# Complementarities and Externalities

Thành Nguyen[*] and Rakesh Vohra[†]

February 24, 2022

**Abstract**

Stable matchings need not exist when one side has preferences over the side that exhibits complementarities. This chapter summarizes some of the approaches taken emphasizing the usefulness of Scarf's lemma.

[*]Krannert School of Management, Purdue University, nguye161@purdue.edu
[†]University of Pennsylvania, rvohra@seas.upenn.edu

## Contents

# 1 Introduction

David Gale and Lloyd Shapley formulated the problem of finding a stable matching and identified a setting where such matchings always exist via the deferred acceptance algorithm (DA) (see **??**). The algorithm is a thing of beauty and no one exposed to it is immune to its charms. It has inspired others to enlarge domains in which the DA algorithm is applicable.

However, in the presence of preference complementarity, stable matchings are not guaranteed to exist (see Section 3 for an example). This is a problem because there are many settings of practical importance in which agents' preferences exhibit complementarities. In school choice, for example, this can arise in two ways. First, parents with multiple children have preferences over *subsets* of schools their children are assigned to rather than individual schools¿ Second, schools themselves may have preferences over the distribution of characteristics of the students they admit.

Furthermore, even if a stable matching does exist, the DA algorithm is not guaranteed to find it. This is because at each iteration of the algorithm when a 'proposing' agent is rejected, the decision is irrevocable. This excludes the possibility of accommodating complementarities in preferences on the accepting side. A proposing agent by themselves may be ranked individually below other agents but in concert with another agent may be ranked high.

To overcome this problem, we present Scarf's lemma which yields an alternative proof for the existence of stable matchings. While Scarf's proof lacks some of the desirable features of the DA, it is much more powerful in an entirely different way: it gives us the tools to incorporate complementarities and externalities while preserving the guaranteed existence of a stable solution.

The question of the existence of a stable matching is closely related to that of non-emptiness of the core of a nontransferable utility cooperative game. A key tool in the study of nontransferable utility cooperative games is Scarf's lemma. It provides a sufficient condition for the non-emptiness of the core of an NTU co-operative game. The lemma makes three important contributions. The first is to frame the problem of finding a matching as a special case of a more general problem about coalition formation. The second is an extension of the notion of stability to 'fractional' assignments of agents to coalitions called domination. The third is establishing the existence of 'stable' fractional assignments. The connection to stability has been neglected, perhaps, because the study of NTU co-operative games fell out of fashion. This is unfortunate. To paraphrase Scarf from another context

> Our message boils down to a simple straightforward piece of advice; if economists are to study stability, the first step is to take our trusty DA algorithm, pack it up carefully in mothballs, and put it away respectfully; it has served us well for many a year. In the presence of complementarities, it simply doesn't do the job it was meant to do.

When complementarities are present, stable matching need not exist (see

Section 3). Further, pairwise stability does not always imply group stability as discussed in **??**. Pairwise stable matchings need not be Pareto-optimal, but group stable matchings are.

This chapter will survey proposals to surmount the problem of non-existence as well as identifying group stable matchings. It emphasizes the role of Scarf's lemma. So as to strike a balance between the perspective of the parachutist and that of the truffle hunter, we restrict our discussion to two-sided many-to-one matching problems.

# 2 Existence of Stable Matching, revisited

## 2.1 Scarf's Lemma

In this section, we state a version of the lemma that is adapted to the matching context. We show how it implies the existence of a stable matching in the setting originally considered by Gale and Shapley.

Let $\mathcal{A}$ be an $n \times m$ non-negative matrix with at least one non-zero entry in each row and $q \in \mathbb{R}^n_+$. Associated with each row $i \in \{1, .., n\}$ of $\mathcal{A}$ is a strict order $\succ_i$ over the set of columns $j$ for which $\mathcal{A}_{i,j} > 0$.

To interpret, suppose for a moment, that $\mathcal{A}$ is a 0-1 matrix. Associate each row of $\mathcal{A}$ with an agent and interpret each column to be the characteristic vector of a coalition of agents. Hence, $\mathcal{A}_{ij} = 1$ means that agent $i$ is in the $j^{th}$ coalition. Then, $\succ_i$ can be interpreted as agent $i$'s preference ordering over all the columns/coalitions of $\mathcal{A}$ of which $i$ is a member. No restrictions on $\succ_i$ are imposed.

Consider the system $\{x \in \mathbb{R}^m_+ : \mathcal{A}x \leq q\}$. The positive elements of a feasible $x$ can be interpreted as describing the feasible coalitions that can be formed. We illustrate with an example.

**Example 2.1.** Consider an instance of the stable matching problem that consists of two hospitals $(h_1, h_2)$, each with capacity 1, and two single doctors $(d_1, d_2)$. This is the setting of Gale and Shapley [1962]. The preferences are as follows: $d_1 \succ_{h_1} d_2; d_1 \succ_{h_2} d_2;\ h_2 \succ_{d_1} h_1; h_2 \succ_{d_2} h_1$.

We describe the set of feasible matchings as the solution to a system of inequalities. The constraint matrix of this system will be the matrix $\mathcal{A}$ that is used when invoking Scarf's lemma.

Introduce variables $x_{(d_i,h_j)} \in \{0, 1\}$ for $i \in \{1, 2\}; j \in \{1, 2\}$ where $x_{(d_i,h_j)} = 1$ if and only if $d_i$ is assigned to $h_j$ and zero otherwise. In the $4 \times 4$ matrix, $\mathcal{A}$, below, each row corresponds to an agent (a hospital or a doctor), and each column corresponds to a doctor-hospital pair. An entry $\mathcal{A}_{ij}$ of the matrix $\mathcal{A}$ is 1 if and only if the agent corresponding to row $i$ is a member of the coalition corresponding to column $j$. Otherwise, $\mathcal{A}_{ij} = 0$. $\mathcal{A}x \leq q$ models the capacity constraints of the hospital and the constraints that each doctor can be assigned to at most one hospital. In this example $q = \mathbf{1}$. For each row $i$ of $\mathcal{A}$, the strict order on the set of columns $j$ for which $\mathcal{A}_{ij} \neq 0$ is the same as the preference ordering of agent $i$. Specifically, we have the following system:

$$\begin{pmatrix} & (d_1,h_1) & (d_1,h_2) & (d_2,h_1) & (d_2,h_2) \\ h_1 & 1 & 0 & 1 & 0 \\ h_2 & 0 & 1 & 0 & 1 \\ d_1 & 1 & 1 & 0 & 0 \\ d_2 & 0 & 0 & 1 & 1 \end{pmatrix} \cdot x \le \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} ; \quad order: \begin{array}{c} col_1 \succ_{h_1} col_3 \\ col_2 \succ_{h_2} col_4 \\ col_2 \succ_{d_1} col_1 \\ col_3 \succ_{d_2} col_4. \end{array}$$

We now define the notion of domination.

**Definition 1.** A vector $x \ge 0$ satisfying $\mathcal{A}x \le q$ **dominates** column $k$ of $\mathcal{A}$ if there exists a row $i$ such that $\sum_{j=1}^{n} \mathcal{A}_{ij} x_j = q_i$ and for all column $l \in \{1, .., m\}$ such that $\mathcal{A}_{i,l} > 0$ and $x_l > 0$, $l \succeq_i k$.

The next example will illustrate that when a dominating vector is $x$ is integral, this is precisely the notion of stability.

