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

John William Hatfield          Scott Duke Kominers
McCombs School of Business     Society of Fellows
University of Texas at Austin  Harvard University

First Version: October 2009
Current Version: October 2014

Abstract

We develop a model of many-to-many matching with contracts which 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. We show that bundling contractual primitives encourages substitutability of agents’ preferences over contracts; however, such bundling also makes the contractual language less expressive. Thus, we see that, in choosing contract language, market designers face a tradeoff between expressiveness and stability.

JEL Codes: C78, C62, D47, L14

Keywords: Many-to-Many Matching, Stability, Substitutes, Contract Design

*This paper contains parts of the earlier working versions entitled “Contract Design and Stability in Matching Markets” and “Many-to-Many Matching with Contracts.” The authors thank Eric Budish, Jeremy Bulow, Peter Cramton, Federico Echenique, Sonia Jaffe, Fuhibo Kojima, David Laibson, Paul Milgrom, Muriel Niederle, David Parkes, Alvin E. Roth, Tayfun Sönmez, Alexander Westkamp, E. Glen Weyl, seminar audiences at Chicago, Harvard, Northwestern, and Stanford, and especially Drew Fudenberg for helpful comments and discussions. Much of this work was completed while Hatfield was at the Stanford University Graduate School of Business and Kominers was at the Becker Friedman Institute of the University of Chicago. Hatfield appreciates the hospitality of Harvard Business School which hosted him during parts of this research. Additionally, Kominers gratefuly acknowledges the support of a National Science Foundation Graduate Research Fellowship, National Science Foundation Grant CCF-1216095, 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. Please e-mail comments or suggestions to john.hatfield@utexas.edu and kominers@fas.harvard.edu.
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 our model is matching with couples, in which pairs of individuals may choose to act as a single agent which receives (at most) two assignments (see Klaus and Klijn (2005); Klaus et al. (2007)). 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)), and buyer–seller markets with heterogeneous and indivisible goods.

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. Moreover, substitutability is necessary for the existence of stable outcomes in the maximal domain sense: that is, substitutable preferences are required in the sense that 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 (2014)).

---

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

2Our 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.

3We 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).
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. As we illustrate, unitarity implicitly restricts the preferences that can be considered substitutable, and may affect the set of stable outcomes.\(^5\) Moreover, for many economic environments, the validity of the unitarity assumption is far from apparent—even in labor markets, agents might contract separately over multiple shifts at the same employer.

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 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 and 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 \(x^\Pi\) the contract with terms \(\Pi \subseteq \{m, a\}\). When morning and afternoon shifts are contracted separately, the doctor’s

\(^4\)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 notion of setwise stability studied by Echenique and Oviedo (2006) and Klaus and Walzl (2009).

\(^5\)Additionally, as Echenique (2012) showed, the unitarity assumption induces an embedding of the many-to-one matching with contracts model into the seemingly simpler matching with salaries framework of Kelso and Crawford (1982).
preferences over contracts are given by
\[ P_d : \{ x^{(m)} \} \succ \{ x^{(a)} \} \succ \emptyset \succ \{ x^{(m)} \}, \]
while the hospital’s are given by
\[ P_h : \{ x^{(m)} \} \succ \{ x^{(m)}, x^{(a)} \} \succ \{ x^{(a)} \} \succ \emptyset. \]

There is no stable contracting outcome: for the set \( \{ x^{(m)}, x^{(a)} \} \), the hospital will not be willing to sign \( x^{(a)} \); for the set \( \{ x^{(a)} \} \), both parties prefer that the doctor work full time; the set \( \{ x^{(m)} \} \) is not individually rational for the doctor; and finally both parties agree that \( \{ x^{(a)} \} \) 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 \( x^{(a)} \) which 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, agents’ preferences reduce to the forms
\[ P_d : \{ x^{(m,a)} \} \succ \emptyset, \quad P_h : \{ x^{(m,a)} \} \succ \emptyset. \]

These preferences are substitutable, and there exists a unique stable outcome \( \{ x^{(m,a)} \} \).

In Section 3, we introduce a theory which formalizes the intuitions from this example regarding how contract language affects the substitutability of preferences and the stability of contract outcomes. This 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 contractual 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.

\[ ^{6} \text{Although using a unitary contract language is beneficial here, this is not always true in general (see Section 3.4).} \]
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 uninterested in the discussion of contract language 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 will 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 $P_f$. 

