# Matching under Non-transferable Utility: Applications*

Tayfun Sönmez M. Utku Ünver

Chapter 3 of the *Handbook of the Economics of Matching*, Volume 1 (Elsevier)
edited by Yeon-Koo Che, Pierre-André Chiaporri, and Bernard Salanié

March 2024

**Abstract**

We survey the literature on applications of matching theory under non-transferable utility. We cover the following six applications in detail: living-donor kidney exchange, living-donor liver exchange, cadet-branch matching in the US Army, affirmative action in India, matching market for entry-level physicians in the US, and course allocation at universities. We also survey other notable applications.

**Keywords:** Matching Theory, Market Design, Kidney Exchange, Liver Exchange, Cadet-Branch Matching, Affirmative Action in India, Matching for Residency Programs, NRMP, Unraveling, Course Allocation, College Admissions, School Choice, Pandemic Resource Allocation, Reserve Systems, Matching under Distributional Constraints, Matching with Reassignment, Balancedness in Matching, Refugee Resettlement

*Sönmez: Boston College, Department of Economics, email: `tayfun.sonmez@bc.edu`.
Ünver: Boston College, Department of Economics, email: `unver@bc.edu`.
We thank Eric Budish, Parag Pathak, and Assaf Romm for their valuable input in this chapter.

# Contents

# 1 Introduction

In the last 25 years, fundamental theory on matching markets and custom-made theory respecting subtle institutional details in various settings has led to several successful applications, in many cases leading to direct policy impact. In addition, matching theory has been used as an analytical tool in understanding the consequences of certain market structures and in formally modeling the reasons for their failures and successes.

How has matching theory become more successful in spearheading direct policy applications than many other branches of economic theory? We will briefly touch on the political economy of market design for matching markets, as it is central to many applications covered in this chapter. We recommend that interested readers also refer to Sönmez (2023) for further details.[1] Here, we highlight two approaches that contribute to this success.

The first approach is *commissioned design*. Policymakers reach out to experienced and well-known scholars when their institutions are in trouble and on the verge of failure. The scholar, bringing both experience and credibility, is tasked with designing a new institution or reforming an existing one. This design may draw on successful ideas from past settings or involve creating innovative solutions from scratch using various technical tools. Two prominent examples of this approach are the redesign of the National Residency Matching Program matching mechanism (Roth and Peranson, 1999) and the New York City high school matching mechanism (Abdulkadiroğlu et al., 2005, Abdulkadiroğlu, Pathak, and Roth, 2009).

The second approach is *aspired design*. Design economists study the operational details of an institution and show that either it fails to fulfill some key policy goals or there are better ways of achieving such goals. Aspiring for reform from outside the system, they propose to *minimally* change the current institution in the hopes of persuading the stakeholders to achieve their goals. Solid theoretical and other scholarly evidence may be needed to persuade the stakeholders, as often many have vested interests in maintaining the status quo. Often, custom-made theory that strongly aligns with the application may be key for policy impact. Some important examples include the redesign of the Boston school district's school choice mechanism (Abdulkadiroğlu and Sönmez, 2003; Abdulkadiroğlu et al., 2005), the foundation of kidney exchange programs (Roth, Sönmez, and Ünver, 2004, 2005a,b, 2007), the redesign of course allocation mechanisms (Budish, 2011; Budish, Lee, and Shim, 2019; Budish and Kessler, 2022), designing policies and mechanisms for equitable allocation of scarce medical

[1] Sönmez (2023) outlines a methodological institution design framework called "minimalist market design," which relies on crafting custom-made theory to enhance practical relevance. Many successful policy impact efforts align with this framework.

resources during the COVID-19 pandemic (Pathak et al., 2023; Sönmez et al., 2021; Rubin et al., 2021; White et al., 2022; McCreary et al., 2023), the redesign of the cadet-branch matching mechanism for the US Army (Sönmez and Switzer, 2013; Sönmez, 2013; Greenberg, Pathak, and Sönmez, 2023), and the design of liver-exchange clearinghouses (Ergin, Sönmez, and Ünver, 2020, Yilmaz et al., 2023b).

In this chapter, we cover six notable applications of matching theory in depth.

1. The design and analysis of *living-donor kidney exchange* clearinghouses are discussed in Section 2, where the economic models and the evolution of institutions through the intervention of design economists are explored. The models are based on the private and mixed ownership economy models covered in Sections 2 and 3 of Chapter 1.
2. The design and analysis of *living-donor liver exchange* clearinghouses are discussed in Section 3, focusing on the evolution of the models starting from those in living-donor kidney exchange (cf. Section 2) and their policy impact. This serves as an application of exchange under private ownership, as detailed in Section 2 of Chapter 1.
3. The design of the *cadet-branch matching* mechanism for the United States Military Academy at West Point and Reserve Officer Training Corps is discussed in Section 4, where the evolution of the market and the theory behind it are covered. A detailed background is provided in Chapter 9 of this handbook under matching with contracts models.
4. *Constitutional design of affirmative action policies* in India is discussed in Section 5, focusing on its legal framework and model. It is based on priority-based entitlements covered in Section 4 of Chapter 1.
5. Centralization of *entry-level physician labor markets* occurred due to a costly unraveling period in the 1950s, leading to the redesign of its matching algorithm. The evolution of the market and economic models related to the chaotic behavior that resulted in centralization in similar markets is covered in Section 6. Its centralized matching model is based on the models discussed in Section 2 of Chapter 1 under two-sided matching markets.
6. *Course allocation* to students at universities is discussed in Section 7, where the market institutions, models, and designs are explored. This serves as an application of common ownership economies covered in Section 3 of Chapter 1.[2]

In the last section of this chapter, we briefly touch upon several other notable applications of matching theory. These discussions are briefer for various reasons:

[2] We are indebted to Eric Budish, who provided his extensive sets of slides to us on his and his coauthors' papers on this topic, that helped us tremendously in preparing this section.

(i) They are covered elsewhere in this handbook.
(ii) They have not yet demonstrated significant policy impact or undergone external validity testing.
(iii) We couldn't find a suitable integration of their theoretical foundations with the fundamental models discussed in Chapter 1.
(iv) There is limited publicly available scholarly work on these applications, which makes documentation challenging.

The field of market design in matching markets is expansive and constantly evolving. Given the depth and breadth of applications across various domains, some noteworthy ones might not have been included despite our best efforts. As the field continues to grow, new applications, advancements, and policy impacts emerge regularly. Our coverage may not encompass every notable application, and this is a testament to the field's rapid growth and reach.

# 2 Living-Donor Kidney Exchange

Kidney exchange, as an unexpected application of market design in early 2000s, has not only enhanced the visibility of this field but also significantly transformed the institutions governing living-donor kidney donation within a remarkably short period. In just a few years after its introduction, this approach revolutionized the practice of living-donor kidney transplantation in many countries, paving the way to save thousands of lives across the globe each year. This section provides an exploration of the institutional background and the meticulous market design strategies adopted by economists to establish kidney exchange clearinghouses.[3]

About 96,000 patients were waiting for a deceased-donor kidney transplant in the United States as of November 2023, a number which stabilized in recent years.[4] In 2022:

- 44,442 new patients joined the deceased-donor queue,
- 19,630 received deceased-donor transplants,
- 5,854 received living-donor transplants,
- 4,673 died while waiting, and
- 15,066 were removed from the queue for other reasons.

[3]This section aligns with and complements previous surveys and perspectives coauthored by the authors of this chapter (Sönmez and Ünver, 2013, 2017, 2023). Kidney exchange has become an prominent interdisciplinary research area spanning bioethics, medicine, health policy, economics, operations research, computer science, and sociology. For example, see Dickerson and Sandholm (2016) and Ashlagi and Roth (2021) for surveys of the topics covered here from the perspectives of computer science and operations research disciplines, respectively. We will mostly focus on economic models in the literature that can be tied to the theoretical background explored in Chapter 1 and how they played a role in developing different exchange paradigms and their resulting policy impact.

[4]Retrieved through https://optn.transplant.hrsa.gov/, National Data option.

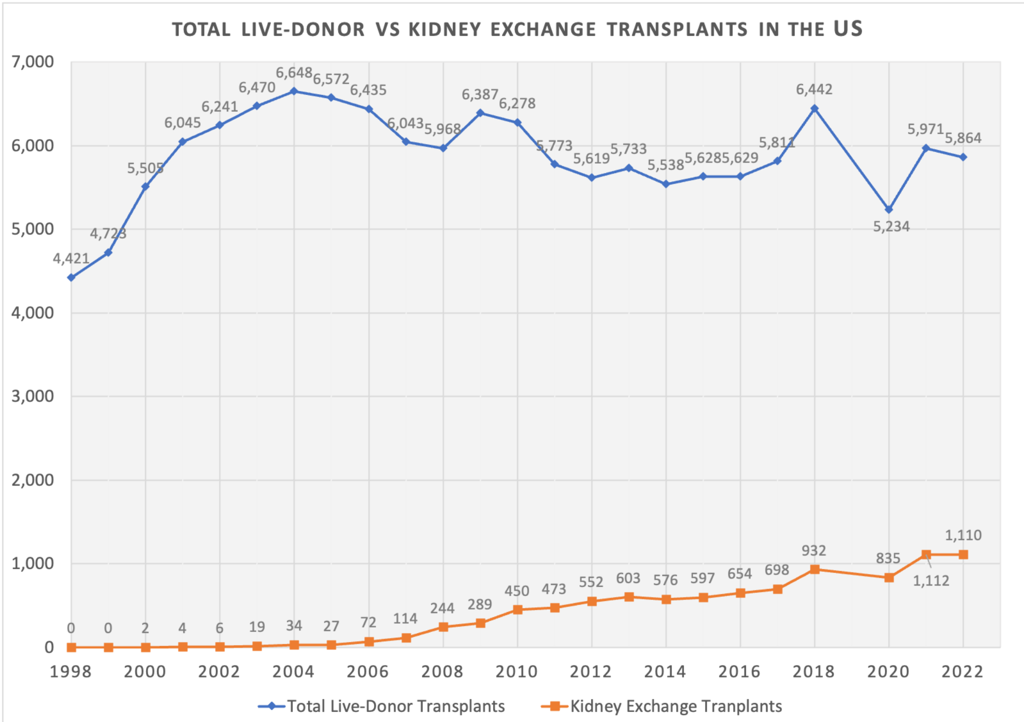



**Figure 1:** The number of total and kidney-exchange transplants in the US since 1999 (source: https://optn.transplant.hrsa.gov/, using the National Data option, retrieved on 11/14/2023.

Of the 5,854 living-donor transplants, 1,110 were through kidney exchanges, significantly increasing since the early 2000's. (See Figure 1). Before 2004, when economists published their first paper on this subject in 2004, the total number of kidney exchange transplants conducted in the US in five years was fewer than 50.

## 2.1 Background

Transplantation is the preferred treatment for the most severe forms of kidney disease. There is a major shortage of transplantation organs all over the world. As buying and selling a body part is illegal in many countries in the world, including the US (National Organ Transplant Act - NOTA 1984/2007), donation is the only legal source of transplant organs.

### 2.1.1 Constraints to Donation: Medical Compatibility

A donor must pass two *medical compatibility* tests, in addition to undergoing a psycho-sociological screening. If any of these tests fail, the donor is deemed incompatible with the patient. The medical compatibility requirements are as follows:

**Blood-type Compatibility:** Although there are different blood type groupings, the most commonly used one for kidney compatibility is the ABO blood grouping. There are four most common human ABO blood types: $O$, $A$, $B$, and $AB$. While a donor of blood-type $O$ can donate to recipients of all blood types, a donor of blood-type $A$ can donate to patients of blood-types $A$ and $AB$, a donor of blood-type $B$ can donate to patients of blood-types $B$ and $AB$, and a donor of blood-type $AB$ can only donate to patients of blood type $AB$ (see Figure 2).

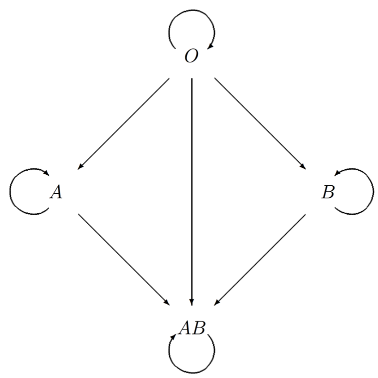



**Figure 2:** Blood-type compatibility.

**Tissue-type Compatibility:** Tissue type or Human Leukocyte Antigen (HLA) type is a combination of several pairs of antigens on Chromosome 6. HLA proteins A, B, DR, and DQ are especially important in determining the tissue type of a donor. Before transplantation, the potential recipient's blood plasma is tested for preformed antibodies against donor HLA. If antibody levels are above a threshold, the transplant cannot be carried out. For a random patient and donor, there is about a 10%-15% chance of such tissue-type incompatibility (Zenios, Woodle, and Ross, 2001). This background probability for each patient is known as *panel reactive antibody* (*PRA*) level. However, there is a huge variance in PRA levels among patients. For example, different background PRA levels imply that certain patients may reject more than 99% of the offered kidneys. At the same time, some of them are almost always compatible with a blood-type compatible donor.[5] It is possible to detect to a high precision level using single HLA level tests whether a patient is tissue-type incompatible with a type of each single HLA.[6]

In the last 20 years, *desensitization* methods have been developed to address both ABO incompatibility and tissue-type incompatibility, enabling incompatible donors to donate to their paired patients. These treatments involve filtering the patient's plasma periodically to remove HLA or ABO antibodies. The process can take up to

[5]See OPTN data for such rates at https://optn.transplant.hrsa.gov/.

[6]This technology enabled the development of kidney exchange programs, without the need of mixing each donor's blood with a patient's blood plasma.

6 months. However, these procedures are expensive, and there is no conclusive evidence that the transplanted organ's longevity matches that of compatible transplants. As a result, neither form of desensitization is typically practiced in the US and much of the Western world.

### 2.1.2 Transplantation Policies

Using two well-established technologies, three strategies for transplantation were developed before the intervention of market design economists:

**Deceased Donation:** Deceased-donor organs are considered a national treasure and are usually allocated to patients on an organ-specific deceased-donor waitlist (or simply waitlist) using a centralized point system in the US and many other countries, akin to a common-ownership economy covered in Section 3 of Chapter 1. Each deceased donor can save up to two kidney patients when both of their kidneys are suitable for transplantation. In the US, the allocation policy evolved over the years.

- Before 2014, it was almost like a *first-in–first-out* (FIFO) queue in different transplant regions of the country, where waiting time got the highest priority criterion (subject to certain exceptions, such as past living donors receiving priority if their remaining kidney ever failed). This system resembled a dynamic variant of the type of priority mechanism covered in Chapter 1.
- In 2014, this policy was amended to create a *reserve system* where the highest quality 20% of the kidneys were entirely allocated to the 20% of the patients with the best chance of long-term survival (e.g., younger or healthier patients) while everybody was eligible for the remaining 80% of the deceased-donor kidneys. This system is a dynamic version of a static reserve system covered in Section 4 of Chapter 1.
- As of 2023, a new methodology that is being developed by the United Network for Organ Sharing (UNOS) in the US aims to use a more granular point system that considers many other variables to form the priority mechanism, called the *continuous allocation* policy.

**Direct Living Donation:** Under this technology, loved ones of a patient come forward, and if one is compatible with the patient, one of their kidneys is transplanted to the patient. A kidney transplanted from a living donor lasts 1.5 to 2 times longer on average than one from a deceased donor. Thus, a living-donor transplant is preferred to a deceased-donor transplant. Also, living donation eliminates the need for a patient to wait on the deceased-donor waitlist – often for many years – while going through a life-quality decreasing and more expensive treatment option called dialysis continuously.

**Living-donor Kidney Exchange:** This is the newest of the three policies and still the least resorted option. It was rarely used before market design economists got involved in designing the clearinghouses that organize kidney exchanges. There were two established modalities of this policy before 2004 in the US:

**Paired Exchange:** If the living donor who came forward for their patient is incompatible, the donor is swapped with the donor of a similar patient-donor pair to find a compatible match for the patient (proposed by Rapaport, 1986, see Figure 3).

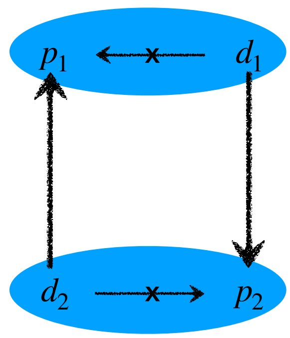



**Figure 3:** A typical two-way exchange

As a logistical constraint, all transplants in one paired exchange have to be done simultaneously to prevent the reneging of a donor whose patient has already received a transplant.

The first kidney exchange was conducted in South Korea in 1991 (Huh et al., 2008). This was followed by a program in the Netherlands (De Klerk et al., 2005). While the first kidney exchange in the US was conducted in 1994, the first exchange program was established in US Transplantation Region 1, consisting of New England, in the early 2000s. The New England program and other transplant centers in the US conducted paired exchanges infrequently until 2004, as databases regarding patients and their paired donors were not adequate, and formal exchange mechanisms were yet to be introduced to find the best combinations of exchanges.

**List Exchange:** If a living donor who came forward for their patient is not compatible with them, the patient receives a priority in the deceased-donor list while the donor's organ is given to the highest-priority compatible patient on the waitlist (introduced by Ross and Woodle, 2000).

This was mainly practiced in US Transplantation Region 1, New England, in the early 2000s, along with paired exchanges (See Figure 4). They were conducted more frequently than paired exchanges, as they were easier to organize.