**Example 2.2.** Recall example 2.1. Every integer solution to $\mathcal{A}x \le \mathbf{1}$ corresponds to a matching and vice versa. Notice, $x = (1,0,0,1)^T$ corresponds to the matching $(d_1,h_1);(d_2,h_2)$. It is not stable because it is blocked by $(d_1,h_2)$. In the language of Scarf's lemma, $x = (1,0,0,1)^T$ is not a dominating solution because $x$ does not dominate the column corresponding to $(d_1,h_2)$. The solution $x = (0,1,1,0)^T$ is a dominating solution and corresponds to a stable matching.

**Lemma 2.1** (Scarf [1967])**.** *Let $\mathcal{A}$ be an $n \times m$ non-negative matrix and $q \in \mathbb{R}^n_+$. Then, there exists an extreme point of $\{x \in \mathbb{R}^m_+ : \mathcal{A}x \le q\}$ that dominates every column of $\mathcal{A}$.*

*Proof.* The proof is by reduction to the existence of a Nash equilibrium in a two-person game. By scaling we can assume that $q$ is the vector of all 1's, denoted $\mathbf{1}$. Let $\mathcal{C}$ be a matrix defined as follows:

- If $\mathcal{A}_{ij} = 0$, then $\mathcal{C}_{ij} = 0$.
- If $\mathcal{A}_{ij} > 0$ and $j$ is ranked $t$-th in the preference list of $i$, then $\mathcal{C}_{ij} = -N^t$, for $N \ge n$.

The $j^{th}$ columns of $\mathcal{A}$ and $\mathcal{C}$ will be denoted $\mathcal{A}^j$ and $\mathcal{C}^j$ respectively.

We associate a 2-person game with the pair $(\mathcal{A}, \mathcal{C})$. The payoff matrix for ROW will be $\mathcal{A}$. The payoff matrix for COLUMN will be $\mathcal{C}$. Let $(x^*, y^*)$ be an equilibrium pair of mixed strategies for this game where $x^*$ is a mixed strategy for ROW player (payoff matrix $\mathcal{A}$) and $y^*$ is the mixed strategy for COLUMN (payoff matrix $\mathcal{C}$).

Let $R^*$ be ROW's expected payoff and $C^*$ be the expected payoff to COLUMN. Clearly $\mathcal{A}x^* \le R^*\mathbf{1}$. We show that $\frac{x^*}{R^*}$ is a dominating solution.

Suppose the columns of $\mathcal{A}$ are sorted so that $x^*_i > 0$ for $1 \le i \le k$ and $x^*_i = 0$ for $k+1 \le i \le n$. COLUMN's expected payoff when playing each column $1, \ldots, k$ is exactly $C^*$, and expected payoff when playing any other column is at most $C^*$. If COLUMN plays one of the first $k$ columns uniformly at random, her expected payoff is also $C^*$, i.e.,

$$\frac{1}{k}y^* \cdot (\mathcal{C}^1 + \ldots + \mathcal{C}^k) = \sum_i y_i^* \frac{(\mathcal{C}_{i1} + \mathcal{C}_{i2} + \ldots + \mathcal{C}_{ik})}{k} = C^*.$$

Choose any column $j$. We show that $x^*$ dominates $j$. As $x^*$ is a best response to $y^*$ it must be that

$$\sum_i y_i^* \mathcal{C}_{ij} \leq C^* = \sum_i y_i^* \frac{(\mathcal{C}_{i1} + \mathcal{C}_{i2} + \ldots + \mathcal{C}_{ik})}{k}.$$

$$\Rightarrow \ \exists y_r^* > 0 \ \text{ s.t. } \ \mathcal{C}_{rj} \leq \frac{(\mathcal{C}_{r1} + \mathcal{C}_{r2} + \ldots + \mathcal{C}_{rk})}{k}.$$

Hence,

$$y_r^* > 0 \ \Rightarrow \ \mathcal{A}_{r1}x_1^* + \ldots + \mathcal{A}_{rk}x_k^* = R^* > 0.$$

Therefore, at least one of $\mathcal{A}_{r1}, \ldots, \mathcal{A}_{rk}$ is non-zero. Hence, at least one of $\mathcal{C}_{r1}, \ldots, \mathcal{C}_{rk}$ is non-zero and $\mathcal{C}_{ij} \neq 0 \ \Rightarrow \mathcal{A}_{ij} \neq 0$.

Assume that among the columns $1, \ldots, k$, that $\mathcal{A}_{r1} \neq 0$ and column 1 is the least preferred. Let $\ell$ be the rank of column 1 in that preference ordering. Recall $\mathcal{C}_{r1} = -N^{\ell}$. Hence,

$$\frac{(\mathcal{C}_{r1} + \mathcal{C}_{r2} + \ldots + \mathcal{C}_{rk})}{k} \leq \frac{\mathcal{C}_{r1}}{k} = \frac{-N^l}{k} < -N^{l-1}.$$

By definition, $N > n \geq k$. Therefore, $\mathcal{C}_{rj} < -N^{\ell-1}$. This shows first, that $\mathcal{C}_{rj} < 0$ and thus $\mathcal{A}_{rj} \neq 0$, and $j$ is in the preference list of row $i$; second, the ranking of $j$ is below column 1. In other words, row $i$ does not prefer $j$ to any of the columns 1 to $k$. $\square$

Scarf himself gave a finite time algorithm for finding a dominating extreme point, however, the problem of finding a dominating solution is PPAD complete. Thus, it has a worst-case complexity equivalent to that of computing a fixed point, but this is not a barrier to implementation. For example, building on Budish [2011], a course allocation scheme that relies on a fixed-point computation has been proposed and implemented at the Wharton School.

**Example 2.3.** Continuing with example 2.2, we see by the Birkhoff-von Neumann theorem, that every non-negative extreme point of the system $\mathcal{A}x \leq \mathbf{1}$ is integral. Therefore, it follows by Scarf's lemma that a stable matching exists. The same theorem shows that the conclusion generalizes to more than two single doctors and more than two unit-capacity hospitals.

We draw the reader's attention to three features of Scarf's lemma. First, it does not distinguish between two-sided vs. multi-lateral settings. Any family of feasible coalitions of agents that can be expressed using linear inequalities with non-negative coefficients is acceptable. Second, the notion of domination adjusts to the set of coalitions encoded in the columns of the $\mathcal{A}$. In this way, one encodes *multilateral* stability and not just pairwise stability (see chapter **?**). Third, it assumes that preferences can be represented as a strict ordering over coalitions. Allowing for indifference simply raises familiar issues about what constitutes a blocking coalition. Must every member be strictly better off or does it suffice for no one to be worse off and at least one member to be strictly better off (see Roth and Postlewaite [1977])? The simplest resolution is to allow for indifferences. To see the difference suppose in definition 1, column $j$ is such that $x_j = 0$. Then, the corresponding row $i$ may be indifferent between column $j$ and all columns $k$ such that $\mathcal{A}_k x_k > 0$. It would be natural in this case to call $x$ a weakly dominating solution. This is what is used in Section 4.2. One can also handle indifferences via a lexico-graphic tie-breaking rule. That is, if a row is indifferent between two columns, then it breaks ties using the preferences of another row. See Section 3.1 for an example.

In many matching settings, it is more common to represent the preferences of at least one side via choice functions rather than a preference ordering. Section 3.1 gives one example of how this can be handled by Scarf's lemma. Section 4.2 gives another.

In the subsequent sections, we discuss applications of Scarf's lemma to matching settings with complementarities. We begin with the problem of matching with couples. It is the simplest and most well-known instance of preference complementarity in two-sided matching. It will illustrate two things. First how the lemma may be deployed even when the underlying preferences are not described by orderings. Second, how the lemma can be used even in a setting where a stable matching is not guaranteed to exist.

## 2.2 Rounding

There is no guarantee that a dominating solution will be integral. When a dominating solution is fractional, it is natural to 'round' it into an integer solution while preserving domination. There is a price to be paid because the rounded solution may be infeasible. For the moment, set this aside. The following result shows that if we round the fractional dominating solution in the right way, then domination is maintained.