5
For any \( f \in F \) and offer set \( Y \subseteq X \), we let
\[
C_f(Y) \equiv \max_{P_f} \{ Z \subseteq Y : x \in Z \Rightarrow f \in x \}
\]
be the set of contracts \( f \) chooses from \( Y \).\(^7\)\(^8\)\(^9\) We let
\[
R_f(Y) \equiv Y \setminus C_f(Y)
\]
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 \( R_D(Y) \equiv Y \setminus C_D(Y) \). 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 \( R_H(Y) \equiv Y \setminus C_H(Y) \).\(^{10}\)

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 \( Y \succ_f Z \).

### 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 \( f \). More formally:

**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 \( z \notin C_f(\{x\} \cup Y \cup \{z\}) \).

\(^7\)We use the term “offer set” instead of “budget set” or “set of alternatives,” as agents are typically allowed to choose only one option (or point, or bundle) from a budget set. (See the definition given on the first page of Chapter 1 of *Mas-Colell et al.* (1995), for instance.) Here, agents may choose any subset of the set of contracts offered.

\(^8\)We use the notation \( \max_{P_f} \) to indicate that the maximization is taken with respect to the preferences of agent \( f \).

\(^9\)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 convention 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 (2014a,b)).

\(^{10}\)Note that \( R_D(Y) = \bigcup_{d \in D} R_d(Y_d) \neq \bigcup_{d \in D} R_d(Y) \) and similarly \( R_H(Y) = \bigcup_{h \in H} R_h(Y_h) \neq \bigcup_{h \in H} R_h(Y) \).
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 monotone, i.e., if for any $Y' \subseteq Y \subseteq X$, we have that $R_f(Y') \subseteq R_f(Y)$.

An alternative characterization of substitutability can be obtained in terms of submodularity of the indirect utility function: We say that the indirect utility function $U$ over offer sets represents preference relation $P_f$ if

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

That is, under $U$, 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

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 $Z \cap A = \emptyset$ 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. 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

---

11In particular, this definition is equivalent to that of Hatfield and Milgrom (2005) in the context of many-to-one matching with contracts.

12Our stability concept allows agents associated with a blocking set to disagree as to whether a contract in the original outcome is maintained while deviating. This possibility may be unrealistic in some markets, such as those where agents can tie one contract to another.
that \(|Z| = 1\); we call this \textit{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 \textit{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.

\textbf{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 \(f \in z_F\).

\section{Contract Language}

We now develop a theory of the contract set \(X\) as a language for expressing bundles 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.

\subsection{Basic Theory of Language}

For each doctor–hospital pair \((d, h) \in D \times H\), there is a set of \textit{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 $\pi(d, h) \cap \pi(d', h') = \emptyset$ for all $(d, h) \neq (d', h')$ so that each primitive uniquely identifies a doctor and hospital. A \textit{primitive outcome} is a collection of primitives
\[
\Lambda \subseteq \bigcup_{(d, h) \in D \times H} \pi(d, h).
\]

A \textit{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 \textit{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 $\mathcal{P}(\pi(d, h)) \setminus \{\emptyset\}$. More generally, a \textit{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 $X_{(d, h)} \subseteq \mathcal{P}(\pi(d, h)) \setminus \{\emptyset\}$ for each $(d, h) \in D \times H$. We say that a primitive outcome $\Lambda$ is \textit{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 $Y$ \textit{expresses} $\Lambda$.

Each $f \in F$ has a strict preference relation $P_f$ over the set $\mathcal{P}(\Pi_f)$ of bundles of primitives involving $f$. For any contract language $X$, this preference relation over primitives induces a preference relation, denoted $P_f^X$, over bundles of contracts in $X$ (i.e. subsets of $\mathcal{P}(X)$). This induced preference relation is not strict, but its only indifferences arise on bundles of contracts $Y, Y' \subseteq X$ that correspond to the same primitive outcome $(\bigcup_{y \in Y} y = \bigcup_{y' \in Y'} y')$. When describing preferences in the sequel, despite indifference between these \textit{primitive-equivalent} sets of contracts, we will typically assume that the preference relation $P_f^X$ is strict, arbitrarily breaking ties among primitive-equivalent contract sets.\footnote{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}
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 will 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.