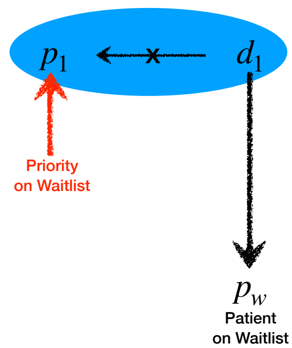



**Figure 4:** A typical two-way paired exchange

### 2.1.3 Kidney Exchange as an Application in Market Design

Roth, Sönmez, and Ünver (2004) observed that the two main types of kidney exchanges conducted in the US, the paired exchange and the list exchange, correspond to the most basic forms of the two types of exchanges in the house allocation problem with existing tenants model (Abdulkadiroğlu and Sönmez, 1999), which we covered in Chapter 1, Section 3. Building on the existing practices in kidney transplantation, they introduced efficient and incentive-compatible mechanisms. They illustrated conceptually and quantified through simulations possible increases in the number of kidney transplants by using their mechanism. The methodology and techniques advocated in their following research program (Roth, Sönmez, and Ünver, 2005b,a, 2007; Saidman et al., 2006; Roth et al., 2006) provided the backbone of several kidney exchange programs established in the US, including the New England Program for Kidney Exchange (NEPKE), Alliance for Paired Donation (APD), National Kidney Registry (NKR), and The National UNOS Paired Donation Program, and in the rest of the world.

## 2.2 A General Kidney Exchange Model

We introduce a general kidney exchange model that encompasses theretical models in Roth, Sönmez, and Ünver (2004, 2005b, 2007), Sönmez and Ünver (2014) and conceptual and simulated models in Roth et al. (2006) and Saidman et al. (2006).

A *kidney exchange problem* is a list $\left[P,\ \Omega,\ A,\ \succsim,\ n^e,\ n^d,\ n^a\right]$ such that

- $P$ is a set of *patients*.
- $\Omega = \left(\Omega_p\right)_{p \in P}$ is a *paired-donor profile*. Each patient $p \in P$ is *paired* with a distinct non-empty set of living donors $\Omega_p$ such that for each distinct pair of patients $p, r \in P$,
$$\Omega_p \cap \Omega_r = \emptyset.$$
- $A$ is a set of *altruistic donors*, who are not paired with any patient so that $A \cap \left(\bigcup_{p \in P} \Omega_p\right) = \emptyset$.

- $\succsim = (\succsim_p)_{p \in P}$ is a *preference profile*. For each patient $p \in P$, $\succsim_p$ is a preference relation (a binary relation that is complete, reflexive, and transitive) over $\left(\bigcup_{r\in P} \Omega_r\right) \cup A \cup \{\emptyset, w\}$ where
  - each $d \in \left(\bigcup_{r\in P} \Omega_r\right) \cup A$ represents receiving a transplant from donor $d$ such that

    $$\text{donor } d \text{ is medically incompatible with patient } p \quad \Longrightarrow \quad \emptyset \succ_p d,$$

  - $w$ represents receiving a priority in the deceased-donor waitlist for a future deceased-donor transplant, and
  - $\emptyset$ represents remaining unmatched.
- $n^e \in \{2, 3, \ldots, |P|\}$ is an integer denoting the maximum *paired exchange size* that is feasible.
- $n^d \in \{0, 1, 2, 3, \ldots, |P|\}$ is an integer denoting the maximum *deceased-donor chain size* that is feasible.
- $n^a \in \{0, 1, 2, 3, \ldots, |P|\}$ is an integer denoting the maximum *altruistic donor chain size* that is feasible.

Medical compatibility is defined in Section 2.1.1. Only medically compatible donors can be acceptable in a patient's preference. On the other hand, some medically compatible donors can be unacceptable for a patient.

We formally explain the implications of maximum paired exchange size $n^e$, maximum deceased-donor chain size $n^d$, and maximum altruistic donor chain size $n^a$ after introducing the concepts of paired exchanges, deceased-donor chains, and altruistic donor chains.

**Matchings.** An outcome of a problem is a *matching*, which is a function $\mu : P \to \left(\bigcup_{p\in P} \Omega_p\right) \cup A \cup \{\emptyset, w\}$ such that

1. each donor can be matched with at most one patient, i.e.,

$$\forall\, d \in \big(\bigcup_{p\in P} \Omega_p\big) \cup A, \quad \left|\mu^{-1}(d)\right| \le 1,$$

2. at most one paired donor of a patient can be matched, i.e.,

$$\forall\, p \in P, \quad \left|\mu^{-1}(\Omega_p)\right| \le 1,$$

3. if a paired donor of a patient is matched, then the patient is matched as well, i.e.,

$$\forall\, p \in P, \quad \left|\mu^{-1}(\Omega_p)\right| = 1 \implies \mu(p) \ne \emptyset.$$

Since the number of patients who are eventually matched with $w$ is envisaged to be small relative to the number of deceased donors that arrive in a relatively short period, the priority in waitlist option $w$ can be assigned to multiple patients.

Conditions 2 and 3 imply that at most one paired donor of a patient is matched, and if one of their donors is matched, they also have to be matched to receive either a deceased-donor transplant through a priority on the waitlist $w$ or a living-donor transplant from a donor, who could be their own paired donor. We do not formally model the patients on the deceased-donor waitlist. As a result, in a matching, it could be the case that a patient is matched, but none of their donors are matched to any patient in $P$. If that is the case, it means that the central authority chooses this patient's exactly one donor

- to donate to a patient on the deceased-donor waitlist or
- to be utilized as a new altruistic donor in the future iterations of the kidney exchange problem when new patients and their donors arrive.

Thus, the opportunity cost of receiving a transplant is exactly giving up one donor for each patient in $P$.

We define the following two important properties of matchings.

A matching $\mu \in \mathcal{M}$ is *individually rational* if $\mu(p) \succeq_p \emptyset$ and $\mu(p) \succeq_p d$ for each $d \in \Omega_p$.

A matching $\mu \in \mathcal{M}$ is *Pareto efficient* if there exists no other matching $\nu \in \mathcal{M}$ that makes each patient weakly better off and at least one patient strictly better off.

**The graph representing a matching.** It will be helpful to represent a matching $\mu \in \mathcal{M}$ as a directed graph. Let $N = P \cup A \cup \{w, \emptyset\}$ be the set of nodes of this graph. A directed edge $(p, x)$ in the graph is an ordered pair of a patient node $p \in P$ and a node $x \in P \cup A \cup \{w, \emptyset\}$. In this graph, we refer to the edge $(p, x)$ as *$p$ points to $x$*. We define the *graph representing $\mu$* as a pair $\langle N, \mathcal{E} \rangle$ where $\mathcal{E}$ is the set of directed edges defined as

$$\mathcal{E} = \left\{ (p,x) \in P \times \big(P \cup A \cup \{w, \emptyset\}\big) \ : \ \begin{array}{c} x \in P \quad \text{and} \quad \mu(p) \in \Omega_x, \\ \text{or} \\ x \in A \cup \{w, \emptyset\} \quad \text{and} \quad \mu(p) = x \end{array} \right\}.$$

The resulting "kidney exchange" graph consists of four types of subgraphs:

1. A *cycle* $(p_1, p_2, ...,, p_k)$ with $k \geq 1$, as illustrated in Figure 5, consists of a list of patients, such that

$$\mu(p_j) \in \Omega_{p_{j+1}} \qquad \forall\, j \in \{1, 2, \ldots, k\} \ (j \text{ in modulo } k).$$

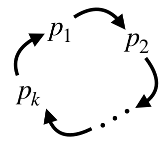


**Figure 5:** A cycle with $k$ patients represented as a subgraph.

2. An *unmatched patient chain* $(p, \emptyset)$, as illustrated in Figure 6, is a list that consists of a single patient and the remaining unmatched option $\emptyset$, such that $\mu(p) = \emptyset$.

$$p \longrightarrow \varnothing$$

**Figure 6:** An unmatched patient chain represented as a subgraph.

3. A *w-chain* $(p_1, p_2, ..., p_k, w)$ with $k \geq 1$, as illustrated in Figure 7, consists of a list of patients and the priority in the waitlist option $w$, such that

$$\begin{aligned} &\mu(p_j) \in \Omega_{p_{j+1}} \quad \forall\, j \in \{1, 2, \ldots, k-1\}, \quad \text{and} \\ &\mu(p_k) = w. \end{aligned}$$

$$p_1 \longrightarrow p_2 \longrightarrow \cdots \longrightarrow p_k \longrightarrow w$$

**Figure 7:** A $w$-chain with $k$ patients represented as a subgraph.

We refer to patient $p_1$ as the *tail patient* of the $w$-chain and $p_k$ as the *head patient* of the $w$-chain.

4. An *altruistic donor chain* $(p_1, p_2, ..., p_k, d_A)$ with $k \geq 1$, as illustrated in Figure 8, is an ordered list of patients and an altruistic donor $d_A \in A$, such that

$$\begin{aligned} &\mu(p_j) \in \Omega_{p_{j+1}} \quad \forall j \ \in \{1, 2, \ldots, k-1\}, \quad \text{and} \\ &\mu(p_k) = d_A. \end{aligned}$$

$$p_1 \longrightarrow p_2 \longrightarrow \cdots \longrightarrow p_k \longrightarrow d_A$$

**Figure 8:** An altruistic donor chain with $k$ patients represented as a subgraph.

We refer to patient $p_1$ as the *tail patient* of the altruistic donor chain and $p_k$ as the *head patient* of the altruistic donor chain.

For the graph representing matching $\mu$, we only account for maximal $w$-chains and maximal altruistic donor chains. For each $x \in \{w\} \cup A$, a chain $(p_1, \ldots, p_k, x)$ is *maximal* if there is no patient $r \in P \setminus \{p_1, \ldots, p_k\}$ such that $(r, p_1) \in \mathcal{E}$. In a maximal chain, none of patient $p_1$'s paired donors are matched in $\mu$ to any patient in $P$.

The graph $\langle N, \mathcal{E} \rangle$ representing matching $\mu$ is a collection of cycles, unmatched patient chains, maximal $w$-chains, and maximal altruistic donor chains such that each patient appears in exactly one cycle, unmatched patient chain, maximal $w$-chain, or maximal altruistic donor chain.

Let $E_1, E_2, \ldots, E_n$ be the list of cycles, maximal $w$-chains, and maximal altruistic chains in $\langle N, \mathcal{E} \rangle$. Besides its functional representation, we also represent matching $\mu$ as the collection of these subgraphs as

$$\mu = \{E_1, E_2, \ldots, E_n\}.$$

For brevity, we do not include unmatched patient chains in this representation. Thus, if a patient $p$ is not in any of the subgraphs $E_1, E_2, \ldots, E_n$, then they are unmatched in $\mu$, i.e., $\mu(p) = \emptyset$, forming an unmatched patient chain $(p, \emptyset)$ of $\langle N, \mathcal{E} \rangle$.

The number of patients in each cycle, maximal $w$-chain, or maximal altruistic donor chain is referred to as its *size*.[7]

A cycle is an elaborate version of a paired exchange defined in Section 2.1.2. When it has size 2, it is equivalent to the paired exchanges practiced in New England prior to 2004 (see Figure 3). We refer to a cycle with a size $k$ as a *k-way paired exchange*.

A maximal $w$-chain is an elaborate version of a list exchange defined in 2.1.2. When it has size 1, this is precisely a list exchange, as practiced in New England prior to 2004 (see Figure 4), in which the patient receives a priority on the waitlist and, in return, one of their donors donates a kidney to a patient on the waitlist. We refer to a maximal $w$-chain with a size $k$ as a *k-way deceased-donor chain*. For reasons explained in Section 2.4, they are conducted less frequently compared to other forms of kidney exchanges.

Altruistic donor chains were not practiced in 2005, and the initial models we report here do not contain them. In most cases, they can be incorporated into the mechanisms we introduce below — as we explain in Section 2.7. At present, they are among the most common forms of kidney exchange conducted in the U.S. We refer to a maximal altruistic donor chain with a size $k$ as a *k-way altruistic donor chain*.

[7] Another way to define the size of a maximal chain with $k$ patients is to set it to $k+1$ by including the patient without a paired donor who receives a kidney on the waitlist if the central authority uses this donor to donate to the patient on the waitlist immediately. We use the definition in the text that does not account for the patient on the waitlist. The size of a chain becomes only relevant in *open* altruistic donor chains that were initially implemented in the Alliance for Paired Donation and are now being implemented in other programs. In that case, the tail patient's donor is utilized in future iterations of the problem in an altruistic donor chain.

**Maximum exchange and chain sizes and feasible matchings.** The maximum paired exchange size of the problem, denoted by $n^e$, indicates that only paired exchanges with a size of $n^e$ or less are possible. Similarly, the maximum deceased-donor chain size, denoted by $n^d$, indicates that only maximal deceased-donor chains with a size of $n^d$ or less are possible. Likewise, the maximum altruistic donor chain size, denoted by $n^a$, indicates that only maximal altruistic donor chains with a size of $n^a$ or less are possible. We refer to a cycle or chain of size less than or equal to the maximum as *feasible*.

Therefore, a matching $\mu$ is considered *feasible* if each cycle, maximal $w$-chain, and altruistic donor chain in the directed graph representing $\mu$ is feasible. Let $\mathcal{M}$ be the set of feasible matchings. Henceforth, we will simply refer to each feasible matching as a matching.

**Mechanisms.** Typically, patient preferences over medically compatible donors and their paired donor sets are private information.[8] When kidney exchange was first formulated as an application in market design, patient incentives regarding preference manipulation were analyzed under mechanisms provided in Roth, Sönmez, and Ünver (2004, 2005b). Later research also focused on patient incentives in declaring the paired-donor set of the patient truthfully, defined as donor monotonicity (Roth, Sönmez, and Ünver, 2005b; Sönmez and Ünver, 2014, and also see Biró et al., 2023).[9] This property is important for a mechanism because it incentivizes each patient to recruit as many paired donors as possible. As a result, other patients may also receive better transplants through exchange when a patient brings forward their full set of paired donors more donors. We analyze both types of incentives here.[10]

[8]There are various reasons for the existence of private information in this model. The doctor of a patient may possess private information regarding the medical condition of the patient and other donors in the system. Also, a patient with a compatible paired donor may or may not have a higher value for the direct donation from their donor than receiving a marginally better kidney from a stranger. There is also an additional source of private information. In real practice, when a patient remains unmatched, they typically remain in the exchange pool. They expect to be matched in a future iteration of the kidney exchange problem that consists of unmatched patients and their donors in the current matching and new patients and their donors that arrive in the meantime. Therefore, remaining unmatched $\emptyset$ also represents a reserve value for a patient. In expected terms, receiving a relatively inferior living-donor kidney or being matched to the generally inferior deceased-donor kidney on the waitlist may be a worse option than remaining unmatched and waiting for the next kidney exchange run.

[9]Donor monotonicity is an important property in other related exchange problems. Most recently, Han, Kesten, and Ünver (2021) formalized donor monotonicity of mechanisms for blood allocation with replacement donors, an environment with multi-unit demand for each patient. In many countries, patients who need blood transfusion also need to bring forward their relatives who are willing to donate blood in return.

[10]Later consensus among practitioners and researchers pinpoints that strategy-proofness for individual patients is not paramount for the success of a kidney exchange mechanism unless it has obvious vulnerabilities. Therefore, later research focused on creating correct incentives for transplant centers to participate

Fix a patient set $P$, their paired-donor profile $\Omega$, a maximum exchange size $n^e$, a maximum deceased donor chain size $n^d$, and a maximum altruistic donor chain size $n^a$. For each patient $p \in P$, let $\mathcal{P}_p$ be the set of preference relations for patient $p$. Let $\mathcal{P} = \bigtimes_{p \in P} \mathcal{P}_p$ be the set of preference profiles. For each patient $p \in P$, let $2^{\Omega_p}$ be the power set of their paired-donor set $\Omega_p$. Let $\Gamma = \bigtimes_{p \in P} 2^{\Omega_p}$. We refer to $\left[P,\ \Gamma,\ A,\ \mathcal{P},\ n^e,\ n^d, ; n^a\right]$ as a *kidney exchange environment*.

Next, fix an environment. Given a donor profile $\Omega^* \in \Gamma$ and preference profile $\succsim \in \mathcal{P}$, the resulting kidney exchange problem is denoted as $[\Omega^*, \succsim]$ and its set of feasible matchings is denoted as $\mathcal{M}[\Omega^*]$.

A kidney exchange *mechanism* is a function $\varphi : \Gamma \times \mathcal{P} \to \mathcal{M}$ that chooses a matching $\varphi[\Omega^*, \succsim] \in \mathcal{M}[\Omega^*]$ for each problem $[\Omega^*, \succsim] \in \Gamma \times \mathcal{P}$. A mechanism is *individually rational* if, in every problem in the environment, its outcome is individually rational. A mechanism is *Pareto efficient* if its outcome is Pareto efficient in every problem of the environment.

Formally, patients can potentially manipulate a mechanism in two different ways:

1. Patients, in consultation with their doctors, may manipulate their preferences over compatible donors.
2. Patient may not reveal their full set of donors to the system.

