^H$. We assume the former case ($z \notin X^D$); the latter is analogous. Now, again as (X^D, X^H) is a fixed point of Φ, we have $X^D = \Phi_H(X^D) = \{x \in X : x \in C_D(X^D \cup \{x\})\}$; hence, we have that $z \notin C_H(X^H \cup \{z\})$. We suppose, by way of contradiction, that $z \in C_H((X^D \cap X^H) \cup \{z\})$. Then, by the irrelevance of rejected contracts condition, there exists at least one $x \in (C_H(X^H \cup \{z\})) \cap (X^H \smallsetminus X^D)$ with $x \neq z$. But then substitutability of hospital preferences implies that $x \in C_H(X^H) = C_H(X^H \cup \{x\})$. Thus, we must have $x \in X^D = \Phi_H(X^D) = \{x \in X : x \in C_D(X^D \cup \{x\})\}$ by the definition of Φ—a contradiction.

Now, we suppose that A is a stable outcome. We let $(X^D, X^H) = \Phi(A, A)$. If $X^D \not\supseteq A$, then $C_H(A) \neq A$, and so A is not individually rational for some hospital, contradicting the stability of A. Analogously, if $X^H \not\supseteq A$, then $C_D(A) \neq A$, and so A is not individually rational for some doctor, contradicting the stability of A. If $z \in (X^D \cap X^H) \smallsetminus A$, then we have $z \in C_D(A \cup \{z\})$ and $z \in C_H(A \cup \{z\})$ (by the definition of Φ and the substitutability of the choice functions of z_D and z_H). It follows that $\{z\}$ blocks A, contradicting the stability of A. Hence, we see that $A = X^D \cap X^H$.

Next we show that $(X^D, X^H) = \Phi(A, A)$ is a fixed point of Φ. First, we consider $\Phi_D(X^H) = \{x \in X : x \in C_H(X^H \cup \{x\})\}$. There are two cases to consider:

1. Suppose that $y \in \Phi_D(X^H) \smallsetminus X^D$. Since $y \in \Phi_D(X^H)$, we have $y \in C_H(X^H \cup \{y\})$, implying by substitutability that $y \in C_H(A \cup \{y\})$; hence, we have $y \in X^D = \Phi_H(A)$, a contradiction.
2. Suppose that $y \in X^D \smallsetminus \Phi_D(X^H)$. Then $y \in C_H(A \cup \{y\})$; hence, if $y \notin \Phi_D(X^H)$, there exists a $z \in X^H \smallsetminus A$ such that $z \in C_H(X^H \cup \{y\})$ by the irrelevance of rejected contracts condition. Then, by substitutability, we must have $z \in C_H(A \cup \{z\})$. But $X^H = \Phi_H(A)$, so $z \in C_D(A \cup \{z\})$, implying that $\{z\}$ blocks A, contradicting the stability of A.

The logic that $\Phi_H(X^D) = X^H$ is analogous.

Finally, we show that there does not exist any fixed point $(\tilde{X}^D, \tilde{X}^H) \neq (X^D, X^H)$ such that $\tilde{X}^D \cap \tilde{X}^H = A$. We first show that $C_D(\tilde{X}^D) = A$: If $C_D(\tilde{X}^D) \subsetneq A$, then A is not individually rational, contradicting the stability of A. If $y \in C_D(\tilde{X}^D) \smallsetminus A$, then $y \in \tilde{X}^H = \Phi_H(\tilde{X}^D)$ (as $(\tilde{X}^D, \tilde{X}^H)$ is a fixed point of Φ), and so $y \in \tilde{X}^H \cap \tilde{X}^D$—a contradiction.

Now, $C_D(\tilde{X}^D) = A$, we have that $(\tilde{X}^D \smallsetminus A) \cap C_D(\tilde{X}^D) = \varnothing$. By substitutability, $(\tilde{X}^D \smallsetminus A) \cap C_D(\tilde{X}^D \cup \{x\}) = \varnothing$ for all $x \in X$. Hence, by the irrelevance of rejected contracts condition,

$$\tilde{X}^H = \{x \in X : x \in C_D(\tilde{X}^D \cup \{x\})\} = \{x \in X : x \in C_D(A \cup \{x\})\} = \Phi_H(A) = X^H.$$

An analogous argument shows that that $\tilde{X}^D = X^D$, so we cannot have $(\tilde{X}^D, \tilde{X}^H) \neq (X^D, X^H)$.