### 3.2 Language and Stability

If a primitive outcome \( \Lambda \) is expressible in the contract language \( X \) by a stable outcome \( Y \), then we say that \( \Lambda \) is stable with respect to the contract language \( X \).\(^\text{14}\)

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 any expression of a primitive outcome \( \Lambda \), the outcome \( Y \subseteq X \) expresses \( \Lambda \) and is stable with respect to \( X \), then for any tie-breaking there is a (possibly distinct) outcome \( Y' \subseteq X \) which expresses \( \Lambda \) 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.

\(^\text{14}\)Unfortunately, although agents are indifferent over contract sets which express the same primitive outcomes, not all expressions of a primitive outcome \( \Lambda \) stable with respect to \( X \) need be stable.
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 $X \supseteq X'$. In this case, we also say that $X'$ is coarser than (or coarsens) $X$ and write $X \triangleright X'$.\(^{15}\)

Refinement of a language $X'$ corresponds to an increase in expressiveness: if $X \triangleright X'$, then each agent may express a richer preference relation over contracts in $X$ than she can over contracts in $X'$.\(^{16}\) With this ordering $\triangleright$, 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 $X \triangleright X'$ and that an outcome $Y \subseteq X'$ is stable in $X$. Then, $Y$ is stable in $X'$.

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 $\pi(d, h) = \{w, \$\}$.\(^{16}\)

\(^{15}\)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 $\triangleright$ to help clarify when we are actively comparing two contract languages.

\(^{16}\)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 $\Pi_f$. 

11
We suppose that agents’ underlying preferences take the natural form
\[ P_d : \{\$\} \succ \{w, \$\} \succ \emptyset, \quad P_h : \{w\} \succ \{w, \$\} \succ \emptyset. \]

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 \( \Pi \). When all contracts are possible—
\[
X = \{x^{\{w\}}, x^{\{\$\}}, x^{\{w, \$\}}\}
\]
preferences over contracts are
\[
P^X_d : \{x^{\{\$\}}\} \succ \{x^{\{w, \$\}}\} \sim \{x^{\{w\}}, x^{\{\$\}}\} \succ \emptyset,
\]
\[
P^X_h : \{x^{\{w\}}\} \succ \{x^{\{w, \$\}}\} \sim \{x^{\{w\}}, x^{\{\$\}}\} \succ \emptyset,
\]
and the unique stable outcome is \( \{x^{\{w, \$\}}\} \) (regardless of the tie-breaking rule used for indifferences in agents’ preferences (see Footnote 13)). If we coarsen \( X \) to \( X' = \{\emptyset, x^{\{w\}}, x^{\{\$\}}\} \) by removing the contract \( x^{\{w, \$\}} \), agents’ preferences reduce to
\[
P^{X'}_d : \{x^{\{\$\}}\} \succ \{x^{\{w\}}, x^{\{\$\}}\} \succ \emptyset,
\]
\[
P^{X'}_h : \{x^{\{w\}}\} \succ \{x^{\{w\}}, x^{\{\$\}}\} \succ \emptyset,
\]
under which only \( \emptyset \) is stable.\(^{17}\) Thus, we see that the stability of the primitive outcome \( \{w, \$\} \) is not preserved under the coarsening.

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. Hence, a natural question is whether a primitive outcome \( \Lambda \) that is stable 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 \( \Lambda \) into a single contract. 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

\(^{17}\)Note that \( \emptyset \) is not blocked by \( \{x^{\{w\}}, x^{\{\$\}}\} \), as \( C^{X'}_d(\{x^{\{w\}}, x^{\{\$\}}\}) = \{x^{\{\$\}}\} \neq \{x^{\{w\}}, x^{\{\$\}}\} \).