A mechanism is $\varphi$ is *immune to preference manipulation* if, for each preference profile $\succsim \in \mathcal{P}$, patient $p \in P$, and alternative preference relation $\succsim'_p \in \mathcal{P}_p$, we have

$$\varphi\Big[\Omega, (\succsim_p, \succsim_{-p})\Big](p) \succsim_p \varphi\Big[\Omega, (\succsim'_p, \succsim_{-p})\Big](p).$$

A mechanism is $\varphi$ is *donor monotonic* if, for each preference profile $\succsim \in \mathcal{P}$, patient $p \in P$, and subset of their donors $\Omega'_p \subsetneq \Omega_p$, we have

$$\varphi\Big[(\Omega_p, \Omega_{-p}), \succsim\Big](p) \succsim_p \varphi\Big[(\Omega'_p, \Omega_{-p}), \succsim\Big](p).$$

A mechanism $\varphi$ is *strategy-proof* if for each preference profile $\succsim \in \mathcal{P}$, patient $p \in P$, paired-donor set $\Omega'_p \subseteq \Omega_p$, and alternative preference relation $\succsim'_p \in \mathcal{P}_p$, we have

$$\varphi\Big[(\Omega_p, \Omega_{-p}), (\succsim_p, \succsim_{-p})\Big](p) \succsim_p \varphi\Big[(\Omega'_p, \Omega_{-p}), (\succsim'_p, \succsim_{-p})\Big](p).$$

We are ready to continue with the first kidney exchange model, the model in Roth, Sönmez, and Ünver (2004).

---

truthfully in exchange programs and incentivizing patients with compatible donors to participate. We discuss these scholarly contributions in later subsections.

## 2.3 The Initial Model and Top-Trading Cycles and Chains Mechanism

In the initial kidney exchange model of Roth, Sönmez, and Ünver (2004), there is no limit on maximum paired exchange and deceased-donor chain size (i.e., $n^e = |P|$ and $n^d = |P|$). There are also no altruistic donors (i.e., $A = \emptyset$ and $n^a = 0$, and thus, we drop them from the definition of a problem). Altruistic donor chains were introduced later as a viable option.

Patients have strict preferences over compatible kidneys, remaining unmatched $\emptyset$, and the priority on the waitlist $w$. We denote the preference relation of a patient $p \in P$ simply as $\succ_p$. This assumption is based on the evidence that emerged in the 1990's in Europe (for example, see Opelz, 1997) that the younger age of the donor and the higher number of matches of HLA tissue types between the patient and the donor (i.e., the number of matches among the two A, two B, and two DR HLAs of the patient and the donor) have a direct relation with the longevity of a transplanted kidney.

Although priority in the waitlist $w$ could be an acceptable option in patient preferences, only some patients might choose it. This is because deceased-donor kidneys are generally considered inferior to living donor kidneys, and the reserve value of remaining unmatched kidneys allows the patient to participate in a future exchange run

We refer to the problem $\left[P,\ \Omega,\ \succ,\ n^e = |P|,\ n^d = |P|\right]$ as a *baseline kidney exchange problem with strict preferences*.

Let $\mathcal{P}^S$ be the set of all allowed strict patient preference profiles.

Next, we introduce a class of mechanisms proposed by Roth, Sönmez, and Ünver (2004), which relies on an iterative algorithm that generates a series of directed graphs to determine its outcome. In the absence of any chains, this algorithm reduces to Gale's top trading cycles algorithm (Shapley and Scarf, 1974), as discussed in Chapter 1, Section 2. In its full generality, with both cycles and chains, this class of mechanisms is inspired by and generalizes the "You-Request-My-House I-Get-Your-Turn" (YRMH-IGYT) mechanism (Abdulkadiroğlu and Sönmez, 1999), as covered in Chapter 1, Section 3.

Before we formally define the mechanism, we introduce some additional concepts.

In the graphs we generate in the algorithm, we determine the matches using the concepts we defined in Section 2.2, namely cycles, unmatched patient chains, and $w$-chains.[11]

[11] However, in the algorithm, these will be tentatively determined at the pointing stage and only be part of the outcome matching when they are removed or fixed. Unlike in the graphs representing matchings, we will not necessarily consider maximal $w$-chains chains. Any size $w$-chain may be relevant to the functioning of the algorithm.

In each step of the algorithm, each remaining patient points to their favorite remaining option if that is $\emptyset$ or $w$. If their favorite remaining option is the paired donor of a remaining patient, they point to this patient. In the resulting directed graph, if any cycle or any unmatched patient chain does not exist, the concept of $w$-chains are utilized to find the patient matches. In such a case, each patient will be a tail patient of a $w$-chain. However, such chains can intersect, i.e., a patient may be pointed to by one or more patients. See Figure 9 for such a scenario.

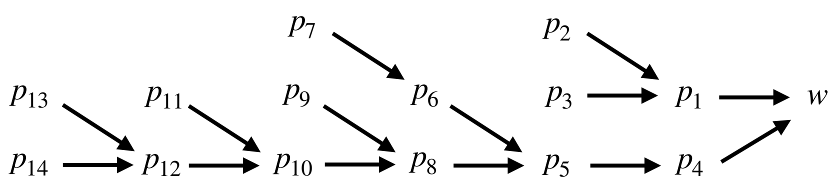



**Figure 9:** Each patient is a tail patient of a $w$-chain.

In such a case, we need a *chain selection* rule to choose which patients will receive their favorite option in this graph.

As an example, in the example in Figure 9, the chain selection rule chooses the $w$-chain with tail patient $p_{10}$, which is $(p_{10}, p_8, p_5, p_4, w)$, as depicted in Figure 10.

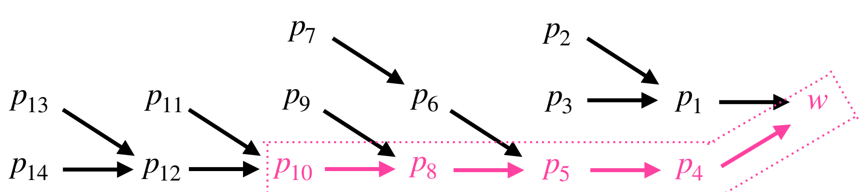



**Figure 10:** As an example, the $w$-chain with tail patient $p_{10}$, $(p_{10}, p_8, p_5, p_4, w)$, is chosen by the chain selection rule.

After a chain is chosen, the selection rule also indicates whether the chain is removed or fixed:

- If the chain is *removed*: Each patient in the $w$-chain and their paired donors are removed from the problem.
- The chain is *fixed*: The chain is preserved so that each patient in the $w$-chain continues pointing to the option they are currently pointing to in every future step of the algorithm unless it is removed as part of a larger $w$-chain.

Each chain selection rule induces a mechanism in this class.

We are now ready to formally introduce the class of mechanisms by Roth, Sönmez, and Ünver (2004). Fix a chain selection rule.

**Top-Trading Cycles and Chains (TTCC) mechanism induced by a given chain selection rule.**

**Step 0.** Initially, all patients and their paired donors are available, and no $w$-chain is fixed.

**Step k.** ($k \geq 1$) Construct a directed graph with the remaining patients in the problem and their paired donors, option $\emptyset$, and option $w$ as follows:

- The previously fixed $w$-chains, if there are any, become part of the graph.
- Each available patient points to their best available option if this option is $w$ or $\emptyset$. Otherwise, they point to the paired-patient of their favorite available donor.

In this graph, either a cycle exists, or an unmatched patient chain exists, or each available patient is the tail patient of a $w$-chain.

**Step k.a.** If there is a cycle or an unmatched patient chain, then

- each patient in a cycle and their paired donors are removed from the problem. Each patient is matched with their favorite donor of the patient they are pointing to, and
- each patient in an unmatched patient chain and their paired donors are removed from the problem. The patient is left unmatched.

We continue with Step k+1.

**Step k.b.** If there is no cycle or unmatched patient chain, we select a $w$-chain using the given chain selection rule.

- If the chosen $w$-chain is removed, then each patient in the chain is matched with their favorite donor of the patient they are pointing to, except the head patient, who is matched with $w$.
- If the chosen $w$-chain is fixed, then each patient in the chain except the tail patient and their paired donors are deemed unavailable. The tail patient of this chain and their paired donors remain available.

Terminate the algorithm if no patient is left in the problem or each available patient is a tail patient of a fixed $w$-chain. In the latter case, each patient, except the head patient of a $w$-chain, is matched with their favorite donor of the patient they are pointing to. The head patient of each $w$-chain is matched with option $w$.

Otherwise, continue with Step k+1.

■

Observe that fixed $w$-chains can grow in the algorithm as other patients can point to the tail patient of a previously fixed chain.

It is easy to see that, regardless of the chain selection rule, TTCC mechanism is individually rational.

Roth, Sönmez, and Ünver (2004) consider the following chain-selection rules.

1. Choose the shortest $w$-chain (subject to a tie-breaker) and remove it.
2. Prioritize patients using a priority order. Choose the $w$-chain starting with the highest-priority available patient and remove it.
3. Prioritize patients using a priority order. Choose the $w$-chain starting with the highest-priority available patient and fix it.
4. Choose the longest $w$-chain (subject to a tie-breaker) and remove it.
5. Choose the longest $w$-chain (subject to a tie-breaker) and fix it.

Although the formal results in Roth, Sönmez, and Ünver (2004) pertain to environments where each patient has a single paired donor, the following two theorems are direct extensions of their results, relying on analogous proofs.

**Theorem 1 (Roth, Sönmez, and Ünver, 2004)** *Fix a baseline kidney exchange environment with strict preferences* $\left[P,\ \Gamma,\ \mathcal{P}^S,\ n^e = |P|,\ n^d = |P|\right]$*. For any chain selection rule that only fixes w-chains but do not remove them, the resulting TTCC mechanism is Pareto efficient.*

Observe that Pareto efficiency in this model is only a welfare measure regarding patients in $P$. When a $w$-chain is removed, a donor of the tail patient donates to the waitlist, and a patient on the waitlist benefits from this removal.

**Theorem 2 (Roth, Sönmez, and Ünver, 2004)** *In a baseline kidney exchange environment with strict preferences* $\left[P,\ \Gamma,\ \mathcal{P}^S,\ n^e = |P|,\ n^d = |P|\right]$*, the TTCC mechanisms with chain selection rules 1,2,3 are immune to preference manipulation, while those with chain selection rules 4,5 are not.*

Krishna and Wang (2007) showed that, under the chain selection rule 3, the TTCC mechanism is equivalent to a version of the YRMH-IGYT mechanism (Abdulkadiroğlu and Sönmez, 1999) in a housing allocation problem with existing tenants with $|P|$ copies of $w$ option.

The following theorem, which can be proven using a related result in Biró et al. (2023), considers donor revelation incentives in environments where option $w$ is unacceptable for each patient. Deceased-donor chains are not widely implemented worldwide; thus, this may represent a viable environment. In this scenario, each TTCC mechanism is equivalent to Gale's TTC algorithm (Shapley and Scarf, 1974). Consequently, in the absence of deceased-donor chains, we refer to this mechanism as the *TTC mechanism*

**Theorem 3 (Biró et al., 2023)** *In a baseline kidney exchange environment with strict pref-*

*erences and no deceased-donor chain possibility $\left[P,\ \Gamma,\ \mathcal{P}^S,\ n^e = |P|,\ n^d = 0\right]$, the TTC mechanism is donor monotonic.*

Moreover, Roth, Sönmez, and Ünver (2003) (the working paper version of the initial kidney exchange paper) extends a result by Ma (1994) for the housing market to this domain when a patient may have multiple donors.[12]

**Theorem 4 (Ma, 1994; Roth, Sönmez, and Ünver, 2003)** *In a baseline kidney exchange environment with strict preferences and no deceased-donor chain possibility $\left[P,\ \Gamma,\ \mathcal{P}^S,\ n^e = |P|,\ n^d = 0\right]$, the TTC mechanism is the only mechanism that is individually rational, Pareto efficient, and immune to preference manipulation.*

Theorems 3 and 4 imply the following corollary:

**Corollary 1** *In a baseline kidney exchange environment with strict preferences and no deceased-donor chain possibility $\left[P,\ \Gamma,\ \mathcal{P}^S,\ n^e = |P|,\ n^d = 0\right]$, the TTC mechanism is strategy-proof.*

## 2.4 Forging a Partnership Between Market Designers and Transplant Surgeons in Creating the New England Kidney Exchange Program

Following the circulation of working paper draft of Roth, Sönmez, and Ünver (2004) in Fall 2003, the authors contacted Dr. Francis Delmonico, the head of New England Organ Bank. Finding the formal approach in the paper intriguing, Delmonico requested the following three modifications from the market designers to their model and proposed mechanism as a prerequisite for implementation (also see Sönmez and Ünver, 2023):

***Only use two-way paired exchanges***. An important ethical and institutional paradigm that dictates kidney exchange is that all operations in a paired exchange—both for the patients and their donors—should be carried out simultaneously. This recommendation is made by the medical consensus statement of Abecassis et al. (2000), which still regulates kidney exchange globally. If exchange transplants are conducted sequentially, the paired donor of a patient who has already received a transplant may become ineligible to donate later or may change their mind. Consequently, a patient involved in the planned paired exchange may not have received a transplant yet, while their paired donor has already donated. In this scenario, the reneging donor puts this patient at risk.

[12]They also employ a different proof technique from Ma, presented by Sönmez (1999b) in a more abstract setting.

Thus, Delmonico recommended starting with the easiest form of paired exchanges to organize: two-way exchanges. This was also the form of paired exchanges implemented in New England prior to 2004.

***Deceased-donor chains should be ruled out.*** Within the medical community, concerns arose about the potential negative consequences of list exchanges for patients with blood type $O$ who are awaiting deceased-donor kidneys. In a list exchange, the patient with a paired donor often has a difficult-to-match blood type, such as blood type $O$. Consequently, when they receive priority on the waitlist, they are matched with an $O$ deceased donor. However, their paired donor, who donates to the waitlist, typically has a blood type such as $A$, $B$, or $AB$, which is not as highly sought after. Consequently, the waitlist effectively loses an incoming $O$ deceased-donor kidney. Even the scholars who proposed the idea of list exchange, Ross and Woodle (2000), acknowledged this potential ethical dilemma and recommended utilizing list exchanges only for patient-donor pairs who are tissue-type incompatible.

An additional ethical concern was that the head patient of a deceased-donor chain does not receive a kidney comparable in-kind to the donor they provide. Instead, they are matched with a less desirable deceased donor, typically with, on average, half the survival duration of a living donor, even though they provide a living donor herself. Some in the medical community found this ethically questionable, even if the patient is willing to accept this option.

At the time, list exchange was mainly practiced in New England and only in cases where a patient could not be matched in a paired exchange. As the simulations using the TTCC mechanism were already showing substantial gains, it was best not to bother anymore with this ethically more questionable form of kidney exchange, according to Delmonico.

***The patients should not be able to rank and choose compatible kidneys***. Most evidence in the US pointed out that almost all compatible donor kidneys have similar longevity (Gjertson and Cecka, 2000; Delmonico, 2004). Moreover, ranking compatible donors to determine a weak or strict preference list was not considered an ethical practice by Delmonico.

As a negative side effect of Delmonico's modeling preferences, by only considering indifferences among compatible kidneys, compatible pairs no longer had any incentive to participate in the exchange.

In response to these requests by Delmonico as prerequisites for partnering with their team of market designers, Roth, Sönmez, and Ünver (2005b) formulated a model with the following features:

- Only two-way paired exchanges are allowed.

- Patients are indifferent over compatible donors.
- Patients with a compatible donor do not participate in the exchange.
- Deceased-donor chains are not allowed.

Thanks to Roth, Sönmez, and Ünver (2005b), Dr. Francis Delmonico and the New England Organ Bank agreed to adopt a priority mechanism incorporating only two-way exchanges without any deceased-donor chains in an agreement reached in 2005 with the market designers Alvin Roth, Tayfun Sönmez, and Utku Ünver. Patients with compatible pairs did not participate in exchanges from the start. As a result, the market designers and Francis Delmonico founded the New England Program for Kidney Exchange (NEPKE) (cf. Roth, Sönmez, and Ünver, 2005a).

## 2.5 Two-way Kidney Exchange with Compatibility-based Preferences

In this subsection, we present the model of Roth, Sönmez, and Ünver (2005b), which paved the way for NEPKE, the world's first kidney exchange program utilizing tools from market design and optimization. This model's game theoretic background and its relationship to graph theory were covered in Chapter 1, Section 2 as the bilateral exchange model with compatibility-based preferences. Here, we introduce its more applied version and discuss its relevant features. Roth, Sönmez, and Ünver (2005b) built this model by generalizing a paper by Bogomolnaia and Moulin (2004) on two-sided matching and proposing a practical deterministic mechanism that became the building block of New England's kidney exchange mechanism.

A *two-way* (or *pairwise*) *kidney exchange problem with compatibility-based preferences* is a list $\left[P,\ \Omega,\ \succsim^C,\ n^e = 2\right]$ such that

- There are no altruistic donors (i.e., $A = \emptyset$), and deceased-donor chains are not allowed. As a result, both parameters $n^a$ and $n^d$ are set to 0 and omitted from the problem.
- For each patient $p \in P$, the relation $\succsim^C_p$ represents a *compatibility-based* (or *dichotomous*) preference relation over paired donors $\cup_{p \in P} \Omega_p$ and remaining unmatched $\emptyset$.[13]

These preferences lead to three indifference classes

1. $\boldsymbol{C}[\succsim^C_p] = \left\{ d \in \bigcup_{r \in P} \Omega_r \ : \ d \ \succ_p \emptyset \right\}$,
2. $\{\emptyset\}$, and
3. $\boldsymbol{I}[\succsim^C_p] = \left(\bigcup_{r \in P} \Omega_r\right) \setminus \boldsymbol{C}[\succsim^C_p]$,

[13]Since deceased-donor chains are not allowed, option $w$ is not feasible. Alternatively, we could define option $w$ to fit this model exactly into the general model's framework in Section 2.2 and declare that $w$ is unacceptable for each patient.

with the following properties:

1. For each $d \in \boldsymbol{C}\big[\succsim_p^C\big]$ and $d' \in \boldsymbol{I}\big[\succsim_p^C\big]$, $\qquad d \succ_p \emptyset \succ_p d'$.
2. For each $d, d' \in \boldsymbol{C}\big[\succsim_p^C\big]$, $\qquad d \sim_p d'$.
3. For each $d, d' \in \boldsymbol{I}\big[\succsim_p^C\big]$, $\qquad d \sim_p d'$.