[48] Our choice functions satisfy the irrelevance of rejected contracts condition because they are induced by strict preference relations (see Aygün and Sönmez (2014, 2012, 2013)).

Proof of Theorem 1

We first verify that the operator Φ is *isotone* with the respect to the ordering $\vdash$, where $(X^D, X^H) \vdash (\tilde{X}^D, \tilde{X}^H)$ if $X^D \subseteq \tilde{X}^D$ and $X^H \supseteq \tilde{X}^H$. In other words, we show that if $(X^D, X^H) \vdash (\tilde{X}^D, \tilde{X}^H)$, then $(\Phi_D(X^H), \Phi_H(X^D)) \vdash (\Phi_D(\tilde{X}^H), \Phi_H(\tilde{X}^D))$, i.e., $\Phi_D(X^H) \subseteq \Phi_D(\tilde{X}^H)$ and $\Phi_H(X^D) \supseteq \Phi_H(\tilde{X}^D)$. To see this, we note that if $x \in C_H(X^H \cup \{x\})$, then $x \in C_H(\tilde{X}^H \cup \{x\})$, as each hospital has substitutable preferences. Hence, we see that $\Phi_D(\tilde{X}^H) \supseteq \Phi_D(X^H)$. The proof that $\Phi_H(\tilde{X}^D) \subseteq \Phi_H(X^D)$ is analogous; hence, we see that Φ is isotone.

As Φ is isotone on the offer set lattice, it follows from Tarski's fixed-point theorem that there exists a nonempty lattice of fixed points of the operator Φ. These correspond to stable outcomes by Lemma 1.

To prove the lattice structure result, we first show that that $A = X^D \cap X^H$ is chosen by the doctors from X^D, i.e., $A = C_D(X^D)$. There are two cases to check:

1. Suppose there exists $z \in C_D(X^D) \smallsetminus A$. Then $z \in C_D(X^D \cup \{z\})$, and hence $z \in \Phi_H(X^D)$. But then, since (X^D, X^H) is a fixed point of Φ, we have $\Phi_H(X^D) = X^H$, so that $z \in \Phi_H(X^D) = X^H$, contradicting the assumption that $z \notin A = X^D \cap X^H$.
2. Suppose there exists $z \in A \smallsetminus C_D(X^D)$. Then there exists a $z \in A = X^D \cap X^H$ such that $z \notin C_D(X^D)$ and, hence, as $z \in X^D$, we must have $z \notin C_D(X^D \cup \{z\})$. But then, we have $z \notin \Phi_H(X^D)$, so that $\Phi_H(X^D) \neq X^H \ni z$, so that (X^D, X^H) cannot be a fixed point of Φ.

Thus, we see that for any fixed point of the lattice (X^D, X^H), we have $C_D(X^D) = X^D \cap X^H$.

The preceding observation implies that for two fixed points (X^D, X^H) and $(\tilde{X}^D, \tilde{X}^H)$ corresponding to the outcomes A and $\tilde{A}$, respectively, if $(\tilde{X}^D, \tilde{X}^H) \vdash (X^D, X^H)$, then $\tilde{X}^D \subseteq X^D$, and so $A = C_D(X^D) \succeq_D C_D(\tilde{X}^D) = \tilde{A}$. Hence, since the set of fixed points is a lattice with respect to $\vdash$, the set of stable outcomes corresponding to those fixed points is a lattice with respect to $\succeq_D$.

Proof of Theorem 2

If the preferences of a hospital h are not substitutable, then there exist contracts $x, z \in X_h$ and a set of contracts $Y \subseteq X \smallsetminus \{x, z\}$ such that $Y_H = \{h\}$ and

$$\begin{aligned} &z \notin C_h(Y \cup \{z\}) \\ &z \in C_h(\{x\} \cup Y \cup \{z\}). \end{aligned}$$

There are two cases to consider.

Case 1: $x_D \neq z_D$. By assumption, there must exist a hospital $h' \neq h$. Furthermore, there must exist contracts x' and z' with $x_D = x'_D$, $z_D = z'_D$ and $x'_H = z'_H = h'$.