12
Suppose that the (substitutable) preferences of the agents are given by:

\[ P^X_h : \{ y, \hat{z} \} \succ \{ \hat{y}, \hat{z} \} \succ \{ y, z \} \succ \{ \hat{y}, z \} \succ \{ z \} \succ \emptyset, \]
\[ P^X_{h'} : \{ y', z' \} \succ \{ y' \} \succ \{ z' \} \succ \emptyset, \]
\[ P^X_d : \{ y, z \} \succ \{ y \} \succ \{ z \} \succ \emptyset, \]
\[ P^X_{d'} : \{ \hat{y}, z' \} \succ \{ \hat{y}, \hat{z} \} \succ \{ y', z' \} \succ \{ y', \hat{z} \} \succ \{ \hat{y} \} \succ \{ z' \} \succ \{ \hat{z} \} \succ \{ y' \} \succ \emptyset. \]

In this case, the only stable outcome is \( Y = \{ y, z, y', z' \} \). However, if we consider the contractual language obtained by binding together \( y \) and \( z \), and \( y' \) and \( z' \)—\( \hat{X} = \{ x, x', \hat{y}, \hat{z} \} \),

where \( x = y \cup z \) and \( x' = y' \cup z' \)—as shown in Figure 2, the preferences of the agents now

\[ \text{Figure 1: Diagram of contracts under language } X. \]

\[ \text{Figure 2: Diagram of contracts under language } \hat{X}. \]
take the form

\[ P^X_h : \{\hat{y}, \hat{z}\} \succ \{x\} \succ \{\hat{z}\} \succ \{\hat{y}\} \succ \emptyset, \]
\[ P^X_{h'} : \{x'\} \succ \emptyset, \]
\[ P^X_d : \{x\} \succ \emptyset, \]
\[ P^X_{d'} : \{\hat{y}, \hat{z}\} \succ \{x'\} \succ \{\hat{y}\} \succ \{\hat{z}\} \succ \emptyset. \]

For this contract language, \( \{x, x'\} \) (which is primitive-equivalent to \( Y \)) is not stable, as \( Z = \{\hat{y}, \hat{z}\} \) constitutes a blocking set. 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

\[ P_d : \{m, a\} \succ \emptyset, \quad P_h : \{m, a\} \succ \emptyset. \]

Both agents want to contract over a full-time job, but 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

\[ P_d^{\{x^{(m,a)}\}} : \{x^{(m,a)}\} \succ \emptyset, \quad P_h^{\{x^{(m,a)}\}} : \{x^{(m,a)}\} \succ \emptyset. \]

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 $X \succ X'$ 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 any collection of contractual relationships between a doctor and a hospital may be bound together into a single contract. This contract structure is in fact required by *Klaus and Walzl* (2009), who allow at most one contract between each doctor–hospital pair. However, imposing this requirement may obscure substitutable structure within agents’ preferences, as the following example shows.

Consider a hospital $h$ with two tasks, $\gamma$ and $\sigma$. We suppose that

- doctor $d$ can do either or both of tasks $\gamma$ and $\sigma$; and
- doctor $d'$ can only do task $\sigma$.
The hospital would like to assign a doctor to task $\gamma$, but would prefer that $d'$, rather than $d$, complete task $\sigma$. These are natural preferences, and intuitively they should be substitutable, as $d'$ “substitutes” for $d$ in performing task $\sigma$. 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 $h$:

- $x^{(d,\gamma)}$, where $d$ performs only task $\gamma$,
- $x^{(d,\sigma)}$, where $d$ performs only task $\sigma$, and
- $x^{(d,\gamma),(d,\sigma)}$, where $d$ performs both tasks.

Using similar notation, we let $x^{(d',\sigma)}$ be the contract between $d'$ and $h$ that specifies that $d'$ performs task $\sigma$. Hence, the set of contracts is given by $X = \{x^{(d,\gamma)}, x^{(d,\sigma)}, x^{(d,\gamma),(d,\sigma)}, x^{(d',\sigma)}\}$.