Let $\mathcal{P}^C$ be the set of compatibility-based preference profiles. $\Big[P, \Gamma, \mathcal{P}^C, n^e = 2\Big]$ denotes a *two-way kidney exchange environment with compatibility-based preferences*.

Let $\mathcal{M}$ be the set of matchings. Since $n^e = 2$ and deceased-donor chains are not feasible, each matching has a representative graph consisting only of two-way exchanges.

Let $\pi$ be a priority order over patients. For notational convenience, relabel the patients such that

$$\pi = p_1 - p_2 - \ldots - p_{|P|}.$$

We next present the mechanism adopted by NEPKE, the *priority mechanism* (Roth, Sönmez, and Ünver, 2005b), which is equivalent to the matching matroid greedy algorithm as discussed in Chapter 1, Section 2.

**Priority mechanism induced by $\pi$.**

**Step 0.** Let $\mathcal{M}^0 \subseteq \mathcal{M}$ be the set of all individually rational matchings.

**Step k.** (k $\geq$ 1) If there exists some matching $\mu \in \mathcal{M}^{k-1}$ such that $\mu(p_k) \neq \emptyset$ then let

$$\mathcal{M}^k = \{\mu \in \mathcal{M}^{k-1} : \mu(p_k) \neq \emptyset\},$$

otherwise, let

$$\mathcal{M}^k = \mathcal{M}^{k+1}.$$

Terminate the algorithm if $k = |P|$. Otherwise, continue with Step k+1.

■

The algorithm's outcome is a matching in set $\mathcal{M}^{|P|}$. Each matching in this set matches the same set of patients that we refer to as $I^\pi \subseteq I$.

Roth, Sönmez, and Ünver (2005b) also shows that the set $I^\pi$ can be constructed in polynomial time, without the need of constructing the sequence of matching sets $\mathcal{M}^0, \mathcal{M}^1, \ldots, \mathcal{M}^{|P|}$ (also see Chapter 1, Section 2).

The following result shows a close connection between the maximum number of transplants achievable and Pareto efficiency:

**Proposition 1 (Roth, Sönmez, and Ünver, 2005b)** *In a two-way exchange problem with compatibility-based preferences, the same number of patients receive a transplant in every Pareto-efficient and individually rational matching.*

One way to maximize the number of transplants with only two way exchanges is through priority mechanisms.

**Theorem 5 (Roth, Sönmez, and Ünver, 2005b)** *In a two-way exchange problem with compatibility-based preferences, any priority mechanism is Pareto efficient and individually rational.*

Moreover, incentives are aligned well in the priority mechanisms for patients to reveal both their preferences and paired-donor sets truthfully:

**Theorem 6 (Roth, Sönmez, and Ünver, 2005b)** *In a two-way exchange environment with compatibility-based preferences, any priority mechanism is strategy-proof.*

Importantly, however, Proposition 1 no longer holds when three-way exchanges are also allowed, as the following example demonstrates:

**Example 1** *Consider a scenario where the maximum exchange size is $n^e = 3$. Let's examine a problem with 4 patients, denoted as $p_1, p_2, p_3, p_4$. Each patient $p_k$ is paired with a single donor $d_k$. Patient $p_1$ has blood type A, and donor $d_1$ has blood type O, but they are medically incompatible due to tissue-type incompatibility. Patient $p_1$ is tissue-type incompatible with all other donors, and no other patient has tissue-type incompatibility with any of the donors. All other patients, however, are blood type incompatible with their paired donors. Specifically, patient $p_2$ has blood type O with a blood-type A paired donor $d_2$, patient $p_3$ has blood type B with a blood-type A paired donor $d_3$, and patient $p_4$ has blood type O with a blood-type B paired donor $d_4$.*

*For this problem, the set of compatible donors for each patient is given as follows:*

$$\boldsymbol{C}\left[\succsim^C_{p_1}\right] = \{d_2, d_3\}, \qquad \boldsymbol{C}\left[\succsim^C_{p_2}\right] = \{d_1\},$$
$$\boldsymbol{C}\left[\succsim^C_{p_3}\right] = \{d_4\}, \qquad \boldsymbol{C}\left[\succsim^C_{p_4}\right] = \{d_1\}.$$

*There are two Pareto-efficient and individually rational matchings, given as*

$$\mu_1 = \Big\{(p_1, p_2)\Big\} \quad \text{and} \quad \mu_2 = \Big\{(p_1, p_3, p_4)\Big\}.$$

*Here, $\mu_1$ consists of a two-way exchange, $(p_1, p_2)$, such that each patient is matched with the paired donor of the other, and $\mu_2$ consists of a three-way exchange, $(p_1, p_3, p_4)$ in which patient $p_1$ is matched with donor $d_3$, patient $p_3$ is matched with donor $d_4$, and patient $p_4$ is matched with donor of $d_1$. Thus, $\mu_1$ matches 2 patients, and $\mu_2$ matches 3 patients.*

Example 1 underscores that, there may be value for kidney exchange programs to develop the logistical capacity to perform three-way or larger exchanges.

## 2.6 The Significance of Three-way Kidney Exchange

While optimizing the number of transplants through two-way kidney exchanges represents a significant milestone in the development of kidney exchange clearinghouses, the discovery of scenarios akin to Example 1 during a NEPKE match run revealed that, even in problems with a large number of patients, there might be substantial marginal gains from the availability of three-way exchanges (Roth, Sönmez, and Ünver, 2007). We next discuss this model.

Assuming there are no altruistic donors (i.e., $A = \emptyset$), we consider a kidney exchange problem $\left[P,\ \Omega^1,\ \succsim^C,\ n^e\right]$ for some maximum exchange size $n^e \geq 2$, such that:

- The preference profile $\succsim^C \in \mathcal{P}^C$ is compatibility-based, and deceased-donor chains and altruistic donor chains are not feasible (i.e., $n^d = n^a = 0$).
- Each patient $p \in P$ is paired with a single donor, so that $\left|\Omega^1_p\right| = 1$.
- There are no patients with compatible paired donors.

In this model, assuming away tissue-type incompatibility, we rely solely on blood-type compatibility for kidney transplants to set an upper bound on the number of exchanges that can be employed for different $n^e$.

When describing a patient $p$ and their paired donor $d_p$, we represent them by their blood types as $X - Y$, where:

- $X$ denotes the blood type of patient $p$,
- $Y$ denotes the blood type of donor $d_p$.

We denote the number of type $X - Y$ patient-donor pairs in the problem as $\#(X - Y)$.

To establish the desired upper bounds, we make four assumptions. The first assumption enables us to focus on only the blood types of patients and their donors.

**Assumption 1 (Upper bound assumption)** *No patient is tissue-type incompatible with another patient's donor.*

Recall that there are four blood types: type $O$, which can donate to all types; types $A$ and $B$, which can donate to their own type and type $AB$; and type $AB$, which can only donate to type $AB$. Thus, a patient with a blood-type compatible donor participates in the exchange only when their donor is tissue-type incompatible with them. On average, this occurs with a frequency of $10 - 15\%$ given a random patient and a random donor. On the other hand, a patient with a blood-type incompatible donor always needs to participate in an exchange to receive a transplant. Therefore, the following assumption typically holds in a relatively large patient set $P$:

**Assumption 2 (Large Population of Incompatible Patient-Donor Pairs)** *Regardless of the maximum number of pairs allowed in each exchange, pairs of types $O - A$, $O - B$, $O - AB$, $A - AB$, and $B - AB$ are on the long side of the exchange in the sense that at least one pair of each type remains unmatched in each individually rational matching.*

Based on this assumption of a large population, we refer to pair types $O - A$,

$O-B$, $O-AB$, $A-AB$, and $B-AB$ as *underdemanded types*. Conversely, we refer to their *reciprocal types* $A-O$, $B-O$, $AB-O$, $AB-A$, and $AB-B$ as *overdemanded types*.

**An Upper-bound to the Number of Patients Matched under Two-way Exchanges.** We first establish an upper bound to the number of patients who can benefit from two-way exchanges under Assumptions 1 and 2.

**Proposition 2 (Roth, Sönmez, and Ünver, 2007)** *Consider a kidney exchange problem* $\left[P, \Omega^1, \succsim^C, n^e\right]$ *with compatibility-based preferences obeying Assumptions 1 and 2. Suppose the maximum paired exchange size is* $n^e = 2$*. Then, the maximum number of patients who can be matched under a feasible matching is:*

$$\begin{aligned}
&2\Big(\#(A\text{-}O) + \#(B-O) + \#(AB-O) + \#(AB-A) + \#(AB-B)\Big) \\
&+\Big(\#(A-B) + \#(B-A) - \Big|\#(A-B) - \#(B-A)\Big|\Big) \\
&+2\left(\left\lfloor \frac{\#(A-A)}{2} \right\rfloor + \left\lfloor \frac{\#(B-B)}{2} \right\rfloor + \left\lfloor \frac{\#(O-O)}{2} \right\rfloor + \left\lfloor \frac{\#(AB-AB)}{2} \right\rfloor\right).
\end{aligned}$$

This result is easy to see, as the highest benefit from two-way exchanges can be obtained by matching $X-Y$ type pairs with their opposite $Y-X$ pairs. Thus, patients with the same blood type as their donors—tissue-type incompatible with their own donors but not with any other donor according to Assumption 1—can all be matched if their number is even; otherwise, one patient remains unmatched (hence, we use the integer floor operator $\lfloor \cdot \rfloor$ in the expression).

**Gains from Three-way Exchanges.** As illustrated earlier by Example 1, matching $X-Y$ type pairs with their opposite $Y-X$ pairs may not maximize the number of matched patients when three-way exchanges are allowed in addition to two-way exchanges (i.e., $n^e = 3$). We reinforce this observation with a more elaborate example:

**Example 2** *Consider a population of 17 incompatible patient-donor pairs. There are 11 pairs of patients who are blood-type incompatible with their donors, with types*

$$O-A,\ O-A,\ O-B,\ A-B,\ A-B,\ A-B,\ A-B,\ B-A,\ A-AB,\ B-AB, \text{ and } B-AB,$$

*and, 6 pairs who are tissue-type incompatible with their donors, with types*

$$A-A,\ A-A,\ A-A,\ B-O,\ AB-O, \text{ and } AB-A.$$

*If only two-way exchanges are possible, i.e.,* $n^e = 2$*, at most 10 patients could be matched in* *three two-way paired exchanges written in terms of the blood types of pairs as*

$$\mu_{\ n^e=2} = \Big\{(A-B, B-A),\ (A-A, A-A),\ (B-O, O-B),\ (AB-A, A-AB),\ (AB-O, O-A)\Big\}.$$

*If three-way exchanges are also feasible (i.e., $n^e = 3$), 14 out of 17 patients could be matched in 3 paired exchanges as:*[14]

$$\mu_{\ n^e=3} = \left\{ \begin{array}{c} (A-B, B-A),\ (A-A, A-A, A-A),\ (B-O, O-A, A-B), \\ (AB-A, A-B, B-AB),\ (AB-O, O-B, B-AB) \end{array} \right\}.$$

*Observe that three-way exchanges allow:*

1. *an odd number of $A-A$ pairs to be transplanted (instead of only an even number with two-way exchanges),*
2. *a $B-O$ and an $AB-A$ pair each to facilitate three transplants rather than only two, given that there are more $A-B$ pairs than $B-A$,*
3. *an $AB-O$ pair to facilitate three transplants rather than only two.*

*This helps an additional 4 patients receive a transplant.*

The insight of the example can be generalized to establish the marginal impact of overdemanded type pairs when three-way exchanges are also feasible (see Figure 11).

We make two additional assumptions to present the next result. Assumption 3 is not necessary for any result; it is assumed simply for notational convenience. Assumption 4 assumes away some very rare situations, also simplifying the expressions for upper bounds.

**Assumption 3** $\#(A-B) \geq \#(B-A)$.

**Assumption 4** *There is either no type $A-A$ pair or there are at least two of them. The same is also true for each of the types $B-B$, $AB-AB$, and $O-O$.*

**Proposition 3 (Roth, Sönmez, and Ünver, 2007)** *Consider a kidney exchange problem $\left[P,\ \Omega^1,\ \succsim^C,\ n^e\right]$ with compatibility-based preferences obeying Assumptions 1-4. Suppose the maximum paired exchange size is $n^e = 3$. Then, the maximum number of patients who*

[14]Here donation direction is rightward for notational convenience, as opposed to leftward in our algorithmic graphs consisting of only patients.

*can be matched under a feasible matching is:*

$$\begin{aligned}
&2\Big(\#(A-O)+\#(B-O)+\#(AB-O)+\#(AB-A)+\#(AB-B)\Big)\\
&+\Big(\#(A-B)+\#(B-A)-\Big|\#(A-B)-\#(B-A)\Big|\Big)\\
&+\Big(\#(A-A)+\#(B-B)+\#(O-O)+\#(AB-AB)\Big)\\
&+\#(AB-O)\\
&+\min\Big\{\Big(\#(A-B)-\#(B-A)\Big),\Big(\#(B-O)+\#(AB-A)\Big)\Big\}.
\end{aligned}$$

Under Assumptions 1 and 4 only, compared to using only two-way exchanges, the marginal impact of utilizing three-way exchanges is as follows:

I. An $AB-O$ pair helps us to match one extra underdemanded type pair in one of two alternative ways (see Column I of Figure 11).

IIa. For each $A-B$ pair in excess of $B-A$ pairs, a $B-O$ or $AB-A$ pair helps us match this pair additionally (see Column II of Figure 11).

IIb. For each $B-A$ pair in excess of $A-B$ pairs, on the other hand, an $A-O$ and $AB-B$ pair helps us match this pair additionally (see Column III of Figure 11).

Even though $AB-O$ type pairs are rarely seen in real-life applications, this impact would rarely diminish even in a large market, as there is often an empirical difference between numbers of $A-B$ and $B-A$ pairs.[15]

[15] The reasons for this asymmetry include differences in the prevalence of these two blood types among different races, variations in exposure to kidney disease across racial groups, and the fact that patients often have multiple donors. For the case of US, some relatively large hospital samples show that $A-B$ pairs are observed more frequently than $B-A$ pairs.

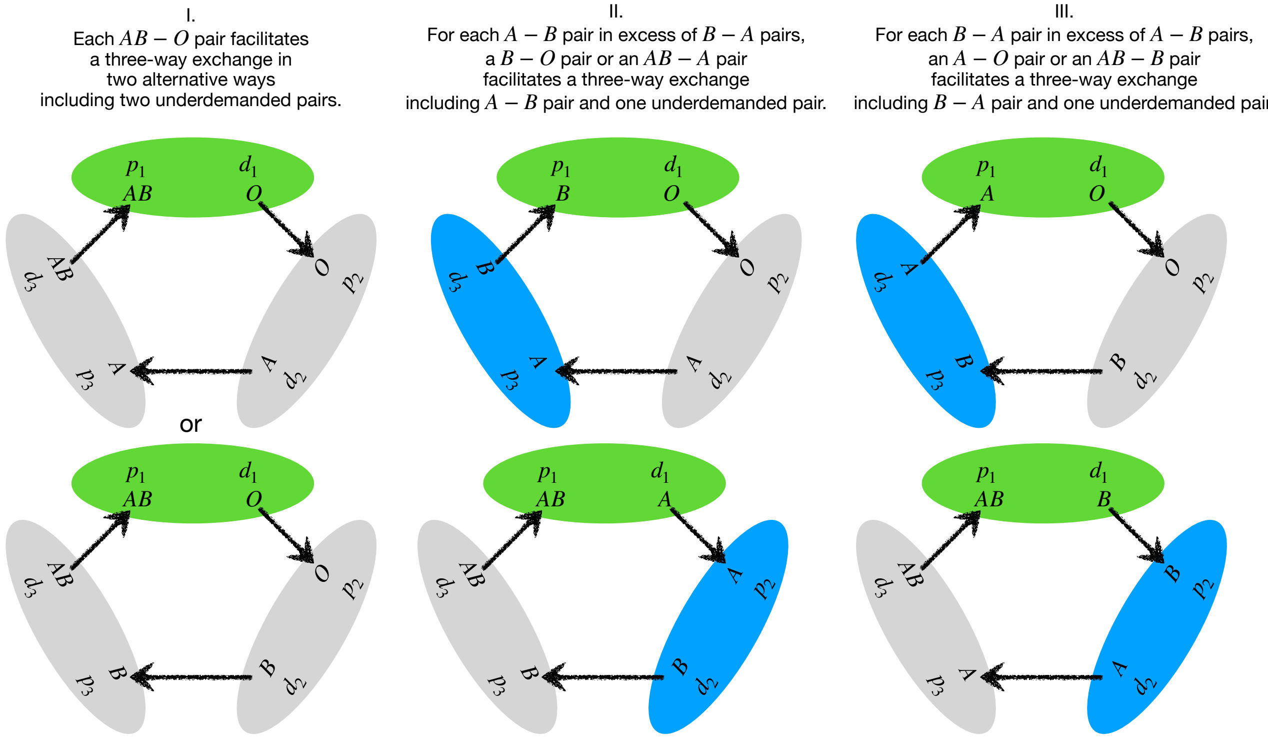



**Figure 11:** Possible gains from utilizing three-way paired exchanges.

To summarize, the marginal number of transplants that the availability of three-way kidney exchanges brings over two-way exchanges is given as:

$$\begin{aligned}
&\#(A-A)+\#(B-B)+\#(O-O)+\#(AB-AB)\\
&-2\left(\left\lfloor\frac{\#(A-A)}{2}\right\rfloor+\left\lfloor\frac{\#(B-B)}{2}\right\rfloor+\left\lfloor\frac{\#(O-O)}{2}\right\rfloor+\left\lfloor\frac{\#(AB-AB)}{2}\right\rfloor\right)\\
&+\#(AB-O)\\
&+\min\left\{\Big(\#(A-B)-\#(B-A)\Big),\Big(\#(B-O)+\#(AB-A)\Big)\right\}.
\end{aligned}$$

**Gains from Four-way Exchanges.** The benefits from larger exchanges largely diminish after three-way exchanges. Allowing for four-way exchanges only helps us through $AB - O$ pairs (see Figure 12). As seen in Column I, when not all $A - B$ pairs can be matched through two-way and three-way exchanges, the $AB - O$ pair can help match an additional $A - B$ pair, in addition to the two underdemanded type pairs that can be matched in a three-way exchange. Alternatively, an $AB - O$ pair can also facilitate an additional transplant through a four-way exchange when $B - A$ pairs are in excess (see Column II).