**Theorem 2.2.** *Given a dominating solution $x$ of $\{\mathcal{A}x \leq q, x \geq 0\}$. Let $\bar{x} \geq 0$ be integral solution such that if $x_i = 0$, then $\bar{x}_i = 0$. Let $\bar{q} := \mathcal{A}\bar{x}$, then $\bar{x}$ is an integral dominating solution of $\{\mathcal{A}x \leq \bar{q}, x \geq 0\}$.*

In the applications discussed below, the system $\{\mathcal{A}x \leq q, x \geq 0\}$ can be expressed as the intersection of two systems of linear inequalities:

$$Ax \leq b \tag{1}$$

$$Cx \leq d \tag{2}$$

$$0 \leq x \leq e \tag{3}$$

Here $e$ is the vector of all 1's.

The constraints in (1) are 'hard' in that they cannot be violated. They correspond to the constraints of the unit-demand side. The constraints in (2) are 'soft', i.e., we allow them to be violated but not by too much.

By adding slack variables, we can always assume that (1) and (2) hold at equality, i.e. $Ax = b$ and $Cx = d$. Given a fractional extreme point solution, $x^*$, to (1-3) we can use the iterative rounding technique to round it into a 0-1 vector $\bar{x}$, such that $\{A\bar{x} = b,\ C\bar{x} = d + \delta d\}$, where $||\delta d||_\infty$ is bounded. Here $||\cdot||_\infty$ refers to the $\ell$ infinity norm. The precise value of $||\delta d||_\infty$ will depend on the proportion of non-zero entries in the matrix $C$ (its sparsity).[1]

We describe one iteration of the technique. Let $\mathcal{Q} = \begin{bmatrix} A \\ C \end{bmatrix}$ and $q = \begin{bmatrix} b \\ d \end{bmatrix}$. Let $J$ be the index set of the *non-integral* components of $x^*$ and $J^c$ its complements. The vector $x^*$ restricted to these components is denoted $x^*_J$. Hence, every component of $x^*_J$ is positive. The components corresponding to $J^c$ are already integer valued and so we leave these untouched. Our focus is on rounding the components in $J$.

Denote by $\mathcal{Q}|_J$, the submatrix of $\mathcal{Q}$ consisting of the columns indexed by $J$. As $x^*$ is an extreme point, the number of non-integral components of $x^*$, $|J|$, is equal to the maximum number of *linearly independent* rows of $\mathcal{Q}|_J$.

Now, let $A^*, C^*$ be the submatrices of $A$ and $C$, that correspond to the constraints that bind at $x^*$, i.e. $\mathcal{Q}|_J = \begin{bmatrix} \mathcal{A}^* \\ \mathcal{C}^* \end{bmatrix}$. Similarly define $b^*$ an $d^*$. Hence, $x^*_J$ satisfies:

$$\mathcal{Q}|_J x^*_J = \begin{bmatrix} b^* \\ d^* \end{bmatrix} - \mathcal{Q}|_{J^c} x^*_{J^c} = q^*. \tag{4}$$

Notice, the components corresponding to $J^c$ have been taken over to the right hand side consistent with the idea that we are leaving them untouched.

By the extreme point property of $x^*$, the matrix $\mathcal{Q}|_J$ is invertible and so of full rank. Suppose we delete a row, say row $i$, from the matrix $C$. Call the reduced matrix $\mathcal{Q}|_J[-i]$ and denote the corresponding right-hand side by $q^*[-i]$. Notice $\mathcal{Q}|_J[-i]x^*_J = q^*[-i]$. The null space of $\mathcal{Q}|_J[-i]$ is one dimensional. Let $z \neq 0$ be a vector in the null space and consider $x^*_J + \epsilon z$ for some $\epsilon \neq 0$. Then, $\mathcal{Q}|_J[-i](x^*_J + \epsilon z) = q^*[-i]$. Hence, the vector $z$ represents a a direction in which we can move $x^*_J$ so as to satisfy all constraints except the one that was removed.

For a suitable choice of $\epsilon$ we can ensure that $e \geq x^*_J + \epsilon z \geq 0$. In fact, by making $|\epsilon|$ sufficiently large we can maintain this property and ensure that $x^*_J + \epsilon z$ has at least one more integral component than $x^*_J$. By executing these steps iteratively, we will arrive eventually at a vector with all integral components.

Now, $x^*_J + \epsilon z$ will violate exactly one constraint, the one corresponding to the $i^{th}$ row of $C$ that was deleted. If $C_i$ denotes this row, the magnitude of this violation will scale with $C_i \cdot \epsilon z$, a quantity bounded by the number of non-zero entries in $C_i$. Hence, when choosing a row to delete, we look for one with a

[1] Statements of this kind can be viewed as a refinement of the Shapley-Folkman-Starr lemma Starr [1969].

small number of non-zero entries. The existence of such a row will follow from the sparsity of $C$. A matrix possessing a small proportion of non-zero entries yet having a large number of non-zero entries in each row is clearly impossible.

# 3 Couples Matching

We describe a version of the problem that is studied, for example, in Roth [1984]. Let $H$ be the set of hospitals, $D^1$ the set of single doctors, and $D^2$ the set of couples. Each couple $c \in D^2$ is denoted by $c = (f, m)$. For each couple $c \in D^2$ we denote by $f_c$ and $m_c$ the first and second member of $c$. The set of all doctors, $D$ is given by $D^1 \cup \{m_c | c \in D^2\} \cup \{f_c | c \in D^2\}$.

Each single doctor $s \in D^1$ has a strict preference relation $\succ_s$ over $H \cup \{\emptyset\}$ where $\emptyset$ denotes the outside option for each doctor. Each couple $c \in D^2$ has a strict preference relation $\succ_c$ over $H \cup \{\emptyset\} \times H \cup \{\emptyset\}$, i.e., over pairs of hospitals including the outside option. The need for ordered pairs arises because couples will have preferences over which member is assigned to which hospital. While a couple consists of two agents they should be thought of as a single agent with preferences over ordered pairs of slots. This is the source of complementarities.

A hospital or an ordered pair of hospitals is acceptable to a single doctor or a couple if they are ranked above the outside option in the doctor's and couple's preferences, respectively.

Each hospital $h \in H$ has a fixed capacity $k_h > 0$. Unlike the doctors, a hospital's preferences are not characterized by an ordering over subsets of $D$ but by a choice function: $Ch_h(.) : 2^D \rightarrow 2^D$. While a choice function can be associated with every strict preference ordering over subsets of $D$, the converse is not true. The information contained in a choice function is sufficient to recover a partial order, only, over the subsets of $D$.

Representing the preferences of one side using a choice function, while, popular is a *modeling choice*. Sometimes it is motivated by a desire to adapt the DA algorithm to a many-to-one setting. In other cases, it can be justified on the grounds that the relevant side is unable to articulate a preference ordering over all subsets but can provide a rule for selection instead. We will return to this matter later in the chapter.

We assume, as is standard in the literature, that $Ch_h(.)$ is responsive (see **?**) which means that $h$ has a strict priority ordering $\succ_h$ over elements of $D \cup \{\emptyset\}$. If $\emptyset \succ_h d$, we say $d$ is not acceptable to $h$. For any set $D^* \subset D$, hospital $h$'s choice from that subset, $Ch_h(D^*)$, consists of the (up to) $k_h$ highest priority doctors among the acceptable doctors in $D^*$. Formally, $d \in Ch_h(D^*)$ if and only if $d \in D^*; d \succ_h \emptyset$ and there exists no set $D' \subset D^* \setminus \{d\}$, such that $|D'| = k_h$ and $d' \succ_h d$ for all $d' \in D'$. Notice, responsiveness rules out preference complementarity on the hospital side.