Let z_D have preferences such that

$$C_{z_D}(W) = \begin{cases} (W \cap (Y \cup \{z\}))_{z_D} & \{z, z'\} \subseteq W \\ (W \cap (Y \cup \{z, z'\}))_{z_D} & \text{otherwise.} \end{cases}$$

That is, z_D is willing to accept any and all of the contracts he is associated with in Y, and z_D wants one of z and z', preferring z, and rejects all other contracts.[49]

Let x_D have preferences such that

$$C_{x_D}(W) = \begin{cases} (W \cap (Y \cup \{x'\}))_{x_D} & \{x, x'\} \subseteq W \\ (W \cap (\{x\} \cup Y \cup \{x'\}))_{x_D} & \text{otherwise.} \end{cases}$$

[49] More precisely, C_{z_D} is induced by a preference relation $\succ_{z_D}$ as follows: Let $\{y^1, \ldots, y^N\} \equiv Y_{z_D}$, and let the score of a set $W \subseteq X_{z_D}$ be given by

$$\mathsf{s}_{z_D}(W) = 2\mathbb{1}_{\{z \in W\}} + \mathbb{1}_{\{z' \in W\}} + \sum_{n=1}^{N} \frac{1}{2^n} \mathbb{1}_{\{y^n \in W\}} - 5\mathbb{1}_{\{\{z,z'\} \subseteq W\}} - 10 \left(\sum_{w \in X_{z_D} \smallsetminus (Y_{z_D} \cup \{z, z'\})} \mathbb{1}_{\{w \in W\}} \right).$$

We then let $W \succ_{z_D} W'$ if $\mathsf{s}_{z_D}(W) > \mathsf{s}_{z_D}(W')$.

That is, x_D is willing to accept any and all of the contracts he is associated with in Y, and x_D wants one of x and x', preferring x', and rejects all other contracts.[50]

Let h' have preferences given by $\{z'\} \succ_{h'} \{x'\} \succ_{h'} \varnothing$, which induces the choice function

$$C_{h'}(W) = \begin{cases} \{z'\} & z' \in W \\ \{x'\} & x' \in W \text{ and } z' \notin W \\ \varnothing & \text{otherwise.} \end{cases}$$

Finally, let every doctor $d \in D \smallsetminus \{x_D, z_D\}$ have preferences such that

$$C_d(W) = (W \cap Y)_d.$$

That is, d is willing to accept any and all of the contracts he is associated with in Y, and rejects all other contracts.[51]

Consider any outcome A; we show A can not be stable.

1. Suppose $C_h(Y \cup \{z\}) \succ_h A_h$. If A is individually rational for all hospitals, then $C_h(Y \cup \{z\})$ blocks A, as all doctors choose their contracts in $C_h(Y)$.
2. Suppose $A_h = C_h(Y \cup \{z\})$. Then $z' \in A$, as otherwise $\{z'\}$ blocks A. Hence, by the individual rationality of h', we have that $x' \notin A$. But then $C_h(\{x\} \cup Y \cup \{z\})$ blocks A.
3. Suppose $C_h(\{x\} \cup Y \cup \{z\}) \succ_h A_h \succ_h C_h(Y \cup \{z\})$. In this case, if A is individually rational for all hospitals, then $A \subseteq \{x, x', z'\} \cup Y \cup \{z\}$; then $x \in A$ as otherwise we could not have $A_h \succ_h C_h(Y \cup \{z\})$. But then, $C_h(\{x\} \cup Y \cup \{z\})$ blocks A.
4. Suppose $C_h(\{x\} \cup Y \cup \{z\}) = A_h$. Then if $z' \in A$, the outcome A is not individually rational for z_D, and if $x' \in A$, the outcome A is not individually rational for x_D; but this implies that $\{x'\}$ blocks A.

Case 2: $x_D = z_D \equiv d$. By assumption, there are two hospitals, h' and h'', such that $h \neq h' \neq h'' \neq h$ and one doctor $\hat{d} \neq d$. Now consider the contracts $x', x'', \hat{x}'$, and $\hat{x}''$ such that $x'_D = x''_D = d$, $\hat{x}'_D = \hat{x}''_D = \hat{d}$, $x'_H = \hat{x}'_H = h'$ and $x''_H = \hat{x}''_H = h''$, which exist by assumption.

Let d have preferences such that

$$C_d(W) = (W \cap Y)_d \cup \tilde{C}_d(W \cap \{x, z, x', x''\})$$

where $\tilde{C}_d(\tilde{W})$ is the responsive choice function over $\{x, z, x', x''\}$ with quota 2 and underlying preference order $x'' \succ z \succ x \succ x'$. That is, d is willing to accept any and all of the contracts he is associated with in Y, and d wants two of x'', z, x, and x', preferring x'' to z to x to x', and rejects all other contracts.[52]