$AB - O$ pairs are usually rare. Not only is $AB$ the rarest blood type (accounting for only around 4% of the population in the US), but $AB - O$ pairs are also blood-type compatible. Therefore, unless they participate in an exchange due to altruistic reasons, their participation is tied to tissue-type incompatibility, which occurs only $10-15\%$ of the time. Hence, the predicted gain from four-way exchanges is marginal in a large exchange pool.

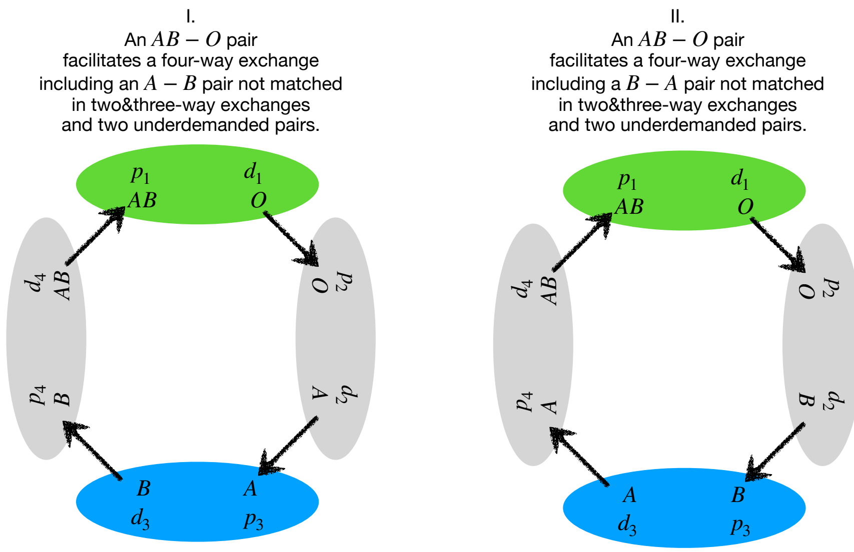



**Figure 12:** Possible gains from utilizing four-way exchanges.

**Proposition 4 (Roth, Sönmez, and Ünver, 2007)** *Consider a kidney exchange problem* $\left[P, \Omega^1, \succsim^C, n^e\right]$ *with compatibility-based preferences obeying Assumptions 1-4. Suppose the maximum paired exchange size is* $n^e = 4$. *Then, the maximum number of patients who can be matched under a feasible matching is:*

$$\begin{aligned}
&2\Big(\#(A-O)+\#(B-O)+\#(AB-O)+\#(AB-A)+\#(AB-B)\Big)\\
&+\Big(\#(A-B)+\#(B-A)-\Big|\#(A-B)-\#(B-A)\Big|\Big)\\
&+\Big(\#(A-A)+\#(B-B)+\#(O-O)+\#(AB-AB)\Big)\\
&+\#(AB-O)\\
&+\min\Big\{\Big(\#(A-B)-\#(B-A)\Big),\\
&\qquad\qquad \Big(\#(B-O)+\#(AB-A)+\#(AB-O)\Big)\Big\}.
\end{aligned}$$

*Therefore, in the absence of tissue-type incompatibilities between patients and other patients' donors, the marginal effect of four-way kidney exchanges is bounded above by the rate of the very rare AB-O type.*

We extend Example 2 to include gains from four-way exchanges

**Example 3** *Consider the problem in Example 2. When $n^e = 4$, we can create the following matching that maximizes the number of patients matched.*

$$\mu_{n^e=4} = \left\{ \begin{array}{c} (A-B, B-A),\ (A-A, A-A, A-A),\ (B-O, O-A, A-B), \\ (AB-A, A-B, B-AB),\ (AB-O, O-A, A-B, B-AB) \end{array} \right\}.$$

*matching 15 patients instead of the 14 matched when $n^e = 3$.*

Finally, under Assumptions 1, 2, and 4, the following theorem shows that there are no further gains from conducting exchanges larger than four-way.

**Theorem 7 (Roth, Sönmez, and Ünver, 2007)** *Consider a kidney exchange problem $\left[P, \Omega^1, \succsim^C, n^e\right]$ with compatibility-based preferences obeying Assumptions 1, 2, and 4. Suppose the maximum paired exchange size is unrestricted, i.e., $n^e = |P|$. Let $\mu$ be any Pareto-efficient matching. Then there exists a Pareto-efficient matching $\nu$, which consists only of two-way, three-way, and four-way exchanges and matches the same patients as matching $\mu$ does.*

As an implication, since any matching that maximizes the number of transplants is Pareto efficient, it is possible to construct another matching that maximizes the number of transplants by matching the same set of patients and utilizing only two-way, three-way, and four-way exchanges.

Subsequent simulations in Roth, Sönmez, and Ünver (2007) and Saidman et al. (2006) showed that in a random population with 100 incompatible patient-donor pairs from the US, assuming each paired donor is blood-type independent from their patient, approximately 49.7% of them could be matched using only two-way exchanges (when $n^e = 2$), while about 59.7% of them can be matched with two- and three-way exchanges (when $n^e = 3$), and 60.35% of them can be matched with two-, three-, and four-way exchanges (when $n^e = 4$). On the other hand, when $n^e = |P|$, so that exchange sizes are not restricted, approximately 60.4% of them can be matched. Hence, three-way exchanges increase the scope of kidney exchange by 20%, while the utilization of larger-size exchanges leads to marginal increases as predicted by the formulas provided in Propositions 2-4 and Theorem 7. Also, the upper-bound formulas provided in the propositions yield close numbers to the simulation averages, especially when $n^e \geq 3$.

### 2.6.1 The Integration of Larger-Size Exchanges to Kidney Exchange

Following Roth, Sönmez, and Ünver (2007) and Saidman et al. (2006) underscoring the importance of 3-way exchanges, NEPKE incorporated three-way paired exchanges in addition to two-way paired exchanges in its mechanism.

Concurrently, an Ohio-based consortium, which had been utilizing an algorithm for organizing kidney exchanges, albeit suboptimal, joined forces with Roth, Sönmez, and Ünver, leading to the establishment of the Alliance for Paired Donation (APD) under the leadership of Dr. Michael Rees (Anderson et al., 2015). From the outset, they implemented three-way exchanges in addition to two-way exchanges, and occasionally pursued four-way exchanges.

The mechanisms designed for NEPKE and APD by the market designers and mechanisms used by numerous other exchange programs can be represented as solutions of an integer program (Roth, Sönmez, and Ünver, 2007; Sönmez and Ünver, 2013; Anderson et al., 2015).

Consider a kidney exchange problem $\left[P, \Omega, \succsim^C, n^e\right]$ where each patient can have multiple paired donors, and the preference profile $\succsim^C \in \mathcal{P}^C$ is compatibility-based. Let $\mathbf{E}_{n^e}$ denote the set of all individually rational paired exchanges with a size of $n^e$ or less. Define $P_E \subseteq P$ as the set of patients matched in any exchange $E \in \mathbf{E}_{n^e}$. We assign a positive weight $W_E > 0$ to each individually rational paired exchange $E \in \mathbf{E}_{n^e}$. The integer programming problem is formulated as:

$$\max_{X \in \{0,1\}^{|\mathbf{E}_{n^e}|}} \sum_{E \in \mathbf{E}_{n^e}} W_E X_E \qquad \text{subject to} \qquad \sum_{E \in \mathbf{E}_{n^e} \,:\, p \in E} X_E \leq 1 \qquad \forall\, p \in P_E \tag{1}$$

A solution vector $X^* = (X^*_E)_{E \in \mathbf{E}_{n^e}}$ is associated with an outcome matching $\mu \in \mathcal{M}$ such that

$$\mu = \left\{E \in \mathbf{E}_{n^e} \;:\; X^*_E = 1\right\}.$$

Exogenous or random tie-breakers can also be employed to choose a unique matching outcome among multiple solutions that may solve this problem.[16]

The NEPKE and APD mechanisms differ in how they set the weight profile $(W_E)_{E \in \mathbf{E}_{n^e}}$. Let $E \in \mathbf{E}_{n^e}$ have size $k$ for some $k \leq n^e$. The exchange weight $W_E$ was determined as follows in NEPKE and APD, respectively:

***NEPKE.*** Each patient $p \in P$ is assigned to a priority tier $t(p) \in \{1, 2, \ldots T\}$ for some positive integer $T$ upper bound. For example, children and patients who had donated one of their kidneys in the past are given a higher priority than the rest. Each priority

[16]This optimization can be solved in polynomial time in the number of variables when $n^e = 2$ by the blossom algorithm of Edmonds (1965) and when $n^e = |P|$ using the maximum flow–minimum cut algorithm of Edmonds and Karp (1972). Otherwise, it is NP-complete (Abraham, Blum, and Sandholm, 2007), i.e., its worst-case solution time is an exponential function of the number of the variables if P $\neq$ NP (using the computer science jargon on algorithmic complexity). In this case, reasonable problem sizes can readily be solved by commercially available or public-domain integer programming software. Abraham, Blum, and Sandholm (2007) and Biró, Manlove, and Rizzi (2009) introduce fast tailored algorithms for the integer programming problem. Anderson et al. (2015) also presents a heuristically fast method that enables solving this problem without constructing set $\mathbf{E}_{n^e}$ in many practical instances.

tier $t$ is associated with a positive weight $W_t$ such that

$$\frac{W_t}{W_{t+1}} > |P| \quad \forall\, t < T.$$

The weight of exchange $E$ is set to

$$W_E = \sum_{p \in P_E} W_{t(p)}.$$

When each patient has a distinct priority tier, any solution matches the same set of patients, and the integer program finds an outcome of a *priority mechanism* for maximum exchange size $n^e$. It is induced by priority order $\pi = p_1 - p_2 - \ldots - p_{|P|}$ where for each $j$, patient $p_j$ is patient in the priority-tier $j$. This mechanism is the generalization of the priority mechanism described in Section 2.4 for maximum paired exchange size 2.

***APD.*** The weight of the $k$-way paired exchange $E$ is determined as

$$W_E = k,$$

which is the number of patients in $E$. In this case, any solution of the integer program (1) corresponds to a matching that maximizes the number of transplants. A tie-breaker was used to choose among multiple matchings that may maximize the number of transplants: a maximum matching that maximizes the sum of a secondary set of patient-specific exogenously given weights was chosen.

We provide additional insights into the solutions of this integer program.

**Theorem 8** *Consider a kidney exchange problem* $\left[P,\ \Omega,\ \succsim^C,\ n^e\right]$ *with compatibility-based preferences. Suppose that the maximum paired exchange size is* $n^e \leq |P|$. *A solution matching of the integer program* (1) *is Pareto efficient for a given profile of positive exchange weights* $(W_E)_{E \in \mathbf{E}_{n^e}}$ *if there exists a positive patient-specific weight profile* $(W_p)_{p \in P}$ *such that for each* $E \in \mathbf{E}_{n^e}$,

$$W_E = \sum_{p \in P_E} W_p.$$

*Conversely, for any Pareto efficient matching* $\mu$, *there exists a profile of exchange weights* $(W_E)_{E \in \mathbf{E}_{n^e}}$ *satisfying the above additivity condition such that all solutions of the integer program* (1) *matches exactly all patients matched in* $\mu$.

The first statement follows because if an outcome matching $\mu$ of the integer program were Pareto dominated by another matching $\nu$, as the set of the patients matched under $\mu$ is a proper subset of patients matched under $\nu$,

$$\sum_{E \in \mu} W_E = \sum_{p \in P \,:\, \mu(p) \neq \emptyset} W_p < \sum_{p \in P \,:\, \nu(p) \neq \emptyset} W_p = \sum_{E \in \nu} W_E,$$

contradicting $\mu$ is an outcome matching of the integer program (1). The second statement can be proven by constructing patient-specific weights as follows for a given Pareto-efficient matching $\mu$: For each $p \in P$, let

$$W_p = \begin{cases} 1 & \text{if } \mu(p) \neq \emptyset \\ \varepsilon & \text{if } \mu(p) = \emptyset \end{cases}$$

such that $0 < \varepsilon < \frac{1}{|P|}$. Then $\mu$ is a solution matching of the integer program (1), as otherwise, any matching $\nu$ that achieves a higher total weight would have to match all patients matched in $\mu$ and at least one additional patient. Matching $\nu$ Pareto dominates $\mu$, leading to a contradiction. Thus, all solutions of the integer program (1) have the same total weight that is achieved by $\mu$. By construction of the weights, then the same patients matched under $\mu$ are matched in all solution matchings.

Therefore, the initially adopted NEPKE and APD mechanisms are both Pareto efficient as NEPKE uses a positive priority-tier-based weight for each patient, APD uses weight 1 for each patient, and the resulting exchange weights are additive of these weights.

As a notable contribution to the kidney exchange model with compatibility-based preferences, Hatfield (2005) finds necessary and sufficient conditions for strategy-proofness of a mechanism in the compatibility-based preference environment when the maximum paired exchange size can be arbitrary. This paper also proves the following result:

**Theorem 9 (Hatfield, 2005)** *Consider a kidney exchange environment* $\left[P, \Omega, \mathcal{P}^C, n^e\right]$ *with compatibility-based preferences. Suppose that the maximum paired exchange size is* $n^e \leq |P|$*. Then, any priority mechanism is strategy-proof.*

## 2.7 Altruistic Donor and Deceased-Donor Chains

Besides adopting a structured, optimization-based matching approach, utilizing three-way exchanges in addition to two-way exchanges has been instrumental in expanding the scope of kidney exchange. Partnerships between market design economists and members of the transplantation community have also paved the way for more elaborate strategies in kidney exchange.

One significant advancement proposed by Roth et al. (2006) introduced the concept of altruistic donor chains, where transplants could be organized sequentially: The head patient of the chain would first receive a transplant from an altruistic donor.

Then, a paired donor of this patient would donate to the second patient in the chain. Subsequently, a paired donor of the second patient would donate to the third, and so on. In this chain, no patient and their paired donor would be harmed in case some other donor reneges, as their paired donor would not donate before the patient receives a transplant. Even if a donor reneges and opts out of the transplant, the remaining patients in the chain would retain their paired donors for a future exchange.

One might question the scope of this innovation, given that altruistic donors are perceived to be rare due to the costly nature of kidney donation. However, as the prevalence of kidney exchanges increased worldwide, more individuals in the US started stepping forward as altruistic donors, providing an opportunity to combine altruistic donations with kidney exchanges.

Traditionally, though not required by legislation, the gift of a kidney from altruistic donors would be treated similarly to deceased donors: their donated kidney would go to a patient on the deceased-donor waitlist determined by the point system utilized for the waitlist. However, there is ample reason to consider establishing altruistic donor chains akin to those proposed for deceased-donor chains by Roth, Sönmez, and Ünver (2004). Moreover, these chains could be perceived as ethically more acceptable, primarily because altruistic donors aren't subject to the same legislative allocation rules governing deceased-donor kidneys. Additionally, all patients within the chain, including the initial recipient, would receive a comparable living donor—an advantage not present in a deceased-donor chain. Notably, case studies and simulation work exploring simultaneous altruistic donor chains were documented by Montgomery et al. (2006), termed "domino paired donation". In this model, the donor of the last recipient contributes back to the waitlist.

The altruistic donor chain expands if the tail donor's kidney does not immediately return to the waitlist and the tail donor waits to initiate a new chain in a future iteration of the problem. Such a tail donor is called a "bridge donor". Yet, the risk lies in any donor in a chain reneging on their promise. Dr. Mike Rees and APD implemented this concept (Anderson et al., 2015).

There is another upside to utilizing altruistic donors to initiate chains. The idea of helping several patients may also increase the value of the gift for the donor and potentially increase the number of altruistic donors.

All mechanisms we introduced so far and the integer program in (1), can incorporate altruistic donors. Also, all corresponding results except Theorems 3 and 4 and Corollary 1 hold in the specified environments below. Let $A \neq \emptyset$ and $n^a > 0$.

1. The TTCC mechanism for an environment with strict preferences and unrestricted cycle and chain sizes $\Big[P,\ \Omega,\ A,\ \mathcal{P}^{S,w},\ n^e = |P|,\ n^d = |P|,\ n^a = |P|\Big]$:

The TTCC mechanism's definition is amended as follows:

- Every "$w$-chain" is replaced by "$w$-chain or altruistic donor chain" in the definition.
- When an altruistic donor chain is removed, its altruistic donor is removed from the problem and assigned to the head patient.
- When an altruistic donor chain is fixed, the altruistic donor in the chain is deemed unavailable, and upon the termination of the algorithm, they are assigned to the head patient of the altruistic donor chain.