A matching $\mu$ is an assignment of every single doctor to a hospital or his/her outside option, an assignment of couples to at most two positions (in the same or different hospitals) or their outside option, such that the total number of doctors assigned to any hospital $h$ does not exceed its capacity $k_h$. A matching satisfies individual rationality if all hospitals receive only acceptable doctors, and all doctors and couples are assigned to acceptable choices.

A matching $\mu$ can be 'blocked' in three different ways. First, by a pair $(d, h)$ such that $d \in D^1$ prefers $h$ to $\mu(d)$ and $h$ would select $d$ possibly over a doctor currently assigned to it. Second, by a couple, $c \in D^2$ and a hospital $h$ such that the couple prefers to be assigned to $h$ over their current assignments and $h$ would accept them, possibly over some of its current assignments. Third, by a couple and two distinct hospitals. In this case, the couple would prefer to be assigned to the two hospitals (one to each) over their current assignment and each of the hospitals would accept a member of the couple over at least one of their current assignment. A matching $\mu$ is **stable with respect to a capacity vector** $k$ if $\mu$ is individually rational and cannot be blocked in any of the three ways just described.

The example below, due to Klaus and Klijn [2005] shows that in the presence of couples a stable matching need not exist.

**Example 3.1.** Suppose two hospitals $h_1$ and $h_2$ and three doctors $\{d_1, d_2, d_3\}$. Doctors $\{d_1, d_2\}$ are a couple while $d_3$ is a single doctor. The capacity of each hospital is 1 and the priority ordering of $h_1$ is

$$d_1 \succ_{h_1} \succ d_3 \succ_{h_1} \emptyset \succ_{h_1} d_2.$$

The priority ordering for hospital $h_2$ is

$$d_3 \succ_{h_2} \succ d_2 \succ_{h_2} \emptyset \succ_{h_2} d_1.$$

The preference ordering of the couple is $(h_1, h_2) \succ_{(d_1,d_2)} \emptyset$, while that of doctor $d_3$ is $h_1 \succ h_2 \succ \emptyset$.

## 3.1 Choice Functions vs Orderings

Examples 2.1, 2.2 and 2.3 suggest how one might approach the couples matching problem. For each single doctor $d$ and hospital $h$, that are mutually acceptable, let $x_{(d,h)} = 1$ if $d$ is assigned to $h$ and 0 otherwise. Similarly, for each couple $c \in D^2$ and distinct $h, h' \in H$, such that $(h, h')$ is acceptable to $c$ and the first and second member of $c$ are acceptable to $h$ and $h'$, respectively, let $x_{(c,h,h')} = 1$ if the first member of $c$ is assigned to $h$ and the second is assigned to $h'$. Let $x_{(c,h,h')} = 0$ otherwise.[2] Finally, for a couple $c$ and a hospital $h$ that are mutually acceptable, let $x_{(c,h,h)} = 1$ if both members of the couple are assigned to hospital $h \in H$ and 0 otherwise. Every 0-1 solution to the following system is a feasible matching and vice versa.

$$\sum_{d \in D^1} x_{(d,h)} + \sum_{c \in D^2} \sum_{h' \neq h} x_{(c,h,h')} + \sum_{c \in D^2} \sum_{h' \neq h} x_{(c,h',h)} + \sum_{c \in D^2} 2x_{(c,h,h)} \leq k_h \ \forall h \in H \tag{5}$$

[2]Note that $x_{(c,h,h')}$ does not represent the same thing as $x_{(c,h',h)}$.

$$\sum_{h\in H} x_{(d,h)} \leq 1 \ \forall d \in D^1 \tag{6}$$

$$\sum_{h,h'\in H} x_{(c,h,h')} \leq 1 \ \forall c \in D^2 \tag{7}$$

Unfortunately, extreme points of this system are not guaranteed to be integer-valued.

Let $\mathcal{A}$ be the matrix whose entries are the coefficients of the system (5-6-7). In (5-6-7), each agent (single doctor, couple and hospital) is represented by a single row. Each column/variable corresponds to a coalition of agents (an assignment of a single doctor to a hospital or a couple to a pair of hospital slots, that are mutually acceptable).

To apply Scarf's lemma we need each of the rows in (5-6-7) to have an ordering over the columns that are in the support of that row. For the rows associated with a single doctor and a couple, we can use their preference ordering over the hospitals (and pairs of hospitals in the case of couples).

For the rows associated with hospitals, this will depend on how the preferences of hospitals are modeled. If they are characterized by an ordering over subsets of doctors it induces an ordering over the columns associated with coalitions involving either a single doctor or a couple and the hospital $h$. Notice, it incorporates preference complementarity on the hospital side. If one is content to model hospital preferences in this way, we can immediately invoke Scarf's lemma to deduce the existence of a *fractional* stable matching.

However, as noted earlier, hospital preferences are more commonly encoded using responsive choice functions rather than preference orderings. We outline how to use each hospital's choice function to induce an ordering over the columns associated with coalitions involving either a single doctor or a couple and that hospital. Formally, we use hospital $h$'s priority ordering, $\succ_h$, to generate an *artificial* ordering $\succ_h^*$ over single doctors and couples with the property that domination with respect $\succ_h^*$ corresponds to stability with respect to the underlying choice function.

Hospital $h$ will order these columns based on its ranking of the corresponding doctors assigned to $h$. If the column corresponds to an assignment of both members of a couple to $h$, then, $h$'s ranking of this column depends on the ranking of the 'worst' member of the couple, as determined by $\succ_h$. Under this ordering, there can be ties between columns that correspond to different ways a couple is assigned. For example, one member is assigned to $h$, while the other is assigned elsewhere. To break the ties between these columns, $h$ uses the preference ordering of the couple. We denote by $\succ_h^*$ this new ordering. We illustrate the construction of $\succ_h^*$ in the following example.

**Example 3.2.** There are two hospitals $h, h'$, one couple $c = (d_1, d_2)$, and a single doctor, $d_3$. Hospital $h$'s priority ordering is $d_1 \succ_h d_3 \succ_h d_2$. We use to $\succ_h$ to induce an ordering $\succ_h^*$ over $x_{(c,h,h')}, x_{(c,h',h)}, x_{(c,h,h)}$, and $x_{(d_3,h)}$. In $x_{(c,h,h')}$, $d_1$ is assigned to $h$; in $x(c, h', h)$, $d_2$ is assigned to $h$; in $x_{(d_3,h)}$, $d_3$ is assigned to $h$, and in $x_{(c,h,h)}$ both $d_1$ and $d_2$ are assigned to $h$. In the induced ordering, hospital $h$ will rank $x_{(c,h,h)}$ based on the member of $c$ which is lower in $h$'s priority order, which is $d_2$. Thus, $x_{(c,h,h')} \succ_h^* x_{(d_3,h)} \succ_h^* x_{(c,h',h)} \sim x_{(c,h,h)}$. The tie between $x_{(c,h',h)}$ and $x_{(c,h,h)}$ is broken based on the preference ordering of $c$. Namely, $x_{(c,h',h)} \succ_h^* x_{(c,h,h)}$ iff $x_{(c,h,h')} \succ_c x_{(c,h,h)}$.

## 3.2 Soft Capacity Constraints

If a stable matching does not exist and one is not prepared to yield on stability, something else must be sacrificed. Nguyen and Vohra [2018] propose relaxing the hospital capacity constraints. While not a universal panacea, there are settings where capacity constraints are 'soft', for example, Correa et al. [2019].