[50] More precisely, C_{x_D} is induced by a preference relation $\succ_{x_D}$ as follows: Let $\{y^1, \ldots, y^N\} \equiv Y_{x_D}$, and let the score of a set $W \subseteq X_{x_D}$ be given by

$$\mathsf{s}_{x_D}(W) = 2\mathbb{1}_{\{x' \in W\}} + \mathbb{1}_{\{x \in W\}} + \sum_{n=1}^{N} \frac{1}{2^n} \mathbb{1}_{\{y^n \in W\}} - 5\mathbb{1}_{\{\{x, x'\} \subseteq W\}} - 10 \left(\sum_{w \in X_{x_D} \smallsetminus (Y_{x_D} \cup \{x, x'\})} \mathbb{1}_{\{w \in W\}} \right).$$

We then let $W \succ_{x_D} W'$ if $\mathsf{s}_{x_D}(W) > \mathsf{s}_{x_D}(W')$.

[51] More precisely, for each d we have that C_d is induced by a preference relation $\succ_d$ as follows: Let $\{y^1, \ldots, y^N\} \equiv Y_d$, and let the score of a set $W \subseteq X_d$ be given by

$$\mathsf{s}_d(W) = \sum_{n=1}^{N} \frac{1}{2^n} \mathbb{1}_{\{y^n \in W\}} - 10 \left(\sum_{w \in X_d \smallsetminus Y_d} \mathbb{1}_{\{w \in W\}} \right).$$

We then let $W \succ_d W'$ if $\mathsf{s}_d(W) > \mathsf{s}_d(W')$.

[52] More precisely, C_d is induced by a preference relation $\succ_d$ as follows: Let $\{y^1, \ldots, y^N\} \equiv Y_d$, and let the score of a set

We let $\hat{d}$ have preferences such that

$$C_{\hat{d}}(W) = \begin{cases} (W \cap (Y \cup \{\hat{x}'\}))_{\hat{d}} & \{\hat{x}', \hat{x}''\} \subseteq W \\ (W \cap (Y \cup Z \cup \{\hat{x}', \hat{x}''\}))_{\hat{d}} & \text{otherwise,} \end{cases}$$

That is, $\hat{d}$ is willing to accept any and all of the contracts he is associated with in Y, and $\hat{d}$ wants one of $\hat{x}'$ and $\hat{x}''$, preferring $\hat{x}'$, and rejects all other contracts.[53]

We let h' have preferences given by $\{x'\} \succ_{h'} \{\hat{x}'\} \succ_{h'} \varnothing$, which induces the choice function

$$C_{h'}(W) = \begin{cases} \{x'\} & x' \in W \\ \{\hat{x}'\} & \hat{x}' \in W \text{ and } x' \notin W \\ \varnothing & \text{otherwise.} \end{cases}$$

We let h'' have preferences given by $\{\hat{x}''\} \succ_{h''} \{x''\} \succ_{h''} \varnothing$, which induces the choice function

$$C_{h'}(W) = \begin{cases} \{\hat{x}''\} & \hat{x}'' \in W \\ \{x''\} & x'' \in W \text{ and } \hat{x}'' \notin W \\ \varnothing & \text{otherwise.} \end{cases}$$

Finally, let every doctor $\bar{d} \in D \smallsetminus \{d, \hat{d}\}$ have preferences such that

$$C_{\bar{d}}(W) = (W \cap Y)_{\bar{d}}.$$

That is, $\bar{d}$ is willing to accept any and all of the contracts he is associated with in Y, and rejects all other contracts.[54]

Consider any outcome A; we show that A can not be stable.

1. Suppose $C_h(Y \cup \{z\}) \succ_h A_h$. Then $C_h(Y \cup \{z\})$ blocks A, as all the doctors choose their contracts in $C_h(Y \cup \{z\})$.

$W \subseteq X_d$ be given by

$$\mathsf{s}_d(W) = 8\mathbb{1}_{\{x'' \in W\}} + 4\mathbb{1}_{\{z \in W\}} + 2\mathbb{1}_{\{x \in W\}} + \mathbb{1}_{\{x' \in W\}} + \sum_{n=1}^{N} \frac{1}{2^n} \mathbb{1}_{\{y^n \in W\}}$$
$$- 16\mathbb{1}_{\{|\{x,z,x',x''\} \cap W| \geq 3\}} - 32 \left(\sum_{w \in X_d \smallsetminus (Y_d \cup \{x,z,x',x''\})} \mathbb{1}_{\{w \in W\}} \right).$$

We then let $W \succ_d W'$ if $\mathsf{s}_d(W) > \mathsf{s}_d(W')$.