2. The priority mechanism for environment $\left[P,\ \Omega,\ A,\ \mathcal{P}^C,\ n^e = 2,\ n^a = 1\right]$: This environment is the simplest exchange environment with the shortest altruistic donor chains in addition to the smallest paired exchanges. To define the priority mechanism, each problem is associated with a problem in an auxiliary environment without altruistic donors as follows: Let preference profile $\succsim^C \in \mathcal{P}^C$ be fixed:
   - Every patient $p \in P$, their donor set $\Omega_p$, and preference relation $\succsim_p$ are included to the auxiliary problem.
   - Each altruistic donor $a \in A$ is paired with an auxiliary patient $p_a$ with donor set $\Omega_{p_a} = \{a\}$. The auxiliary patient is endowed with a preference relation $\succsim^C_{p_a}$ such that

     $$\boldsymbol{C}[\ \succsim^C_{p_a}\ ] = \bigcup_{p \in P} \Omega_p.$$

     That is, each auxiliary patient finds all donors acceptable, with the exception of altruistic donors of the original problem.

   Let $P^* = P \cup \{p_a : a \in A\}$ be the patient set and $[P^*, (\Omega_p)_{p\in P^*}, (\succsim^C_p)_{p\in P^*}, n^e = 2]$. We define a priority mechanism induced by a priority order $\pi$ over $P$ as follows: Construct the auxiliary problem and find a matching using the priority mechanism of the auxiliary environment induced by a priority order $\pi^*$ such that

   (a) patients in $P$ are ordered using $\pi$ at the beginning,

   (b) auxiliary patients are ordered at the end in an arbitrary order.

   Moreover, it is possible to minimize the use of altruistic donors subject to generating a priority matching (see Sönmez and Ünver, 2014 how this can be done in a related environment.)

3. The integer programming solution for an environment with arbitrary maximum paired exchange and altruistic donor chain sizes, $\left[P,\ \Omega,\ A,\ \mathcal{P}^C,\ n^e,\ n^a\right]$: Let $\mathbf{E}_{n^e,n^a}$ be the set of all feasible cycles and altruistic donor chains given a preference profile $\succsim^C \in \mathcal{P}^C$. In the formulation of integer program (1), we replace

$\mathbf{E}_{n^e}$ by $\mathbf{E}_{n^e,n^a}$.

At present, a significant fraction of kidney exchange transplants in the US are conducted through altruistic donor chains (Agarwal et al., 2019), as they are easier to organize than paired exchanges. Additionally, occurrences of donors reneging or failing to donate for other reasons are exceedingly rare. According to Cowan et al. (2017), out of 344 chains they reported, only 20 were broken, with only 6 due to reneging and the rest due to other reasons.

Although not as prevalent as altruistic donor chains, the concept of non-simultaneous deceased-donor chains is also feasible, as discussed by Roth et al. (2006). Despite ethical concerns hindering their wide-spread adoption in the US, these chains found viability in countries where altruistic donors were scarce, such as in Italy (Furian et al., 2020).[17]

## 2.8 Worldwide Market Design Initiatives for Kidney Exchange

The launch of the UNOS National Kidney Paired Donation Program in 2010 marked the integration of NEPKE into this program, effectively transitioning NEPKE into the US National Program. Concurrently, the National Kidney Registry (NKR), an independent non-profit based in New York, solidified its position as the largest kidney exchange program in both the US and the world. As of November 2023, NKR holds the top position, with APD trailing as second in the US. The UNOS Program ranks as the third-largest kidney exchange program in the US.

While our discussion has primarily centered on kidney exchange practices within the US, it's important to note that research and implementation of kidney exchange programs have gained significant traction globally. Following the establishment of NEPKE in the US, countries across Europe, Asia, and Australia have seen notable progress in this field. Despite this global momentum, there are still countries, such as Germany and Japan, that have yet to fully embrace kidney exchange programs. Korea, once a leader in the practice of kidney exchange, has shifted its focus towards incompatible transplants.[18] Germany prohibits kidney exchange, permitting trans-

[17]Deceased-donor chains are already incorporated in the initial model of Roth, Sönmez, and Ünver (2004) and TTCC mechanism is tailored for their use. A priority mechanism can also be used here, as we proposed for altruistic donors, by treating $w$ like an altruistic donors and pairing it with $|P|$ different auxiliary patients. The integer programming solution is also straightforward to generalize. Yilmaz (2011) proposes an egalitarian lottery mechanism when unrestricted length deceased-donor chains and paired exchange sizes are feasible. They consider the deceased-donor option $w$ as an inferior option to living donors. Also see Cheng and Yang (2021) for the characterizations of the gains from larger exchanges when chains are possible in large kidney exchange pools, extending the analysis in Roth, Sönmez, and Ünver (2007).

[18]See Chun, Heo, and Hong (2021) for a market design study exploring kidney exchange when blood-type incompatible and tissue-type incompatible transplants are feasible, aiming to minimize them. A sim-

plants only through relatives of patients, while Japan's reluctance is rooted in cultural factors within the transplant community.

On a positive note, several European countries have thriving kidney exchange programs. For example, the UK, which pioneered the use of market design techniques in Europe (Manlove and O'Malley, 2012), as well as programs in Scandinavia (Andersson and Kratz, 2019; Kratz, 2021; Weinreich et al., 2023), and The Netherlands, which was one of the first European countries to implement such programs. While these European programs may operate on a more modest scale compared to the US, they continue to facilitate kidney exchange transplants successfully. Most of these programs utilize various integer-programming-based schemes, primarily focusing on limited-size paired exchanges. Notably, Italy has implemented deceased-donor chains (Furian et al., 2020), a departure from the typical scheme. Further insights into these European developments can be found in the works of Biró et al. (2019) and Biró et al. (2021), which explore these initiatives from a market design perspective.

## 2.9 Frictions that Potentially Increase the Scope of Kidney Exchange

The primary frictions in kidney transplantation have long been recognized as blood-type incompatibility and tissue-type incompatibility. The kidney exchange paradigm has leveraged some of these frictions notably tissue-type incompatibility. Typically, patients with compatible and highly sought-after blood-type donors rarely participate in kidney exchange programs. However, in such cases, the occurrence of less common tissue-type incompatibility within these pairs can be leveraged to assist unfortunate blood-type incompatible pairs through kidney exchange. As a result, patients with highly sought-after blood types but tissue-type incompatible donors have been playing a crucial role in facilitating paired exchanges, as discussed in Section 2.6.

Two innovative ideas, akin to leveraging tissue-type incompatibility, have emerged in the medical literature. One of these concepts has been embraced as a more viable and ethical concept, while the other has sparked more controversy.

### 2.9.1 Leveraging Temporal Incompatibility

The first concept revolves around recognizing *temporal incompatibility* as a pivotal factor in kidney transplantation and as a facilitator in kidney exchanges. Veale et al. (2017) point out that some patients experience declining kidney function but do not immediately need a kidney transplant. Many of these patients often have older, compatible paired donors who are willing to donate a kidney. Due to their advanced ages, these donors have a narrow time window for donation. Veale et al. (2017) report

ilar study by Sönmez, Ünver, and Yilmaz (2018) contrasts the tradeoffs of blood-type incompatible versus tissue-type incompatible transplants in kidney exchange.

successful implementation of donor chains initiated by these older donors, who act like altruistic donors. In return, the patient paired with the older donor is promised a kidney transplant as a terminal tail patient when they need a kidney transplant.

### 2.9.2 Leveraging Financial Incompatibility through Global Kidney Exchange

On the other hand, a more contentious policy intervention arises from recognizing *financial incompatibility* as a barrier to kidney transplantation. The concept of *global kidney exchange* (Rees et al., 2017), championed by economist Alvin Roth and Dr. Michael Rees of APD, aims to address the plight of impoverished patients in low and middle-income countries where kidney transplantation is either prohibitively expensive or logistically challenging. Despite having a compatible paired donor, such patients often cannot undergo the transplant operation locally due to financial constraints. Under global kidney exchange, they can participate in a kidney exchange, with patients in high-income countries–such as the US–who have incompatible donors. This approach yields substantial financial savings by eliminating dialysis costs in participating high-income countries, with these savings then directed toward financing the operation of the compatible patient-donor pair facing financial hardship in low and middle-income countries.

Though heavily promoted and organized by APD, since 2015, over the span of 7 years, only 52 such transplants have taken place (Rees et al., 2022), largely due to significant ethical concerns raised by the medical community regarding this practice. Ambagtsheer et al. (2020) provides a comprehensive analysis of the opposition to global kidney exchange. Opposing organizations include the Council of Europe Committee on Organ Transplantation, the European Union's National Competent Authorities on Organ Donation and Transplantation, and the Declaration of Istanbul Custodian Group against organ trafficking. Most arguments against global kidney exchange contend that it exploits poor countries and individuals, with assistance to impoverished patients coming at the cost of "donated" organs, thereby constituting organ trafficking and increasing the risk of organs being sourced from paid donors. Furthermore, inadequate donor and patient medical follow-up in their home countries may pose significant risks to their well-being. As a result, many argue that global kidney exchange undermines various ethical norms, a stance echoed by Dr. Francis Delmonico (Delmonico and Ascher, 2017), a pivotal figure in the initial collaboration between market designers and medical doctors, who also led the statement of The Declaration of Istanbul Custodian Group (2020) on the issue.

## 2.10 How to Maximize the Benefit from Kidney Exchange?

Before concluding this section, we discuss two potential interventions that could significantly enhance the welfare gains from kidney exchange.

The first proposal involves implementing a token point system to address inefficiencies stemming from misaligned incentives among transplant centers.

The second intervention focuses on incentivizing compatible pairs to participate in exchange programs. This innovation aims to mitigate the efficiency loss caused by donors with highly sought-after blood types donating to their paired patients with less sought-after blood types (e.g., donors with blood type O donating to patients with blood type A), thereby substantially expanding the scope of kidney exchange.

### 2.10.1 Addressing Inefficiencies in Collaborative Kidney Exchange Programs

Large-scale kidney exchange programs usually include many transplant centers as member institutions. Transplant centers submit the database of their eligible patients and donors to the centralized clearinghouse of the collaborative exchange program. Then, the exchange program executes a mechanism to find a match. The resulting transplants are often carried out at the patient's home center. Donors either travel or their kidney grafts are transported to the home center of the receiving patient.

Kidney exchange programs exhibit fragmentation, particularly in the US: there are multiple large-scale nationwide programs such as the NKR, APD, and UNOS National Program. Additionally, many large individual transplant centers are member institutions in one or more of these collaborative exchange programs, and they each also conduct kidney exchanges internally to match their own patients.

A recent study by Agarwal et al. (2019) showed that these centers tend to participate in collaborative exchange programs only with their more challenging-to-match patients, those who cannot be matched internally. This partial participation practice leads to substantial inefficiencies.

Understanding the incentives driving transplant centers' participation decisions is crucial. Typically, each transplant center aims to maximize the number of its patients who receive transplants.[19] Unfortunately, there is no collaborative kidney exchange mechanism that maximizes the number of transplants and makes full participation a dominant strategy for each transplant center as the following example from 2005c demonstrates:

**Example 4** *Suppose there are two centers, A and B, that are members of a collaborative kidney exchange program. Each center aims to maximize the number of its patients who receive*

[19]Both due to altruistic reasons and their revenues being tied to the transplants they conduct.

*transplants. Center $A$ has four patients with paired donors $p_1^A, p_2^A, p_3^A$, and $p_4^A$, while Center $B$ has three patients with paired donors $p_1^B, p_2^B$, and $p_3^B$. The set of individually rational exchanges is as follows (the patients in Center $B$ are shown in lighter gray) (see Figure 13):*

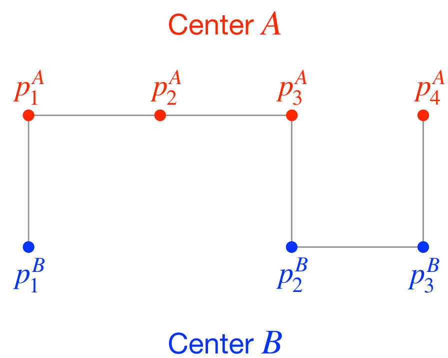



**Figure 13:** Exchanges in Example 4. With only two-way exchanges feasible, each undirected edge of the graph represents an individually rational donor exchange.

$$\mathbf{E} = \left\{(p_1^B, p_1^A), (p_1^A, p_2^A), (p_2^A, p_3^A), (p_3^A, p_2^B), (p_2^B, p_3^B), (p_3^B, p_4^A)\right\}.$$

*Each center can internally match only two of its patients without participating in the centralized exchange program. Center $A$ can conduct internally exchange $(p_1^A, p_2^A)$ and Center $B$ can conduct $(p_2^B, p_3^B)$.*

*Suppose we use a centralized exchange mechanism that maximizes the number of transplants (simply, a maximum mechanism). If each center truthfully reports their patients and paired donors, there are two maximum matchings.*

$$\mu^1 = \left\{(p_1^A, p_2^A), (p_3^A, p_2^B), (p_3^B, p_4^A)\right\}, \qquad \mu^2 = \left\{(p_1^B, p_1^A), (p_2^A, p_3^A), \{p_2^B, p_3^B)\right\}.$$

*Thus, any maximum mechanism chooses either $\mu^1$ or $\mu^2$ for this problem. Of the two, matching $\mu^1$ leaves one patient from Center $B$ unmatched, while matching $\mu^2$ leaves one patient from Center $A$ unmatched.*

*If it chooses $\mu^1$, Center $B$ can manipulate it by reporting to the system only $p_1^B$ and internally matching $p_2^B, p_3^B$ with each other: In this case the centralized mechanism would choose*

$$\mu' = \left\{(p_1^B, p_1^A), (p_2^A, p_3^A)\right\}$$

*as a result all three patients of Center $B$ would be matched (see Figure 14).*

*If it chooses $\mu^2$, Center $A$ can manipulate it by reporting to the system only $p_3^A, p_4^A$ and internally matching $p_1^A, p_2^A$ with each other: In this case the centralized mechanism would*

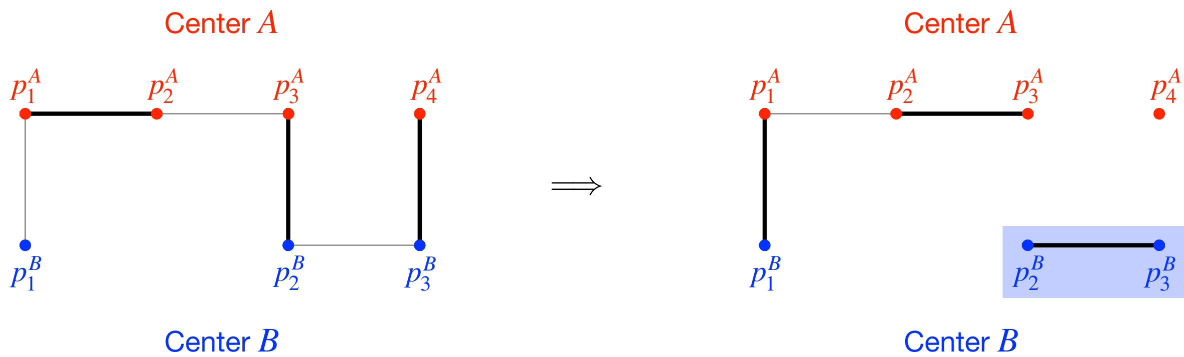



**Figure 14:** In Example 4, when $\mu^1$ in the first panel is chosen then Center $B$ can manipulate by withholding pairs $p_2^B$ and $p_3^B$ as seen in the second panel.

*choose*

$$\mu'' = \left\{ (p_3^A, p_2^B), (p_3^B, p_4^A) \right\}$$

*as a result, all four patients of Center A would be matched (see Figure 15).*

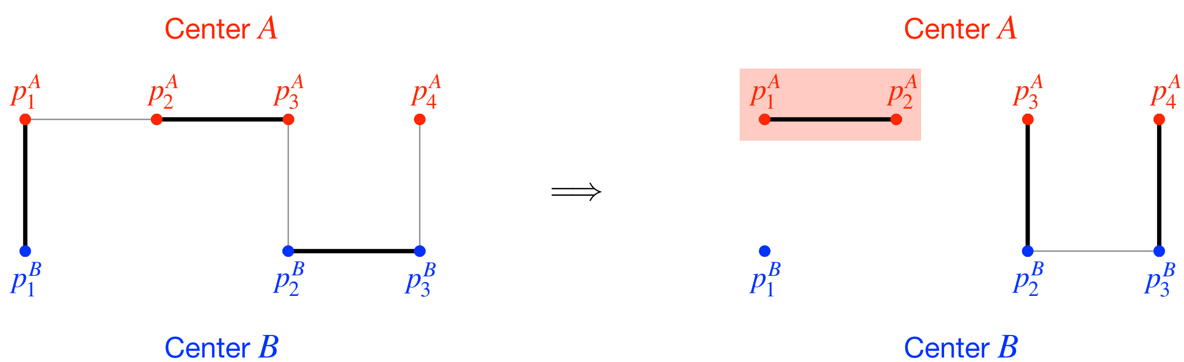



**Figure 15:** In Example 4, when $\mu^2$ in the first panel is chosen then Center $A$ can manipulate by withholding pairs $p_1^A$ and $p_2^A$ as seen in the second panel.

Agarwal et al. (2019) utilizes empirical evidence to demonstrate the widespread nature of the fragmentation problem within kidney exchange programs in the US. They illustrate that centers routinely conduct internal *within-center* exchanges, compromising efficiency in the process. Notably, within-center exchanges account for 62% of kidney exchange transplants in the US. In contrast, the collaborative exchange programs predominantly facilitate exchanges across centers, targeting harder-to-match, highly sensitized patients, differing significantly from the characteristics of within-center exchanges. Additionally, more than 20% of $O$ blood-type donors are matched with non-$O$ blood-type patients in within-center exchanges, surpassing the rate seen in exchanges organized by collaborative programs. This is given as evidence of the inefficiency of within-center exchanges.