How does the softness of the hospital capacity constraints help? Scarf's lemma gives us a fractional dominating solution to the system (5-6-7). We then apply the iterative rounding procedure to obtain an integer dominating solution. The rounded solution will violate (5) only. For this to be useful the amount by which the constraints (5) are violated should be modest. Informally, for any instance of the stable matching problem with couples, find a stable matching with respect to a 'nearby' instance, which is obtained by altering the initial capacities of the hospitals. The next result shows this to be possible. Subsequently, we outline why such a result is possible.

A matching with respect to a capacity vector $k$ is a 0-1 solution to (5-6-7). It is stable with respect to a capacity vector $k$ if it is dominating.

**Theorem 3.1.** *For any capacity vector $k$, there is a capacity vector $k^*$ and a stable matching with respect to $k^*$, such that* $\max_{h \in H} |k_h - k_h^*| \leq 2$. *Furthermore,* $\sum_{h \in H} k_h \leq \sum_{h \in H} k_h^* \leq \sum_{h \in H} k_h + 4$.

Theorem 3.1 shows that by judicially *redistributing* capacities between hospitals only a small injection of additional capacity is needed to ensure the existence of a stable matching.

If one imposes restrictions on a couple's preferences one can improve these bounds. One such restriction is when couples prefer to be together, rather than apart, and a hospital must accept either both members of the couple or none. Dur et al. [2018] report on a real-world instance where this restriction is imposed by the designer. While this restriction relaxes the problem because each blocking coalition only involves the preferences of a single hospital, a stable matching is still not guaranteed to exist. One might expect that under this restriction some modification of the DA algorithm could be useful. Indeed, Dean et al. [2006] adapt the DA algorithm to this instance. It yields the first bound in Theorem 3.1, i.e. $\max_{h \in H} |k_h - k_h^*| \leq 2$. However, it can give no bound to the aggregate increase in capacity. In contrast, the technique outlined here improves the bound on individual hospitals to $\max_{h \in H} |k_h - k_h^*| \leq 1$, while preserving the bound on the aggregate increase in capacity.

# 4 Complementarity via Constraints

When preference complementarity is present on the hospital side, it is convenient to incorporate them via distributional constraints on the set of feasible matchings. In residency matching, the planner may wish to limit the number of residents assigned to a particular geographical *region*. In the context of school admissions, the planner wishes to ensure some level of representation of disadvantaged groups within each school. The introduction of additional constraints necessitates revisiting the definition of stability. Specifically, is a blocking coalition allowed to violate the additional constraints?

## 4.1 Regional Capacity Constraints

We examine a stable matching problem with regional capacity constraints proposed by Kamada and Kojima [2015]. It involves single doctors only and hospitals that are assigned to one or more sets called 'regions'. Let $R^1, R^2, \ldots R^t$ be the set of regions. Interpreted literally, regions would correspond to sets that partition the set of hospitals but we do not insist on such a restriction. In addition to the capacity constraint of each hospital, each region, $R^s$, has a cap, $K_s$, on the number of residents that can be assigned to it that may be less than the total capacity of all hospitals within it.

Using the same notation as before we can describe the set of feasible matchings as the 0-1 solutions of the following:

$$\sum_{d \in D^1} x_{(d,h)} \leq k_h \ \forall h \in H \tag{8}$$

$$\sum_{h \in H} x_{(d,h)} \leq 1 \ \forall d \in D^1 \tag{9}$$

$$\sum_{h \in R^s} \sum_{d \in D^1} x_{(d,h)} \leq K_s \ \forall s = 1, \ldots, t \tag{10}$$

As before, each single doctor has a strict preference ordering over the hospitals in $H$. Each hospital has responsive preferences. As there are no couples we can interpret hospital $h$'s priority ordering, $\succ_h$ as a preference ordering over the doctors in $D^1$.

Kojima et al. [2018] deal with the issue of stability by endowing each region with a preference ordering over single doctors. This is essential to make the problem well defined. To see why, suppose two hospitals in the same region each with capacity 1. The regional capacity constraint is 1 as well. Hence, at most one of these hospitals can be matched to a doctor. Suppose two doctors have a preference for being matched to the hospitals in this region. One doctor prefers the first hospital and the other the second. Each hospital's preference ordering alone cannot determine which one of these doctors will be matched to the region. Thus, the need for a regional ordering over doctors. Hence, each constraint of (8-10) has an ordering over the columns/variables that 'intersect' it. In this way all the conditions needed to invoke Scarf's lemma are present.[3]

If the collection of regions forms a laminar family, that is for any pair $R^s$ and $R^p$ either one is contained in the other or they are disjoint. A standard result (e.g. Edmonds and Giles [1977]) says that if the regions form a laminar family, then, every extreme point of (8-10) is integral. Hence, this system admits an integral dominating solution, i.e., a pairwise stable matching exists. Kojima et al. [2018] and Fleiner and Kamiyama [2012] derive the same result using the DA algorithm. However, if one drops the laminarity restriction on the regions, then, the DA algorithm fails. As is shown next Scarf's lemma can still be deployed and generate a matching that satisfies the stronger condition of group stability.

## 4.2 Multiple Dimensional Knapsack Constraints

In this section only, we depart from the doctor-hospital metaphor as it will not bear the strain of the variations we will now consider. The setting now is that of assigning families who happen to be refugees to localities within their country of refuge (see Delacrétaz et al. [2019], Nguyen et al. [2019]). It will generalize the problem of couples matching in two ways. First, on one side is a set $F$ of families of varying sizes. Couples matching corresponds to families of size at most 2. The hospital side corresponds to localities. However, unlike hospitals, the preferences of the localities cannot simply be described by a responsive choice function because families consume *multiple* resources, not just space, and do so at different intensities. For example, the educational resources consumed will depend on the number of school-age children. Senior care resources consumed will depend upon the number of elderly. Each locality is also interested in the likelihood of family members finding employment. This setting will also be used to show how one can ensure group stability rather than just pairwise stability.

Let $L$ be the set of localities and for each $\ell \in L$ and $f \in F$ let $v_{1,f}$ be a cardinal value $v_{\ell,f}$ that increases with the probability of member of $f$ finding employment in $\ell$.

For each $f \in F$, let $a^1_f, a^2_f, a^3_f$ be the number of children, adults, and elderly in the family $f$. For each $l \in L$ let $c^1_\ell, c^2_e ll, c^3_\ell$ be the limit on the number of children, adults, and elderly that locality $\ell$ can absorb.

Given a subset of families $K$, denote by $ch_\ell(K)$ the choice function (correspondence) of locality $\ell$. Locality $\ell$ will select a subset of families in $K$ that satisfies its capacity constraints on children, adults and the elderly and maximizes the sum of $v_{\ell,f}$s chosen. If we set $z_{\ell f} = 1$ if family $f$ is selected by locality $\ell$ and zero otherwise, $ch_\ell(K)$ can be expressed as a solution of the following

[3] In Kamada and Kojima [2015], hospitals have choice functions rather than preference orderings. See Section 3.1 for a discussion of how one can adapt Scarf's lemma to the case of choice functions.

integer program:

$$\max \quad \sum_{f\in K} v_{\ell f} z_{\ell f} \tag{11}$$

$$\text{s.t.} \quad \sum_{f\in K} a^s_f z_{\ell f} \le c^s_\ell, \text{ for } s \in \{1,2,3\} \tag{12}$$

$$z_{\ell f} \in \{0,1\} \tag{13}$$

Now $ch_\ell(K)$ does not satisfy the substitutes property. In fact, the preferences here can exhibit both substitutes and complementarity. A pairwise stable matching need not exist and even if it did, it needn't be Pareto optimal, see Nguyen et al. [2019] for an example.