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), the preferences of $h$ would take the form

$$R_h (\{x^{(d,\gamma),(d,\sigma)}, x^{(d,\gamma)}\}) = \{x^{(d,\gamma)}\} \not\supseteq \{x^{(d,\gamma),(d,\sigma)}\} = R_h (\{x^{(d,\gamma),(d,\sigma)}, x^{(d,\gamma)}, x^{(d',\sigma)}\}).$$

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

$$P^X_h : \{x^{(d,\gamma)}, x^{(d',\sigma)}\} \succ \{x^{(d,\gamma)}, x^{(d,\sigma)}\} \succ \{x^{(d,\gamma)}\} \succ \{x^{(d',\sigma)}\} \succ \{x^{(d,\sigma)}\} \succ \emptyset.$$

---

19Perhaps $\sigma$ is a specialist task, such as radiology work, while $\gamma$ is a general-practice task. Doctor $d$ is a general practitioner, and hence can perform both tasks $\gamma$ and $\sigma$. Meanwhile, doctor $d'$ is a specialist, and therefore can only perform task $\sigma$, but can perform $\sigma$ better than $d$ can. Here, $(d, \gamma)$, $(d, \sigma)$ and $(d', \gamma)$ are the contractual primitives, and

$$\pi(d, h) = \{(d, \gamma), (d, \sigma)\}, \quad \pi(d', h) = \{(d', \sigma)\}.$$

---

20Note that the preferences do not take the form used in our framework, as the set $\{x^{(d,\gamma),(d,\sigma)}\}$ is primitive-equivalent to $\{x^{(d,\gamma)}, x^{(d,\sigma)}\}$, and so in our model the hospital $h$ should be indifferent between these two sets of contracts.
This rewritten preference relation makes clear the intuitive fact that \( d' \) substitutes for \( d \) in performing task \( \sigma \). 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 our subsequent discussion, we assume the possibility of multiple contracts between doctor–hospital pairs.\(^{21}\) As the example just presented suggests, the class of substitutable preferences in our framework therefore includes many sets of preferences that are naturally substitutable but were not considered substitutable in previous 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 \( \Phi \); we show that fixed points of \( \Phi \) 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 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). We let

\[
\Phi(X_D, X_H) \equiv (\Phi_D(X_H), \Phi_H(X_D))
\]

\[
\Phi_D(X_H) \equiv \{x \in X : x \in C_H(X_H \cup \{x\})\}
\]

\[
\Phi_H(X_D) \equiv \{x \in X : x \in C_D(X_D \cup \{x\})\}.
\]

\(^{21}\)This is a substantive assumption on the contract set \( X \), but a very weak one.
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 $\Phi$, 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 $\Phi$, 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 $\Phi$, 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 $\Phi$ 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 $x \in X^D \setminus X^H$, then $x$ is desired by $x_H$ but not by $x_D$. If $x \in X^H \setminus X^D$, then $x$ is desired by $x_D$ but not by $x_H$. Finally if $x \in X \setminus (X^D \cup X^H)$, then $x$ is desired by neither $x_D$ nor $x_H$.

When all agents’ preferences are substitutable, the operator $\Phi$ is *isotone* in the sense that if $X^D \subseteq \bar{X}^D$ and $X^H \supseteq \bar{X}^H$, then $\Phi_D(X^H) \subseteq \Phi_D(\bar{X}^H)$ and $\Phi_H(X^D) \supseteq \Phi_H(\bar{X}^D)$. Hence, by Tarski’s fixed-point theorem, there exists a nonempty lattice of fixed points of $\Phi$. 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 $d \in D$. 18
Theorem 1. If all agents’ preferences are substitutable, 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 $\tilde{A}$, if $A$ is preferred by all the doctors to $\tilde{A}$ (i.e. $A \succeq_D \tilde{A}$), then all the hospitals prefer $\tilde{A}$ to $A$. In particular, doctor-optimal and doctor-pessimal stable outcomes exist, and they are the hospital-pessimal and hospital-optimal stable outcomes, respectively.\(^{22}\)

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) and Hatfield and Kominers (2014)).\(^{23}\) 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 $F \setminus \{f\}$ such that no (many-to-one) stable outcome exists.\(^{24}\)

If the preferences of a hospital $h$ are not substitutable, then there exist contracts $x, z \in X$ and a set of contracts $Y \subseteq X$ (with $z \notin Y$) such that $z \notin C_h(Y \cup \{z\})$ and $z \in C_h(\{x\} \cup Y \cup \{z\})$. 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, $x_D$ and $z_D$ are taken to have opposing preferences over contracts with hospitals

---

\(^{22}\)Analogous opposition of interests results have been identified in most matching settings, including those of Roth (1984b), Hatfield and Milgrom (2005), and Echenique and Oviedo (2006).

\(^{23}\)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 (see Hatfield and Kojima (2008)).