To incentivize truthful participation in collaborative exchange programs, they propose a token-money-based system where token rewards for each patient-donor pair transplanted within the collaborative system are tied to the concept of *marginal product* of patient-donor pairs, inspired by the theory of the firm within general equilibrium theory.[20] If a center runs out of its token budget, it cannot submit new pairs to the system. If one of its easy-to-match overdemanded pair without highly sensitized patients is matched in the collaborative system, the center earns tokens, while if one of its difficult-to-match pairs is matched, then the center is charged tokens.[21]

### 2.10.2 Incentivizing Compatible Pairs to Participate in Exchange

Patients with compatible donors rarely participate in kidney exchange, as there may not be a tangible benefit. However, this convention was partly influenced by the evolving partnership between market designers and members of the transplantation community. The initial model proposed by Roth, Sönmez, and Ünver (2004) incentivizes compatible pairs to participate by offering them a higher quality donor if they are matched in an exchange, as discussed in Section 2.3.

However, as highlighted in Section 2.4, this option remained underutilized as the model with compatibility-based preferences (Roth, Sönmez, and Ünver, 2005b, 2007) became the cornerstone of real-life exchange systems. This shift was essentially a prerequisite set by the head of New England Organ Bank, Dr. Francis Delmonico, for collaborating with our team of market design economists to improve their kidney exchange program.

As a by-product of this new approach, no naturally induced "biological" scheme existed for incentivizing compatible patient-donor pairs to participate, leading to al-

[20] As explained in Section 2.6, an overdemanded type pair, such as an $A - O$ pair, can facilitate the matching of at least one underdemanded type pair, such as an $O - A$ pair. Thus, such an overdemanded type pair's marginal product is at least 2 (it can be more larger when three-way exchanges or altruistic donor chains are routinely practiced). On the other hand, an underdemanded pair, like the $O - A$ pair, is not needed to facilitate an exchange on the margin, as there are many other pairs of the same type, and they cannot participate in an exchange without the help of an overdemanded type pair. Thus, an underdemanded type pair has a marginal product of 0. Leveraging a comprehensive dataset from NKR, Agarwal et al. (2019) estimate the empirical marginal products of patient-donor pairs. They find inefficiency in the current system, as the overdemanded type pairs with easy-to-match patients (i.e., with patients who are not highly sensitized) are not always matched in kidney exchange. However, whenever they are matched, their marginal product becomes close to 1.7. On the other hand, pairs with highly sensitized patients or underdemanded types have marginal products close to zero. Their proposed system involves a token reward of each matched pair $i$ with estimated marginal product $MP_i$ and an estimated probability of transplantation $\Pi_i$ is set to $\frac{MP_i}{\Pi_i} - 1$. The subtraction by 1 ensures we exclude the pair itself in the reward calculation.

[21] Besides these two studies, center participation incentives are further studied by Ashlagi et al. (2015) for worst-case efficiency scenario of strategy-proof mechanisms, Ashlagi and Roth (2014) for providing an asymptotically Bayesian incentive-compatible and efficient mechanism, Toulis and Parkes (2015) for analytically characterizing tradeoffs of fragmentation vs centralization in a random graph-based model.

most negligible participation of compatible pairs. However, a few centers, such as the Methodist San Antonio Hospital in Texas (Bingaman et al., 2012), continued to include compatible pairs in their exchanges.

Starting with Roth, Sönmez, and Ünver (2004), many papers in the economics literature, including Roth, Sönmez, and Ünver (2005a), Nicolò and Rodríguez-Álvarez (2012), Sönmez and Ünver (2014), Nicolò and Rodríguez-Álvarez (2017), Sönmez and Ünver (2015), and Sönmez, Ünver, and Yenmez (2020), have considered "biological" or "institutional" incentive mechanisms to compel compatible pairs to participate in kidney exchange. The medical community (Veatch, 2006; Kranenburg et al., 2006; Gentry et al., 2007; Steinberg, 2011; Ferrari et al., 2017) has also found the idea of including compatible pairs to be plausible and important.

More recently, a new medical technology called *eplet matching*, which is a more advanced and accurate version of the older HLA matching technology leveraged in the simulations of Roth, Sönmez, and Ünver (2004), is introduced for use in kidney exchange programs. The largest kidney exchange program in the US, NKR, is incentivizing compatible patient-donor pairs to participate by promising to find better matches using this technology.[22]

Among all the discussed interventions, compatible pair participation is by far the most important innovation to the current kidney exchange paradigm, with the potential to increase kidney exchange transplants by 1.6 times its current amount, or 1800 additional transplants per year (Sönmez, Ünver, and Yenmez, 2020).

Sönmez and Ünver (2015) and Sönmez, Ünver, and Yenmez (2020) propose an "institutional" incentive scheme to encourage the participation of compatible pairs in kidney exchange. Consider a compatible pair that is "overdemanded," such as an $A$ patient with a compatible $O$ paired donor. Under the incentive scheme, if the pair participates in kidney exchange rather than opting for direct transplantation, the patient receives priority on the deceased-donor waiting list in case they require a re-transplant due to graft failure in the future. Therefore, patients from overdemanded pairs are provided with "insurance" against future graft failure in exchange for a more efficient matching for their paired donor. Since patients typically lead healthy lives until their kidney transplant fails, it is likely that they would still be fit to undergo a

[22] Their webpage reads

*"Eplets are small patches of polymorphic amino acids on the surface of HLA antigens. These amino acids are the targets of HLA antibodies.*
*While utilizing traditional HLA antibody matching, scientists noticed that some transplanted kidneys were significantly outperforming their projected survival times. Upon closer investigation, they discovered that these better-performing kidneys had fewer mismatched HLA eplets."*

citing, Wiebe et al. (2017) on benefits of eplet matching for longer survival of a kidney transplant. See https://www.kidneyforlife.org/for-centers/about-eplet-matching retrieved on 12-29-2023.

second transplant. The expected survival period of a successful living donor kidney transplant is only about 16 years.

The authors employ a continuum large market model and, using numerical calibration, calculate that an additional 180 transplants would be conducted for every 10% participation of such compatible pairs in exchange in the US per year. Thus, the participation of even a modest number of compatible pairs in exchange has the potential to substantially increase the number of kidney transplants.[23]

This policy does not inherit the ethical problems associated with "list exchange" or "deceased-donor chains." The patient who would be prioritized on the waitlist is not of blood type *O* and, therefore, does not join the blood type *O* waitlist, which typically has one of the longest waiting times. Indeed, this scheme was favorably received by the transplant community (Gill et al., 2017).

# 3 Living-Donor Liver Exchange

In this section, we discuss another life-saving application of matching theory, market design for liver exchange, and its successful implementation at İnönü University Liver Transplant Institute in Malatya, Turkey. Although the principles of this application are similar to those of kidney exchange, it also exhibits several notable differences.

In Chapter 2, while discussing market design for kidney exchange, we observed that including compatible patient-donor pairs where the donor has a more sought-after blood type than the recipient into the exchange would substantially increase the efficacy of donor exchanges. Additionally, we have explored various sources of friction in living donation. These can be utilized to introduce "institutional" or "biological" incentive schemes aimed at compelling such pairs to participate in kidney exchange. The organization of paired exchanges for living-donor liver transplantation (LDLTs) exploits a few other liver-specific sources of frictions that can be the basis of other forms of "biological" incentive schemes, thus expanding the scope of liver exchange.

## 3.1 Background

LDLT differs from living-donor kidney transplantation in a few important ways. Unlike kidneys, the liver is a single organ. Therefore, living donation involves dissecting and removing only a lobe of the liver from the donor to transplant into the pa-

[23] In addition to the literature and advances cited here, several studies consider dynamic aspects of kidney exchange, mostly pointing out that matching "almost" as many pairs as possible at every instance is nearly optimal using the technologies at hand (e.g., see Ünver, 2010; Kerimov, Ashlagi, and Gurvich, 2023). Others discuss exploiting gains through unpaired exchanges over time (Ausubel and Morrill, 2014; Akbarpour et al., 2020). Dynamic matching is explored in more detail in Section 12 of this handbook.

tient. The liver comprises two well-defined lobes: a larger right lobe, which averages around 60-70% in size, and a smaller left lobe. Additionally, for pediatric patients, two segments of the left lobe, referred to as "Segments 2-3," can be transplanted. (See Figure 16 for an illustration of liver segments and lobes.)

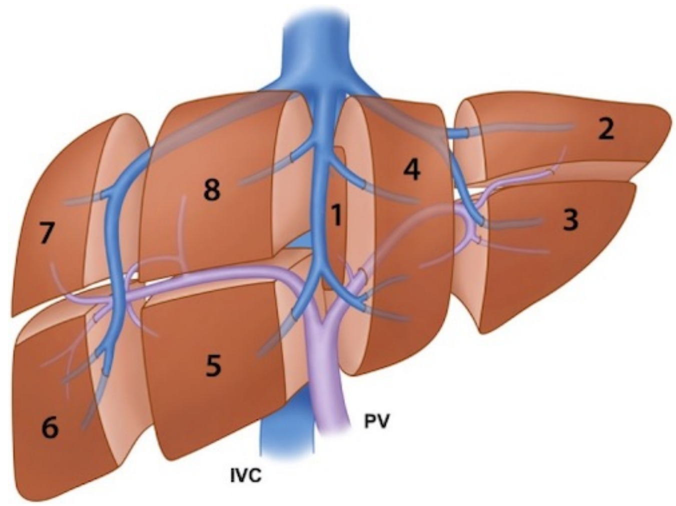



**Figure 16:** Liver segments, from Orcutt et al. (2016). The left lobe (Segments 1, 2, 3, 4) comprises approximately 30-40%, while the right lobe (Segments 5, 6, 7, 8) constitutes around 60-70% of the liver. These percentages exhibit significant variation between individuals. Segments 2 and 3 of the left lobe can also be used as a graft in transplantation for pediatric patients.

The two liver portions, the *remnant* remaining in the donor and the *graft* transplanted to the patient, each grow back so that each person has a completely functioning liver weeks after the transplant operation.

Just like kidney transplantation, *blood-type compatibility* is an essential requirement for liver transplantation. While, as with kidneys, some Asian countries perform blood-type incompatible transplants, this practice is discouraged in much of the western world. Unlike kidney transplantation, however, tissue-type incompatibility is not a concern for liver transplantation, and it is often not checked. Tissue-type incompatible livers are routinely transplanted without any long-term harm.

If blood-type compatibility were the only factor for liver transplantation, then the only possible donor exchanges between incompatible pairs would involve two-way exchanges between:

- Blood-type $A$ patients with blood-type $B$ donors, and
- Blood-type $B$ patients with blood-type $A$ donors.

Thus, the scope of liver exchange would be very limited. However, key to our model and analysis, there are several important considerations regarding the "size" of the liver graft for transplantation.

The patient requires a substantially large graft to survive because they are only receiving a portion. The widely accepted norm is that a patient needs to receive a graft that is at least 0.8% of their body weight (the whole liver is around 2% of the body weight, making it the second largest human solid organ after the skin). However, since the received graft needs to fit in the abdominal liver cavity, it cannot be too large either (with around 2% being the upper limit).

After a potential donation, it's crucial for the donor to retain at least 30% of their liver for a safe operation. Consequently, if the right lobe constitutes more than 70% of the liver, right lobe donation becomes unfeasible for the donor.[24]

Thus, for safety reasons concerning both patients and donors, *size compatibility* assumes a significant role in liver transplantation, unlike in kidney transplantation.

The relevance of graft size extends beyond *size compatibility* in LDLT. A compatible patient-donor pair may seek to participate in liver exchange for various reasons, such as reducing donor risk through a less risky left lobe donation or obtaining a better-sized graft for the patient. For instance, if the donor can only donate their right lobe to their paired patient, the pair may opt for an exchange where they can donate their left lobe to mitigate risks associated with donation.

Moreover, compatible pairs may also aim to participate in liver exchange to improve the long-term success of the transplant by receiving a blood-type identical graft rather than a blood-type compatible one.

These inherent transplantation barriers have significantly influenced the expanded scope of recent developments in liver exchange.

Given the considerably higher donor risk associated with living liver donation compared to kidney donation, it is more commonly pursued in countries where deceased donation is not a primary source of organ procurement. Such countries are predominantly located in the Middle East and Eastern Asia, notably including India, Pakistan, South Korea, Japan, Taiwan, and Turkey. However, in recent years, the number of LDLTs in the US has been increasing due to the acute shortage of transplant livers.

The global-first liver exchange was conducted in South Korea in 2002 (Hwang et al., 2010), although the scope of liver exchange has been limited in this country. The number of liver exchange transplants remained at 52 from 2002 to 2018, representing 0.4% of the LDLTs conducted in South Korea in this period (Kim, 2022). One important challenge has been the complexity in organizing liver exchange, with multi-center collaborations being much less common than in kidney exchange. This traditional setup has imposed significant barriers on the number of transplants that can be conducted

[24] Additionally, other factors such as fatty liver or anatomical variations may render the donor ineligible to donate either lobe.

simultaneously, as a single center may lack the necessary resources to facilitate multi-way liver exchanges.

The simplified liver exchange model we discuss in this section, with two individual/graft sizes, is based on the working paper 2018, later published in 2020 for the more general case with an arbitrary number of individual/graft sizes. Reflecting the logistical capacity of the vast majority of liver transplant programs worldwide, liver exchanges are restricted to two-way exchanges in our formal model.[25]

## 3.2 Liver Exchange Model with Two Individual/Graft Sizes

Let $I$ be a set of liver patients, with each patient *paired* with a single living donor. We sometimes refer to $i \in I$ as a *pair* when it is convenient.[26] For the simple version of the model presented in this section, individuals come in two sizes: large $l$ or small $s$. Let $\mathcal{S} = \{s, l\}$ denote the set of sizes and $\mathcal{B} = \{O, A, B, AB\}$ denote the set of blood types. Thus, $\mathcal{B} \times \mathcal{S}$ represents the set of *individual types*.

For the benchmark case of the model as a reference, we start by considering *left-lobe* transplants only.

A patient and a donor are *left-lobe compatible* if

1. the patient is blood-type compatible with the donor, and
2. the donor is not smaller than the patient.

Formally, the left-lobe compatibility relation is defined as the *liver donation partial order* $\trianglerighteq$ on the set $\mathcal{B} \times \mathcal{S}$ of individual types. The partially ordered set $(\mathcal{B} \times \mathcal{S}, \trianglerighteq)$ forms a lattice depicted in Figure 17.

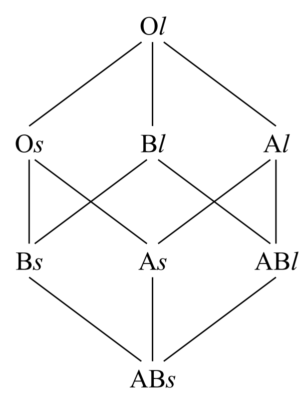



**Figure 17:** Left-lobe compatibility lattice $(\mathcal{B} \times \mathcal{S}, \trianglerighteq)$.

[25] Apart from the Banu Bedestenci Sönmez Liver Paired Exchange System, the two authors have been jointly operating with the Liver Transplant Institute at İnönü University, Malatya, Turkey, later discussed in Section 3.3.2, only two other liver exchange programs have ever performed exchanges larger than two-way, with each conducting one three-way exchange as of March 2024. Thus, while larger exchanges, up to 6-way, have been performed regularly at İnönü University, their experience has not yet been replicated elsewhere.

[26] Since each donor has a single donor, we do not denote a patient's donor separately.

This lattice is order-isomorphic to the standard partial order $\geq$ over the corners of the three-dimensional unit cube, or the binary cube $\{0,1\}^3$ (see Figure 18) such that each individual type $t \in \mathcal{B} \times \mathcal{S}$ is associated with the following vector $X \in \{0,1\}^3$:

$$\begin{aligned} X_1 = 0 &\iff t \text{ has the } A \text{ antigen} \\ X_2 = 0 &\iff t \text{ has the } B \text{ antigen} \\ X_3 = 0 &\iff t \text{ is small} \end{aligned}$$

Here, it is helpful to note that, individuals with blood type $AB$ have both $A$ and $B$ antigens, individuals with blood type $A$ have $A$ antigen only, individuals with blood type $B$ have $B$ antigen only, and individuals with blood type $O$ have neither $A$ nor $B$ antigen. Blood type compatibility means that the donor does not have any antigen the patient lacks. That is the basis of the order isomorphism.

For notational simplicity, we will work with the representation $(\{0,1\}^3, \geq)$.

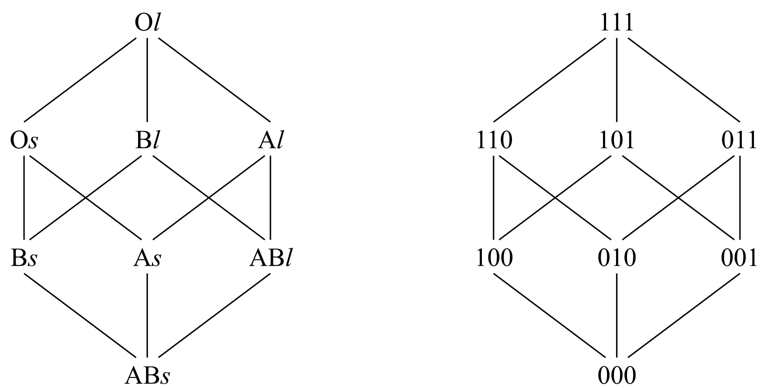



**Figure 18:** Isomorphism between $(\mathcal{B} \times \mathcal{S}, \trianglerighteq)$ and $(\{0,1\}^3, \geq)$.

The *type* of a patient-donor pair is represented through the individual types of its patient and donor, respectively, as $X - Y \in \left(\{0,1\}^3\right)^2$.