If one follows the approach taken thus far, each column of the matrix $\mathcal{A}$ corresponds to a family and location pair. Therefore, domination would only capture pair-wise stability. To capture group stability, we need the columns to correspond to all potential blocking coalitions. The resulting $\mathcal{A}$ matrix is dense which means the rounding step cannot guarantee small violations of the soft capacity constraints.

Nguyen et al. [2019] overcome this problem by introducing contracts that specify both a family-locality match as well as a "price" for each one of the scarce resources that are consumed by the family at that locality. A *contract* between family $f$ and locality $\ell$ is represented by a variable $x_{f,\ell,p} \in [0,1]$, where $p$ is a 3-dimensional price vector specifying one price $p^s_{f,\ell}$ for each constraint or resource $s$ at $\ell$. Interpret $v_{f,\ell}$ as the total value of the match between family $f$ and locality $\ell$. The prices must be chosen so as to apportion the value between the family and the locality by resources consumed, i.e.,

$$\sum_{s=1}^{3} a^s_{f,\ell} \cdot p^s_{f,\ell} \le v_{f,\ell}. \tag{14}$$

Let $P_{f,\ell}$ be the set of all feasible price vectors $p$ satisfying equation (14).

The following infinite dimensional linear system describes the set of all feasible fractional matchings.

$$\sum_{\ell\in L}\sum_{p\in P_{f,\ell}} x_{f,\ell,p} \le 1 \qquad \text{for every family } f. \tag{15}$$

$$\sum_{f\in F}\sum_{p\in P_{f,\ell}} \frac{a^s_{f,\ell}}{c^s_\ell} \cdot x_{f,\ell,p} \le 1 \qquad \text{for every locality } \ell \text{ and } s \in \{1,2,3\}, \tag{16}$$

If we discretize the sets $P_{f,\ell}$, we ensure that (15-16) is finite-dimensional. In fact, Nguyen et al. [2019] show that there is a sufficiently small but positive discretization such that a dominating solution with respect to the discretized system is dominating with respect to the original system.

It remains to specify the ordering of each row over the columns that are in the support of that row.

- Consider the constraint (15). Its associated ordering is based on the preference list $\succ_f$ of family $f$ over localities. The row corresponding to a family $f$ is indifferent between any two columns $x_{f,\ell,p}$ and $x_{f,\ell,p'}$ for any $p \neq p'$.
- Consider constraint (16) which corresponds to the locality–service pair $(\ell, s)$. The corresponding ordering is based on the decreasing order of the price $p^s_{f,\ell}$. Therefore, the $(\ell, s)$-row prefers $x_{f,\ell,p}$ to $x_{f',\ell,p'}$ if $p^s_{f,\ell} > p'^s_{f',\ell}$. Hence, the row corresponding to $(\ell, s)$-row, prefers the contract that "pays" a higher price for $s$. The $(\ell, s)$-row is indifferent between the other columns.

It is not obvious that these orderings capture any information about the preferences of localities, but they do, as shown in Nguyen et al. [2019]. The key is that the prices of the contracts in a dominating solution correspond to the dual variables associated with (16).

Let $x^*$ be a dominating solution to (15-16). For each pair $(f, \ell)$, there is at most one $p \in P_{f,\ell}$ such that $x^*_{f,\ell,p} > 0$.

By setting

$$z^*_{f\ell} = \sum_{p \in P_{f,\ell}} x^*_{f,\ell,p},$$

we obtain a fractional matching. Furthermore, for each locality $\ell$, define the set of families $K$ to be the families that are assigned with positive probability to a less desirable locality compared with $\ell$. One can show that $z^*_{f\ell}$ is an optimal fractional solution to (11-13) which can be verified by using the price vector $p$ for which $x^*_{f,\ell,p} > 0$ to generate dual variables that complement the chosen primal solution. This shows that $z^*_{f\ell}$ is a fractional (weakly) group-stable matching.

Using the iterative rounding technique, discussed in Section 2.2, we can round the fractional group-stable matching $z^*$ into an integral $\bar{z}$. $\bar{z}$ satisfies (15) and a relaxed (16), whose right hand side is replaced by $1 + \Delta$ where $\Delta := \max_{f,\ell}\{\sum_{s=1}^{3} \frac{a^s_{f,\ell}}{c^s_\ell}\}$. Thus, the rounding technique yields an integral matching $\bar{z}$ that corresponds to a group-stable matching with respect to new capacities $\bar{c}^s_\ell$, where $\bar{c}^s_\ell \leq (1 + \Delta)c^s_\ell$.

## 4.3 Proportionality Constraints

The application of Scarf's lemma requires the constraint matrix that describes all feasible assignments of agents to coalitions, $\mathcal{A}$, which must be non-negative. When distributional concerns are expressed in terms of proportions i.e., the fraction of students of a particular type must be within some specified percentage of all students, this condition is not met.

One can replace proportionality constraints with capacity constraints. Suppose, for example, the constraint states that at most 20% of the students in a 100 seat school can be of a particular type. Replace this with the requirement that at most 20 students in the school must be of the relevant type. This switch is not benign because it assumes that the school will be at full capacity. Whether this is true or not depends on student preferences and the algorithm used to assign them to schools. Hence, we insist on the proportionality constraints but we will treat them as being 'soft' which is consistent with the way rules are written in some school choice settings. We now show how to describe the set of feasible matchings using linear inequalities.

As before, let $D^1$ be the set of single doctors. For each hospital $h$, let $D^h := \{d : d \succ_h \emptyset, h \succ_d \emptyset\}$ be the set of single doctors acceptable to $h$ and who find $h$ acceptable. Each $D^h$ is partitioned into $T_h$ sets: $D^h = D^h_1 \cup D^h_2 \cup \ldots \cup D^h_{T_h}$. A doctor $d \in D^h_t$ is said to be of type $t$ for hospital $h$. No two hospitals need to have the same partition. Thus, one hospital may choose to partition doctors by race and another by socio-economic status.

The proportionality constraint is that at each hospital $h$, the proportion of doctors of type $t$ for each $t$ must be at least $\alpha^h_t$ of the total number of doctors assigned to hospital $h$. The set of feasible matchings corresponds to all feasible 0-1 solutions to the following:

$$\sum_{h \in H} x_{dh} \leq 1 \ \forall d \in D^1 \tag{17}$$

$$\sum_{d \in D^1} x_{dh} \leq k_h \ \forall h \in H \tag{18}$$

$$\alpha^h_t [\sum_{d \in D^1} x_{dh}] - \sum_{d \in D^t_h} x_{dh} \leq 0 \ \forall t = 1, \ldots, t_h \ \forall h \in H \tag{19}$$

Here $0 \leq \alpha^h_t \leq 1, \sum_t \alpha^h_t \leq 1$. Notice, setting all variables to zero is a feasible solution.

In the presence of (19), one needs to modify the usual notion of blocking to rule out blocking pairs that violate (19). One might define $(h, d)$ to be a blocking pair if $d$ prefers $h$ to her current match and either *i.)* $h$ can accept $d$ without violating its capacity and proportionality constraints, or *ii.)* $h$ can replace a lower-ranked doctor (according to $\succ_h$) with $d$ so that $h$'s capacity and proportionality constraints are not violated. This is a weak notion of stability that can lead to a matching that is "wasteful". The empty matching, for example, would be stable! Even if one excluded empty matchings as candidates for being considered stable, we would have a problem, as shown in the next example.