\(^{24}\)Our maximal domain result is stronger than the specialization of the analogous result of Hatfield and Kominers (2012) to our setting.
$h$ and $h'$ (with $x_D$ preferring $h'$). Hospital $h'$ prefers $z_D$, so that whenever $z_D$ accepts a contract with $h'$, he blocks $x_D$ from doing so; 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 $h$. The intuition 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 this argument, 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.

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 many outcomes are pairwise stable that we would not expect to be stable in practice. For example, consider the following preferences:

$$P_h : \{x, z\} \succ \emptyset,$$
$$P_{x_D} : \{x'\} \succ \{x\} \succ \emptyset,$$
$$P_{h'} : \{z'\} \succ \{x'\} \succ \emptyset,$$
$$P_{z_D} : \{z\} \succ \{z'\} \succ \emptyset.$$

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, as a deviation to $\{x, z\}$ seems quite likely. And indeed, there are no (many-to-one) stable outcomes for these preferences.

## 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., Kojima (2007a), Klaus et al. (2007), Klaus and Klijn (2007)). The answer to this question is of practical importance for real world applications.
such as the NRMP \cite{Roth-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 essentially 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.

4.2 The Structure of the Set of Stable Outcomes

The preferences of $f$ satisfy the law of aggregate demand if for all $X'' \subseteq X' \subseteq X$, $|C_f(X'')| \leq |C_f(X')|$.\footnote{This condition was introduced by Hatfield and Milgrom \cite{Hatfield-2005}. Alkan and Gale \cite{Alkan-2003} introduced a related condition called “size monotonicity.”} Under this condition, we obtain an analogue of the rural hospitals theorem of Roth \cite{Roth-1984a}: each agent signs the same number of contracts at every stable outcome (see Appendix B). 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.\footnote{To see this, suppose that preferences are given by

\begin{align*}
P_h : \{x^{(20)}\} & \succ \{x^{(40)}\} \succ \emptyset, \\
P_d : \{x^{(40)}\} & \succ \{x^{(20)}\} \succ \emptyset.
\end{align*}

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

21
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 (2007b); 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, Theorems 5.10 and 5.14 of Roth and Sotomayor (1990) provide an example where the unique stable outcome is not weakly Pareto optimal for the hospitals, and furthermore one hospital has an incentive to misstate its preferences to a mechanism choosing the hospital-optimal stable outcome.27

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.28

4.3 Strong Substitutability and Strong Stability

Echenique and Oviedo (2006) introduced a condition, strong substitutability, that is more

27However, when every agent’s preferences satisfy the law of aggregate demand, for the set of doctors who exhibit unit demand, the doctor-optimal stable mechanism is (group) strategy-proof and the doctor-optimal stable outcome is weakly Pareto optimal (see Hatfield and Kominers (2012)).

28This 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.
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 notion of stability 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 $Z \subseteq X \setminus A$, such that 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 Echenique and Oviedo (2006), a similar concept which imposes the additional requirement that the deviating agents agree on which contracts in the original outcome $A$ to drop, i.e., for all $y \in A$, $y \in Y^{yd}$ if and only if $y \in Y^{yh}$.

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.\footnote{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 a pair of agents.}
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
- 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. Consider the setting where \( D = \{i, j, k\}, H = \{h\} \), and \( X = \{x, y, z\} \) where \( x_D = i, y_D = j \), and \( z_D = k \). 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 \( X = \{x, y, z\} \) where \( x_D = \hat{x}_D = x'_D = i \) and \( y_D = j \). Let the preferences of the agents be
given by

\[ P_h : \{\hat{x}, y\} \succ \{\hat{x}, x\} \succ \{y\} \succ \{x\}, \]

\[ P_{h'} : \{x'\}, \]

\[ P_i : \{x', x\} \succ \{\hat{x}, x\} \succ \{x\} \succ \{x'\} \succ \{\hat{x}\}, \]

\[ P_j : \{y\}. \]

Here, only the preferences of \( h \) and \( i \) are not strongly substitutable, but the only stable outcome—\( \{x', y\} \)—is not strongly unblocked.\(^{30}\)

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 this framework, we have shown that substitutable preferences are sufficient and necessary (in the maximal domain sense) for the existence of stable outcomes.