[53] More precisely, $C_{\hat{d}}$ is induced by a preference relation $\succ_{\hat{d}}$ as follows: Let $\{y^1, \ldots, y^N\} \equiv Y_{\hat{d}}$, and let the score of a set $W \subseteq X_{x_D}$ be given by

$$\mathsf{s}_{\hat{d}}(W) = 2\mathbb{1}_{\{\hat{x}' \in W\}} + \mathbb{1}_{\{\hat{x}'' \in W\}} + \sum_{n=1}^{N} \frac{1}{2^n} \mathbb{1}_{\{y^n \in W\}} - 5\mathbb{1}_{\{\{\hat{x}', \hat{x}''\} \subseteq W\}} - 10 \left(\sum_{w \in X_{x_D} \smallsetminus (Y_{x_D} \cup \{\hat{x}', \hat{x}''\})} \mathbb{1}_{\{w \in W\}} \right).$$

We then let $W \succ_{\hat{d}} W'$ if $\mathsf{s}_{\hat{d}}(W) > \mathsf{s}_{\hat{d}}(W')$.

[54] More precisely, for each $\bar{d}$ we have that $C_{\bar{d}}$ is induced by a preference relation $\succ_{\bar{d}}$ as follows: Let $\{y^1, \ldots, y^N\} \equiv Y_{\bar{d}}$, and let the score of a set $W \subseteq X_{\bar{d}}$ be given by

$$\mathsf{s}_{\bar{d}}(W) = \sum_{n=1}^{N} \frac{1}{2^n} \mathbb{1}_{\{y^n \in W\}} - 10 \left(\sum_{w \in X_{\bar{d}} \smallsetminus Y_{\bar{d}}} \mathbb{1}_{\{w \in W\}} \right).$$

We then let $W \succ_{\bar{d}} W'$ if $\mathsf{s}_{\bar{d}}(W) > \mathsf{s}_{\bar{d}}(W')$.

2. Suppose $A_h = C_h(Y \cup \{z\})$. Since d does not obtain x or z, he desires both x' and x''. Hence, if A is stable, we must have that $x' \in A$. Furthermore, since $\hat{d}$ does not obtain $\hat{x}'$ since A is individually rational for h'. Hence, for A to be stable, we must have $\hat{x}'' \in A$. Hence, if A is stable, $\{\hat{x}'', x'\} \subseteq A$ and thus $x'' \notin A$ by individual rationality for h''. In that case, $C_h(\{x\} \cup Y \cup \{z\})$ blocks A.
3. Suppose $C_h(\{x\} \cup Y \cup \{z\}) \succ_h A_h \succ_h C_h(Y \cup \{z\})$. Then $x \in A$, so $C_h(\{x\} \cup Y \cup \{z\}) \smallsetminus \{x\}$ blocks A, as d will always choose z and the other doctors in Y will always accept offers of any and all contracts in Y.
4. Suppose $C_h(\{x\} \cup Y \cup \{z\}) = A_h$. If $\hat{x}' \notin A$, then $\{\hat{x}'\}$ blocks A. (Note that if $x' \in A$, then $\{x, z, x'\} \subseteq A$, and so A is not individually rational for d.) But $\hat{x}' \in A$ implies that $\hat{x}'' \notin A$. Hence $\{x''\}$ blocks A. (Note that $x'' \notin A$, as then $\{x'', x, z\} \subseteq A$, and so A is not individually rational for d.)

Proof of Theorem 3

The forwards direction is trivial, hence we show only the reverse direction. Without loss of generality, we assume that all hospital preferences are strongly stable. Now, we fix preferences, and consider any stable outcome A. If A is not strongly stable, then there exists a set Z such that for each $f \in Z_F$ there exists an individually rational Y^f such that $Z_f \subseteq Y^f \subseteq Z \cup A$ and $Y^f \succ_f A$. Now, consider a doctor $d \in Z_D$. Since $Y^d \succ_d A_d$, $C_d(Z \cup A) \neq A_d$ and hence, as A_d is individually rational for d, there exists $x \in C_d(Z \cup A)$ such that $x \in Z \smallsetminus A$. Hence, by substitutability, we have $x \in C_d(\{x\} \cup A)$. Now, if $x \in C_{x_H}(\{x\} \cup A)$, then $\{x\}$ blocks A, contradicting the stability of A. Hence, $x \notin C_{x_H}(\{x\} \cup A)$, but $x \in C_{x_H}(Y^{x_H})$ and

$$C_{x_H}(Y^{x_H}) = Y^{x_H} \succ_{x_H} C_{x_H}(A) = C_{x_H}(\{x\} \cup A).$$

But $x \in ((A \cup \{x\}) \cap C_{x_H}(Y^{x_H})) \smallsetminus C_{x_H}(A \cup \{x\})$, so the preferences of x_H are not strongly substitutable.