**Example 4.1.** Consider a single hospital $h$ with capacity 100 and 100 doctors $d_1, \ldots, d_{100}$. All doctors strictly prefer to be matched to $h$ than remain unmatched, while the priority order of the hospital is $d_1 \succ_h d_2 \succ_h \cdots \succ_h d_{100}$. The set of doctors are divided into 2 subgroups, $D^h_1 = \{d_1, d_3, \ldots, d_{99}\}$ and $D^h_2 = \{d_2, d_4, \ldots, d_{100}\}$. The proportionality constraint is that at least 50 percent of the doctors in each subgroup are accepted.

Under the naive stability definition above, the matching that assigns $d_1, d_2$ to $h$ is stable because no *single* doctor can form a blocking coalition with $h$ because of the proportionality constraints. This matching is undesirable compared to the one that assigns all doctors to $h$.

To overcome the waste of 98 positions in example 4.1, one must allow each hospital to accept subsets of doctors who demand it without violating the constraints. Therefore, stability in this context cannot be restricted to just pairwise stability. One must allow for "coalitional" blocks that contain multiple students. We refer the reader to Kamada and Kojima [2017], Nguyen and Vohra [2019] for further details.

Once one settles on an appropriate definition of stability, the second difficulty is how to apply Scarf's lemma. Some of the coefficients in inequality (19) are negative. Thus a direct application of Scarf's lemma is impossible. However, Nguyen and Vohra [2019] propose a work around. Denote the constraints (17-18) by $\mathcal{A}x \leq b$. Denote (19) by $\mathcal{M}x \geq 0$. Now, $\{x \in \mathbb{R}^n_+ | \mathcal{M}x \geq 0\}$ is a polyhedral cone and can be rewritten as $\{\mathcal{V}z | z \geq 0\}$, where $\mathcal{V}$ is a finite non-negative matrix. The columns of $\mathcal{V}$ correspond to the **generators** of the cone $\{x \in \mathbb{R}^n_+ | \mathcal{M}x \geq 0\}$. The 'trick' is to apply Scarf's lemma to $\mathcal{P}' = \{z \geq 0 : \mathcal{A}\mathcal{V}z \leq b\}$ and use rounding to yield an integer $x^*$ that is stable such that

$$\sum_{h \in H} x^*_{dh} \leq 1 \ \forall d \in D^1$$

$$\sum_{d \in D^1} x^*_{dh} \leq k_h \ \forall h \in H$$

$$\bar{\alpha}^h_t [\sum_{d \in D} x^*_{dh}] \leq \sum_{d \in D^t_h} x^*_{dh} \ \forall t \ h \in H,$$

where

$$|\alpha^h_t - \bar{\alpha}^h_t| \leq \frac{2}{\sum_{d \in D} x^*_{dh}}.$$

Notice, the violation of proportionality constraints at hospital $h$ is bounded by $\frac{2}{\# \text{ assigned students at } h}$. If a hospital accepts more than 100 students, the matching violates the diversity constraints by at most 2 percent. However, the violation increases as the number of accepted students decreases which can be interpreted as smaller schools having "softer" diversity constraints. What determines whether a school receives a small or large number of accepted students are student preferences. The set of "small" schools cannot be determined a priori from capacity information alone. A large capacity school that is unpopular could end up with a number of students that is well below its capacity. Thus, relaxing their proportionality constraints allows them to recruit more students.

# 5 Other Methods

In this section, we briefly discuss other methods for dealing with complementarities and externalities in matching.

## 5.1 Restricting Preferences

In the matching with couples problem, we can guarantee the existence of a stable matching by restricting the preferences of couples so as to limit the degree of complementarity exhibited. Klaus and Klijn [2005], for example, propose a restriction (called weakly responsive) based on interpreting a couple's preference ordering over-ordered pairs as being a particular aggregation of the individual preferences of each member of the couple.

Under this restriction, Klaus and Klijn [2005] shows that a stable matching exists via the DA algorithm. However, the DA algorithm must have access to the preference orderings of each member of the couple as well as the couple's ordering. Hence, whether one can use the DA algorithm in practice will depend on how couples are asked to communicate their preferences.

Other examples of restrictions are (Cantala [2004], Pycia [2012], and Sethuraman et al. [2006].) Like Klaus and Klijn [2005], they rule out certain complementarities in preferences. Interestingly, Tang et al. [2018] has shown that under some of these preference restrictions, all dominating extreme points are integral. This is surprising given that not all the extreme points of (5-7) are integral.[4]

On the flip side, by allowing for 'extreme' kinds of complementarities such as 'all or nothing', one can guarantee the existence of stable matchings. See Rostek and Yoder [2020] for an example.

When the hospital side has distributional preferences, it is common to modify their choice function so as to favor various groups. If the modified choice function is specified in the right way, the DA algorithm (or some variant) will find a stable matching. However, there is no ex-post guarantee on the realized distribution.[5] Examples of this approach can be found in Ehlers et al. [2014]. In lieu of an ex-post guarantee, some authors focus on priorities that will produce distributions that are closest to a target distribution, see Erdil and Kumano [2012] and Echenique and Yenmez [2015].

## 5.2 Large Markets

In other settings (see Aumann [1964]) it is well known that one can wash away the problems caused by complementarities by making the underlying economy large. The same is true in the matching context as well. See **??** for more details.

Kojima et al. [2013] and Ashlagi et al. [2014], for example, show that in a setting where applicant preferences are drawn independently from a distribution, as the size of the market increases and the proportion of couples approaches 0, a stable matching exists and it can be found with high probability using a modification of the DA algorithm. However, Ashlagi et al. [2014] shows that when

[4]In the case of weakly responsive preferences, the algorithm for finding a dominating extreme point only requires access to the couple's ordering and not the individual ones.

[5]If one is not careful, there is also a "circularity" problem, in that stability is defined with respect to the modified choice function.

the proportion of couples is positive, the probability that no stable matching exists is bounded away from 0 even when the market's size increases.

Beyond the large but finite setting is the continuum. Greinecker and Kah [2020] provides a nice summary of the measure theoretic issues associated with both defining what constitutes a matching as well as stability with a continuum of agents on both sides. They propose that matchings be interpreted as joint distributions over the characteristics of the populations to be matched and introduce a novel stability notion for pairwise matchings based on sampling. They show that stable matchings exist and, importantly, correspond precisely to limits of stable matchings for finite-agent models.

Prior work assumes a continuum of agents on one side only or that there are a finite number of distinct types of agents. This has the advantage of side-stepping some of the measure theoretic issues present in Greinecker and Kah [2020] as well as providing 'interpretable' characterizations of stability. Azevedo and Leshno [2016], Azevedo and Hatfield [2015] as well as Che et al. [2015], for example, guarantee the existence of a stable matching and characterize it in terms of thresholds which are the analog of prices.

Even in these cases, Scarf's lemma may be useful. Wu [2018], for example, introduces a convexity condition for matching problems that via Scarf's lemma implies that the core is nonempty. It allows for arbitrary contracting networks, multilateral contracts, and complementary preferences. The condition is not related to preferences but the space of feasible matchings. It requires that the set of feasible allocation is convex and that for each potential block, the set of unblocked allocations is also convex. This convexity condition is present in Azevedo and Hatfield [2015] which allows him to extend that paper to doctors having continuous preferences over a continuum of alternatives.

### 5.3 Relax Stability

A third approach is to relax the stability requirement. Manlove et al. [2017], for example, proposes finding a matching that admits the minimum number of blocking pairs. However, they show that this problem is also NP-hard and difficult to approximate. Even were this not the case, this relaxation of stability is problematic. Consider a stable matching with a single blocking coalition. It is possible that was this block to form, it would trigger an entire chain of blocks that were not initially present. Others modify the notion of stability so that some modification of the DA algorithm will succeed. See for example, Klijn and Masso [2003] and Dur et al. [2018]. Depending on the context, these modifications need not capture the original spirit of the notion of stability.