In related work (Hatfield and Kominers, 2014), we apply the results obtained here in the context of many-to-one matching with contracts: We identify a class of preferences which 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 class of preferences is relevant for a broad array of applications including military cadet–branch matching (Sönmez and Switzer, 2013; Sönmez, 2013) and the design of affirmative action mechanisms (Kominers and Sönmez, 2013).

\(^{30}\)To see this, take \( Z = \{\hat{x}, x\} \) in the definition of strong unblockedness.
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. Optimal selection of the contract language depends upon application-specific parameters which the market designer must assess; hence, our work leaves substantial room for market design.

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\(^{31}\) 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.\(^{32}\) Nevertheless, we believe that a centralized matching mechanism is likely to see continued participation if its contractual language is both expressive and guarantees the existence of stable outcomes.

Our work 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.\(^{33}\) Of course, these issues are magnified when more complicated complementarities in preferences are present, as in the case of combinatorial package auctions (Ausubel and Milgrom, 2002; Milgrom, 2004; Kwasnica et al.,

\(^{31}\text{Roth (1984a, 1991), and Roth and Xing (1994) provided empirical evidence that the stability of the outcome recommended by a centralized match is essential to the long-run success of the matching system.}\)

\(^{32}\text{Clearly, there is no a priori reason why those agents should deal within the match’s contract language.}\)

\(^{33}\text{Recent work by Kojima et al. (2013) and Ashlagi et al. (2014) has argued that large-market effects may explain this phenomenon.}\)
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.
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 $U$ which represents these preferences. Clearly, $U(Y) = U(Y \cup \{z\})$ and $U(\{x\} \cup Y \cup \{z\}) > U(\{x\} \cup Y)$, and so

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

hence, $U$ 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

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

Let $U(Z^n) = 1 - 2^{-n}$. 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 $U(\{x\} \cup Z) - U(Z) \geq 2^{-n-1} \geq 2^{-n'-1} \geq U(\{x\} \cup Y) - U(Y)$ and so

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

so that $U$ is submodular.

Proof of Proposition 2

We prove a lemma which directly implies the result.

---

34A 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.
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

$$Z \subseteq C_H(Y \cup Z) \quad (1)$$

by definition.

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

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 $Z' \subseteq X'$. Since $X \triangleright X'$, we have $Z' \subseteq X' \subseteq X$. But $Z' \not\subseteq Y'$ and by construction we must have $Z'_j \subseteq C^X_f(Z' \cup Y')$ for each $f \in F$, contradicting the stability of $Y$ with respect to $X$.

Proof of Proposition 4

If $P^X_f$ is not substitutable for some $f \in F$, then there exist $z, x \in X'$ and $Y \subseteq X'$ such that $z \not\in C^X_f(Y \cup \{z\})$ but $z \in C^X_f(Y \cup \{z, x\})$. But $P^X_f$ is just the restriction of $P^X_f$ 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 $P^X_f$.

\[\text{\footnote{It is clear that } Y' \text{ is individually rational, so } Y' \text{ can only be unstable if it is blocked.}}\]
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.\(^{36}\)

**Definition 6.** A choice function \(C^f\) satisfies the irrelevance of rejected contracts condition if for all \(Y \subseteq X\) and \(z \in X \setminus Y\), 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 \(\Phi\), 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

\[x \in C_D(X^D \cup \{x\}) = C_D(X^D).\] (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 \(\Phi\), 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 \(x \in (C_H(X^H \cup \{z\})) \cap (X^H \setminus X^D)\) 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 \(\Phi\)—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 \(z \in (X^D \cap X^H) \setminus A\), then we

\(^{36}\)Our choice functions satisfy the irrelevance of rejected contracts condition because they are induced by strict preference relations (see Aygün and Sönmez (2014a,b)).
have \( z \in C_D(A \cup \{ z \}) \) and \( z \in C_H(A \cup \{ z \}) \) (by the definition of \( \Phi \)). 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 \( \Phi \). 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 \( y \in \Phi_D(X_H) \setminus X_D \). 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 \( y \in X_D \setminus \Phi_D(X_H) \). Then \( y \in C_H(A \cup \{ y \}) \); hence, if \( y \notin \Phi_D(X_H) \), there exists a \( z \in X_H \setminus A \) 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 \( C_D(\tilde{X}_D) \subseteq A \), then \( A \) is not individually rational, contradicting the stability of \( A \). If \( y \in C_D(\tilde{X}_D) \setminus A \), then \( y \in \tilde{X}_H = \Phi_H(\tilde{X}_D) \) (as \( (\tilde{X}_D, \tilde{X}_H) \) is a fixed point of \( \Phi \)), and so \( y \in \tilde{X}_H \cap \tilde{X}_D \) — a contradiction.