B. The Rural Hospitals Theorem, Strategy-Proofness, and the Weak Pareto Property

We show an analogue of the rural hospitals theorem of Roth (1984a) and Hatfield and Milgrom (2005).

Theorem 4. *If preferences are substitutable and satisfy the law of aggregate demand, then each agent signs the same number of contracts at every stable outcome.*

Proof. Consider any stable outcome A, and the doctor-optimal stable outcome A^*. Since every hospital prefers A to A^* from Theorem 1 and surrounding discussion, if follows from from the law of aggregate demand that the number of contracts signed by each hospital is weakly smaller at A^*, hence $|A^*| \leq |A|$. Hence, if any doctor receives strictly more contracts at A^* than at A, some doctor must receive strictly fewer contracts at A^* than at A. This cannot happen, as every doctor is weakly better off at A^* than at A, and every doctor's preferences satisfy the law of aggregate demand. Thus every doctor receives the same number of contracts at every stable outcome.

An analogous argument shows the result for hospitals. □

Theorem 4 is an immediate consequence of the law of aggregate demand and the lattice structure obtained in Theorem 1. Since for any stable outcome A, every hospital prefers A to the doctor-optimal stable outcome A^*, the fact that hospitals' preferences satisfy the law of aggregate demand guarantees that $|A^*| \leq |A|$. But no doctor can receive strictly more contracts at A^* than at A unless some other doctor receives strictly fewer contracts at A^* than at A. This cannot happen because every doctor is weakly better off at A^* than at A, and every doctor's preferences satisfy the law of aggregate demand as well.

C. Nash Implementability

We show that the stable outcome correspondence is Nash implementable whenever it is nonempty and there are at least three agents. Informally, this means that all stable outcomes can be achieved non-cooperatively, through strategic interactions in equilibrium. The requirement of three agents is as sharp as possible, since Kara and Sönmez (1996) have already proven that the stable matching correspondence is not Nash implementable in the setting of one-to-one matching when there are fewer than three agents.

First, we review some standard terminology and notation. A *generalized matching mechanism* is a pair $(\mathcal{M}, o)$, where $\mathcal{M} \equiv \prod_{f\in F} \mathcal{M}_f$ denotes a set of strategy profiles and o is an *outcome function* mapping strategy profiles into outcomes.[55] As is standard, we identify a mechanism $(\mathcal{M}, o)$ with its outcome function, o. For a given profile P of agents' true preferences, a mechanism o induces a non-cooperative strategic form game $\Gamma_o(P)$, in which the outcome $o(m)$ of a strategy profile $m \in \mathcal{M}$ is evaluated using agents' true preferences.

We write NE$(\cdot)$ for the Nash equilibrium correspondence. A mechanism o is said to *Nash implement solution* φ if, for all possible profiles P,

$$\varphi(P) = o(\mathrm{NE}(\Gamma_o(P))).$$

That is, o Nash implements φ if the set of outcomes in $\varphi(P)$ are exactly those which are the outcomes (under o) of Nash equilibria of $\Gamma_o(P)$.

Now, we state our implementability result.

Theorem 5. *If $|F| \geq 3$, then the stable outcome correspondence is Nash implementable whenever it is nonempty.*

Theorem 5 subsumes the analogous results of Kara and Sönmez (1996, 1997) and Haake and Klaus (2009a,b) for less-general matching settings. The proof of Theorem 5 is a straightforward generalization of the argument used by Haake and Klaus (2009a) in the setting of many-to-one matching with contracts; hence, we omit it.[56]

Combining Theorem 5 with Theorem 1 shows in particular that the stable outcome correspondence is Nash implementable when all agents' preferences are substitutable. An additional consequence of Theorem 5 is that the stable matching correspondence is monotonic in the sense of Maskin (1999).

[55] We use the adjective "generalized" to indicate that, unlike in standard formulation of a matching mechanism, here we consider as input a generalized space of strategies rather than the space of preference profiles.

[56] In fact, the argument follows that of Haake and Klaus (2009a) directly, but is slighlty simpler in our framework. Specifically, the first subargument of Step 3 in the proof given by Haake and Klaus (2009a) can be omitted, since in our framework both doctors and hospitals may accept multiple contracts.