## 6 Open Questions

An attractive feature of the DA algorithm is that it is strategy-proof for the proposing side but it is not a property enjoyed by all dominating solutions. A natural question is to determine whether it is possible to identify a dominating solution in a strategy-proof way.

Just as the DA algorithm can be modified to find a stable matching that favors one side or the other, one may ask whether the same is true for dominating solutions. In two-sided settings how does one find a dominating solution that favors one side over the other?

Finally, it would be interesting to identify other instances where determining a dominating solution can be executed in polynomial time.

# 7 Notes

The stable matching problem was introduced in Gale and Shapley [1962]. Examples of extensions of Gale and Shapley [1962] can be found in Fleiner [2003], Hatfield and Milgrom [2005], Ostrovsky [2008], and Hatfield and Kojima [2010].

Complementarities and side constraints are studied in many applications of matching markets. For examples of complementarities in school choice see Delacrétaz [2019] and Correa et al. [2019]. Abdulkadiroglŭ and Sönmez [2003], Biró et al. [2010]) and Huang [2010] deal with proportionality constraints by converting them into constraints on absolute numbers.

The Scarf quote being paraphrased appears in Scarf [1994]. The published version of Scarf's lemma is Scarf [1967]. The version of the lemma we state appears in Scarf [1965]. A modernized version can be found in Király and Pap [2008]. The proof given follows Scarf [1965] with a simplification that avoids a double limit argument. The first finite time algorithm for finding a dominating extreme point is given in Scarf [1967]. The proof that the problem of finding a dominating extreme point is PPAD complete is in Kintali [2008]. The iterative rounding technique is described in Lau et al. [2011].

# Acknowledgements

We thank two anonymous referees for their valuable suggestions.

# References

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

I. Ashlagi, M. Braverman, and A. Hassidim. Stability in large matching markets with complementarities. *Operations Research*, 62(4):713–732, 2014.

R. J. Aumann. Markets with a continuum of traders. *Econometrica*, 32(1/2): 39–50, 1964.

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

E. M. Azevedo and J. D. Leshno. A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, 124(5):1235–1268, 2021/05/10 2016.

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

E. Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061 – 1103, 2011.

D. Cantala. Matching markets: the particular case of couples. *Economics Bulletin*, 3(45):1–11, 2004.

Y.-K. Che, J. Kim, and F. Kojima. Stable matching in large economies. 2015.

J. Correa, R. Epstein, J. Escobar, I. Rios, B. Bahamondes, C. Bonet, N. Epstein, N. Aramayo, M. Castillo, A. Cristi, et al. School choice in Chile. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 325–343, 2019.

B. C. Dean, M. X. Goemans, and N. Immorlica. The unsplittable stable marriage problem. In *Fourth IFIP International Conference on Theoretical Computer Science-TCS 2006*, pages 65–75. Springer US, 2006.

D. Delacrétaz. Stability in matching markets with sizes, 2019.

D. Delacrétaz, S. D. Kominers, A. Teytelboym, et al. Matching mechanisms for refugee resettlement. Technical report, 2019.

U. Dur, T. Morrill, and W. Phan. Family ties: School assignment with siblings. Technical report, mimeo, 2018.

F. Echenique and M. B. Yenmez. How to control controlled school choice. *American Economic Review*, 105(8):2679–94, 2015.

J. Edmonds and R. Giles. A min-max relation for submodular functions on graphs. In *Annals of Discrete Mathematics*, volume 1, pages 185–204. Elsevier, 1977.

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

A. Erdil and T. Kumano. Prioritizing diversity in school choice. *Unpublished working paper, Washington University*, 1, 2012.

T. Fleiner. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research*, 28(1):103–126, Feb. 2003.

T. Fleiner and N. Kamiyama. A matroid approach to stable matchings with lower quotas. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 135–142. SIAM, 2012.

D. Gale and L. S. Shapley. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

M. Greinecker and C. Kah. Pairwise stable matchings in large economies. 2020.

J. W. Hatfield and F. Kojima. Substitutes and stability for matching with contracts. *Journal of Economic Theory*, 145(5):1704–1723, 2010.

J. W. Hatfield and P. R. Milgrom. Matching with Contracts. *American Economic Review*, 95(4):913–935, September 2005.

C.-C. Huang. Classified stable matching. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1235–1253. Society for Industrial and Applied Mathematics, 2010.

Y. Kamada and F. Kojima. Efficient matching under distributional constraints: Theory and applications. *American Economic Review*, pages 67–99, 2015.

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

S. Kintali. Scarf is PPAD-complete. *CoRR*, abs/0812.1601, 2008.

T. Király and J. Pap. Kernels, stable matchings, and Scarf's lemma. *Working paper, Egerváry Research Group, TR-2008-13*, 2008.

B. Klaus and F. Klijn. Stable matchings and preferences of couples. *Journal of Economic Theory*, 121(1):75–106, 2005.

F. Klijn and J. Masso. Weak stability and a bargaining set for the marriage model. *Games and Economic Behavior*, 42:91–100, 2003.

F. Kojima, P. A. Pathak, and A. E. Roth. Matching with couples: Stability and incentives in large markets. *The Quarterly Journal of Economics*, 128 (4):1585–1632, 2013.

F. Kojima, A. Tamura, and M. Yokoo. Designing matching mechanisms under constraints: An approach from discrete convex analysis. *Journal of Economic Theory*, 176:803–833, 2018.

L. C. Lau, R. Ravi, and M. Singh. *Iterative methods in combinatorial optimization*, volume 46. Cambridge University Press, 2011.

D. F. Manlove, I. McBride, and J. Trimble. "almost-stable"matchings in the hospitals / residents problem with couples. *Constraints*, 22(1):50–72, 2017.

H. Nguyen, T. Nguyen, and A. Teytelboym. Stability in matching markets with complex constraints. In *EC*, page 61, 2019.

T. Nguyen and R. Vohra. Near-feasible stable matchings with couples. *American Economic Review*, 108(11):3154–69, 2018.

T. Nguyen and R. Vohra. Stable matching with proportionality constraints. *Operations Research*, 67(6):1503–1519, 2019.

M. Ostrovsky. Stability in supply chain networks. *American Economic Review*, 98(3):897–923, 2008.

M. Pycia. Stability and preference alignment in matching and coalition formation. *Econometrica*, 80(1):323–362, 2012. ISSN 1468-0262.

M. Rostek and N. Yoder. Matching with complementary contracts. *Econometrica*, 88(5):1793–1827, 2020.

A. E. Roth. The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy*, 92(6):pp. 991–1016, 1984.

A. E. Roth and A. Postlewaite. Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics*, 4(2):131–137, 1977.

H. Scarf. The allocation of resources in the presence of indivisibilities. *Journal of Economic Perspectives*, 8(4):111–128, 1994.

H. E. Scarf. An elementary proof of a theorem on the core of an $n$ person game. *Cowles Foundation Working Paper*, 1965.

H. E. Scarf. The core of an $n-$ person game. *Econometrica*, 35(1):pp. 50–69, 1967.

J. Sethuraman, C.-P. Teo, and L. Qian. Many-to-one stable matching: Geometry and fairness. *Mathematics of Operations Research*, 31(3):581–596, 2006.

R. M. Starr. Quasi-equilibria in markets with non-convex preferences. *Econometrica*, pages 25–38, 1969.

D. Tang, T. Nguyen, V. Subramanian, and R. Vohra. Computational testing of scarf's algorithm for near feasible stable matching with couples. Technical report, University of Michigan, Ann Arbor, 2018.

X. Wu. Core of convex matching games: A Scarf's Lemma approach. Working paper, Columbia University, Department of Economics, 2018.