Now, \( C_D(\tilde{X}_D) = A \), we have that \( (\tilde{X}_D \setminus A) \cap C_D(\tilde{X}_D) = \emptyset \). By substitutability, \( (\tilde{X}_D \setminus A) \cap C_D(\tilde{X}_D \cup \{ x \}) = \emptyset \) 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 \( (\tilde{X}_D, \tilde{X}_H) \neq (X_D, X_H) \).

**Proof of Theorem 1**

We first verify that the operator \( \Phi \) 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 \(\Phi\) is isotone.

As \(\Phi\) 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 \(\Phi\). 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 \(z \in C_D(X^D) \setminus A\). 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 \(\Phi\), 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 \(z \in A \setminus C_D(X^D)\). 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 \(\Phi\).

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 \(A = C_D(X^D) \triangleright_D C_D(\tilde{X}^D) = \tilde{A}\). 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 \(\triangleright_D\).
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 $Y \subseteq X \setminus \{x, z\}$ such that $Y_H = \{h\}$ and

$$z \notin C_h(Y \cup \{z\})$$
$$z \in C_h(\{x\} \cup Y \cup \{z\}).$$

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. Let $x_D$ have preferences such that

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

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. Let $h'$ have preferences such that

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

Let every doctor $d \in D \setminus \{x_D, z_D\}$ have preferences such that

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

Consider any outcome $A$; we will 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

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

where $\hat{C}_d(W)$ is the responsive choice function over $\{x, z, x', x''\}$ with quota 2 and underlying preference order $x'' \succ z \succ x \succ x'$. We let $\hat{d}$ have preferences such that

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

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. We let $h'$ have preferences such that

$$C_{h'}(W) = \begin{cases} 
\{x'\} \quad x' \in W \\
\{\hat{x}'\} \quad \hat{x}' \in W \text{ and } x' \notin W \\
\emptyset \quad \text{otherwise.}
\end{cases}$$
We let $h''$ have preferences such that
\[
C_{h'}(W) = \begin{cases} 
\{\hat{x}''\} & \hat{x}'' \in W \\
\{x''\} & x'' \in W \text{ and } \hat{x}'' \notin W \\
\emptyset & \text{otherwise.} 
\end{cases}
\]

Finally, let every doctor $\bar{d} \in D \setminus \{d, \hat{d}\}$ have preferences such that
\[
C_{\bar{d}}(W) = (W \cap Y)_{\bar{d}}.
\]

Consider any outcome $A$; we will 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 $C_h(Y \cup \{z\})$.

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 $C_h(\{x\} \cup Y \cup \{z\}) \succ_h A_h \succ_h C_h(\{x\} \cup Y \cup \{z\}) \setminus \{x\}$ 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 \( x \in Z \setminus A \). 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 \( x \in ((A \cup \{x\}) \cap C_{x_H}(Y^{x_H})) \setminus C_{x_H}(A \cup \{x\}) \), so the preferences of \( x_H \) are not strongly substitutable.

\section*{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).

\textbf{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.

\textit{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 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.

This theorem is an immediate and elegant 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. 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 $\text{NE}(\cdot)$ for the Nash equilibrium correspondence. A mechanism $o$ is said to *Nash implement solution* $\varphi$ if, for all possible profiles $P$,

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

That is, $o$ Nash implements $\varphi$ 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.

---

37 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.
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.\(^\text{38}\)

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).

\(^{38}\text{In fact, the argument follows that of Haake and Klaus (2009a) directly, but is slightly simpler in our framework. Specifically, the first subargument of Step 3 in the proof given by Haake and Klaus (2009a) can be omitted, since in our framework both doctors and hospitals may accept multiple contracts.}\)
References