A *liver exchange problem with two sizes* is represented as a pair $[I, \tau]$, where each $i \in I$ corresponds to a pair, denoted as $\tau(i) = X - Y$, with a patient of type $X \in \{0,1\}^3$ and a donor of type $Y \in \{0,1\}^3$.[27] Sometimes, we denote type $X - Y$ as $X_1X_2X_3 - Y_1Y_2Y_3$.

A *(left-lobe) direct transplant* consists of a single pair $i$ of type $X - Y$ such that $Y \geq X$. Such pairs are called *(left-lobe) compatible* pairs.

Note that, $Y \geq X$ means that:

1. $Y_1 \geq X_1$. The donor of pair $i$ lacks the $A$ antigen if the patient of pair $i$ lacks the $A$ antigen.

[27] If the set of individual types is given as $\{0,1\}^2 \times \{0,1,\ldots,S-1\}$ for some $S > 2$, then we refer to this problem as a *liver exchange problem* (with $S$ sizes).

2. $Y_2 \geq X_2$. The donor of pair $i$ lacks the $B$ antigen if the patient of pair $i$ lacks the $B$ antigen.
3. $Y_3 \geq X_3$. The donor of pair $i$ is at least as large as the patient of pair $i$.

A *(left-lobe-only two-way) liver exchange* consists of a pair $i$ of type $V - W$ and another pair $j$ of type $X - Y$ such that $Y \geq V$ and $W \geq X$, and it is represented as $\{i, j\}$.

We assume that there is endowment bias, so that no patient with a left-lobe compatible pair participates in a left-lobe-only liver exchange.

A *matching* is a collection of mutually exclusive exchanges and direct transplants. In this arrangement, if a pair is compatible, it participates in a direct transplant. Therefore, *individual rationality* is implicitly embedded in this definition for brevity.

### 3.2.1 Matchings with Left-Lobe-Only Two-Way Exchanges

We begin by examining the structure of left-lobe-only exchanges and identifying the types of matchings that are Pareto efficient. To facilitate our exploration, we introduce a concept that will prove highly valuable.[28]

The *value* of a pair type $\underbrace{X_1X_2X_3}_{=X} - \underbrace{Y_1Y_2Y_3}_{=Y}$ is defined as

$$v(X - Y) = \sum_{k=1}^{3} (Y_k - X_k).$$

We can conceptualize a scenario with left-lobe-only transplants as one where each pair of type $X - Y$ "consumes" three goods, represented by pairs $X_\ell - Y_\ell$ ($\ell = 1, 2, 3$). Here, if $Y_\ell < X_\ell$ for any $\ell \in \{1, 2, 3\}$, direct transplantation is precluded. For instance, if $Y_1 < X_1$, then $Y_1 = 0$ and $X_1 = 1$, indicating that the donor possesses an $A$ antigen absent in the patient. Note that, due to the endowment bias, a pair has nothing to gain from an exchange unless $Y_\ell < X_\ell$ for some $\ell \in \{1, 2, 3\}$. Conversely, a pair has nothing to offer to another pair in an exchange unless $Y_\ell > X_\ell$ for some $\ell \in \{1, 2, 3\}$.

Therefore, we arrive at the following observation:

**Observation 1** *In any liver exchange problem, the only types that could be part of a two-way exchange are*

$$X - Y \in \left(\{0, 1\}^3\right)^2 \ \text{such that} \ X \not\geq Y \text{ and } Y \not\geq X.$$

*Consequently, only types of values* $-1$, $0$, *or* $1$ *can be part of a left-lobe-only two-way exchange with two sizes.*

[28] We should note that this concept is most useful for the two-size model with the binary 3-cube lattice compatibility structure, where the size plays an analogous role to $A$ and $B$ antigens. It's important to mention that for models with more than two sizes, the intuitions we develop here do not apply.

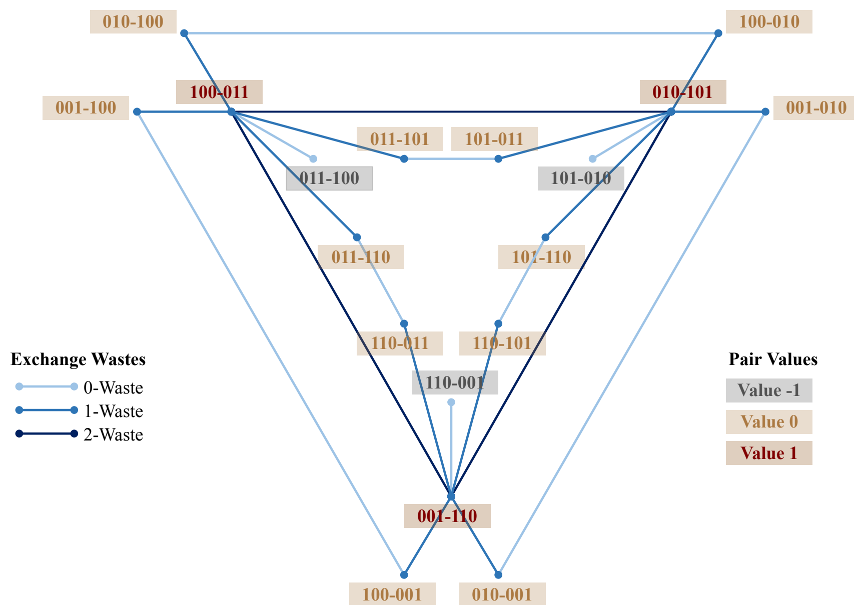



**Figure 19:** Feasible left-lobe-only two-way exchanges and their wastes when there are two sizes. An undirected edge between two pair types designates a feasible two-way exchange between pairs of these types .

The *waste* of an exchange between pair types $V - W$ and $X - Y$ is defined as

$$v(V - W) + v(X - Y).$$

Then, all feasible left-lobe-only exchanges have non-negative waste.

**Observation 2** *In a liver exchange problem with two sizes, any two-way left-lobe-only exchange is either* 0*-waste,* 1*-waste, or* 2*-waste.*

Figure 19 depicts all feasible two-way exchanges and their wastes in the two-size model between pair types.

Using the matroid theory results introduced in Section 2 of Chapter 1 and Proposition 1 in Section 2.4, as the two-way exchange problem spans the matching matroid, a left-lobe-only two-way matching is Pareto-efficient if and only if it maximizes the number of transplants. Moreover, one way to design a Pareto-efficient mechanism is by directly using the matroid greedy algorithm.[29] However, we will take a more direct approach, exploiting the lattice-compatibility structure and the concept of waste. Note that, in a two-way left-lobe-only exchange, if it results in 1-waste or 2-waste, then at least one of the patients is "using up" a more sought-after liver graft than

[29] Recall that the matroid greedy algorithm requires the independent system of a matroid to be known or to check whether a constructed set is independent or not using an external algorithm. Edmonds (1965) algorithm can be utilized for this latter purpose for this problem.

they need. For instance, this could occur when a small patient receives a graft from a large donor, or when a blood type $A$ patient (having antigen $A$) receives a graft from a blood type $O$ donor (lacking antigen $A$). Therefore, it is intuitively clear that lower waste exchanges are associated with higher efficiency.

The following algorithm, which corresponds to the sequential minimization of waste in choosing exchanges, is due to Ergin, Sönmez, and Ünver (2018):

**Two-Size Left-Lobe-Only Sequential Two-way Exchange Algorithm.**

Fix a priority order over patients.

**Step 0.** Clear all feasible direct transplants.

**Step 1.** Clear 0-waste exchanges following the given priority order.

**Step 2.** Clear 1-waste exchanges following the given priority order among the remaining patients.

**Step 3.** Clear 2-waste exchanges following the priority order among the remaining patients.[30]

■

We depict the functioning of this algorithm in Figure 20. This graph representation will help us also decipher the main mechanism we introduce when right-lobe transplants are also feasible.

We have the following result.

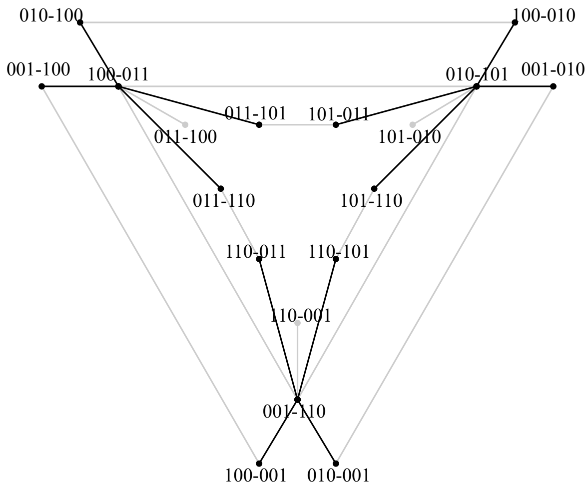
**Theorem 10 (Ergin, Sönmez, and Ünver, 2018)** *Given a liver exchange problem with two sizes, the left-lobe-only sequential two-way exchange algorithm maximizes the number of left-lobe-only two-way exchanges.*

### 3.2.2 Incentives for Right-Lobe Donation

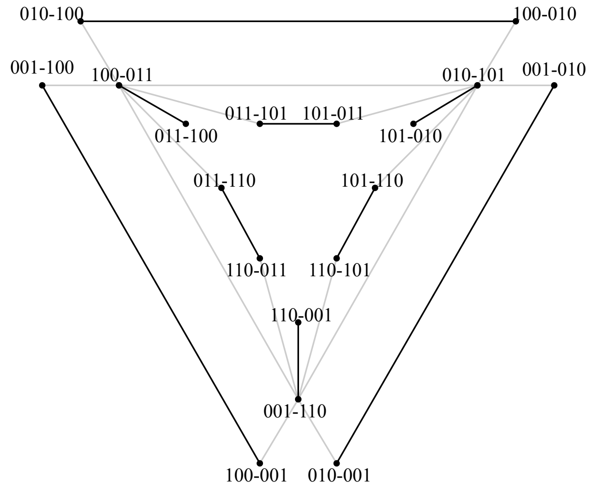
We next consider the possibility of right-lobe transplantation in addition to the left-lobe transplantation. Thus, we have the following two LDLT technologies:

- Left-lobe donation: It is less risky for the donor. It requires left-lobe compatibility, i.e., a blood-type compatible donor should be at least as large as the patient.

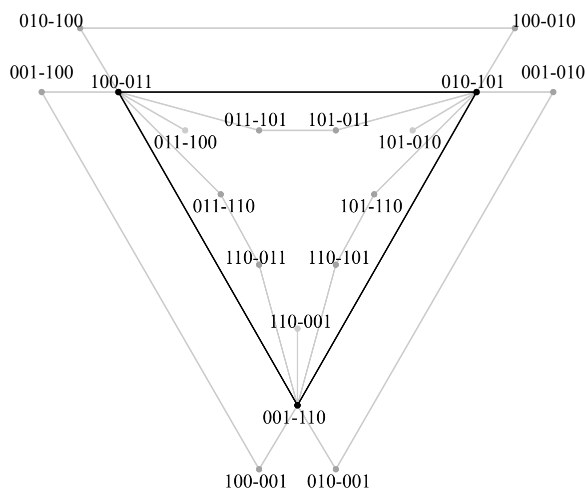
[30] This part is not as straightforward because of the triangular feasible exchange structure governing pairs with 2-waste exchanges; but the structure is simple enough that we can immediately tell how this process will go forward: Suppose at the beginning of Step 3 there are $n_1$ pairs of type $\tau_1 = 100 - 011$, $n_2$ pairs of $\tau_2 = 010 - 101$, and $n_3$ pairs of $\tau_3 = 001 - 110$. W.l.o.g. suppose $n_1 \geq n_2 \geq n_3$ so that the other orderings are symmetrically handled. If $n_1 \geq n_2 + n_3$, then all $\tau_2$ and $\tau_3$ type pairs are matched with $\tau_1$ type pairs who are chosen by priority. If $n_1 < n_2 + n_3$, then all remaining pairs of these three types are matched if their total number is even, and only the lowest-priority one among them remains unmatched if their total number is odd. After determining which pairs will be matched, among them, we can match $\tau_3$ type pairs with any of the other two types so that an equal number of pairs of the other two types $\tau_1$ and $\tau_2$ are remaining to be matched. Then, we match those two groups with each other.




Step 1. Clear 0-waste exch.




Step 2. Clear 1-waste exch.




Step 3. Clear 2-waste exch.

**Figure 20:** Steps 1-3 of the left-lobe-only sequential two-way exchange algorithm.

- Right-lobe donation: It is more risky for the donor. It allows a blood-type compatible donor to donate to a larger patient.

For right-lobe donation feasibility, in addition to left-lobe compatibility, we define right-lobe-only compatibility: a donor of type $Y \in \{0,1\}^3$ is *right-lobe-only compatible* with a patient of type $X \in \{0,1\}^3$ if $X \not\leq Y$ but $X \leq Y_1 Y_2 1$, meaning that $Y$ is not left-lobe-compatible with $X$ but right-lobe compatible.

Observe that a donor of type $Y$ is right-lobe-only compatible with a patient of type $X$ if, and only if (i) $X_1 \leq Y_1$, (ii) $X_2 \leq Y_2$, and (iii) $X_3 = 1$ and $Y_3 = 0$.

These technologies and the associated donor risk profile motivate the following layered preferences for a pair:

- Donating the donor's left lobe is always preferable to donating the donor's right lobe or not donating at all.
- However, the preference between donating the right lobe versus not donating at all is not clear: The pair may prefer donating the right lobe to not donating at all, or they may prefer not donating at all to donating the right lobe.

Thus, depending on their willingness for the right-lobe donation of the pair, there are two possible preferences for pairs indicated with their types: The *type willing* ($w$), who prefer right lobe donation to not donating at all, and the *type unwilling* ($u$), who prefer not donating to right lobe donation.

An *outcome* for a donor-patient pair is denoted as $(x, y)$, where $x$ refers to the type of donation the donor endures, and $y$ denotes the type of transplant it participates in. The set of possible outcomes is given as:

$$\Big\{(\emptyset,\emptyset), \text{(Donate Left, Direct)}, \text{(Donate Left, Exchange)}, \\ \text{(Donate Right, Direct)}, \text{(Donate Right, Exchange)}\Big\}$$

Depending on their willingness type, the two possible preference relations over outcomes for each patient-donor pair $i$, starting with the best outcome at the top, are given as:

| **Willing preferences $R_i^w$** | **Unwilling preferences $R_i^u$** |
|---|---|
| (Donate Left, Direct) | (Donate Left, Direct) |
| (Donate Left, Exchange) | (Donate Left, Exchange) |
| (Donate Right, Direct) | $(\emptyset,\emptyset)$ |
| (Donate Right, Exchange) | $\vdots$ |
| $(\emptyset,\emptyset)$ | |
| $\vdots$ | |

As before, for each $i \in I$, we focus only on *individually rational exchanges* given a willingness profile $R = (R_i)_{i \in I}$ with $R_i \in \{R_i^w, R_i^u\}$ for each $i \in I$:

- A left-lobe compatible pair does not participate in an exchange but participates only in a direct transplant.
- A right-lobe-only compatible pair participates in an exchange only if its donor donates their left lobe; otherwise, it participates in a direct right-lobe transplant.

We assume that the willingness type of a pair is private information. As before, we study direct revelation mechanisms. In this case, they are used to elicit willingness types.

We denote the type of a pair of type $X_1X_2X_3 - Y_1Y_20$ together with their willingness type as $X_1X_2X_3 - Y_1Y_20t$, where $t \in \{u, w\}$ denotes whether they have unwilling $R_i^u$ or willing $R_i^w$ preferences, respectively. Thus, when a pair of type $X_1X_2X_3 - Y_1Y_20w$ is chosen to donate the right lobe, it is treated as if it is of type $X_1X_2X_3 - Y_1Y_21$. We refer to this treatment as a *transformation*.

When the use of a type $X_1X_2X_3 - Y_1Y_21$ covers both *native* $X_1X_2X_3 - Y_1Y_21$ type pairs and *transformed* $X_1X_2X_3 - Y_1Y_20w$ type pairs, we refer to it as an *auxiliary* type.

We have the following intermediate result:

**Lemma 1** *In a liver exchange problem with two sizes and a given willingness profile $R$, if a type $X - Yw = X_1X_2X_3 - Y_1Y_20w$ pair can participate in a two-way exchange with some auxiliary type $V - W$ pair then either*

- *$X \leq W$ and $V \leq Y$, in which case type $X - Yw$ pair donates a left lobe, or*
- *$X \leq W$ and $V \leq Y_1Y_21$ but not $V \not\leq Y$, in which case type $X - Yw$ pair donates a right lobe.*

Depending on their types, the following result characterizes all possible transplant options for pairs:

**Lemma 2 (Individually Rational Two-way Matchings)** *Suppose both left-lobe and right-lobe transplantation are feasible. Then, in a liver exchange problem with two sizes, given a willingness profile $R$, a pair type $X - Y$ belongs to one of the following seven disjoint categories:*

*Cat. 0.* $X > Y_1Y_21$*: A pair of this type cannot participate in an exchange or a direct transplant.*

*Cat. I.* $X \leq Y$*: A pair of this type participates in a direct left-lobe transplant.*

*Cat. II.* $Y_3 = 0$ *&* $X = Y_1Y_21$*: A pair of this type participates in a direct right-lobe transplant only if they are willing.*

*Cat. III.* $Y_3 = 1$ *&* $X \not\geq Y$ *&* $X \not\leq Y$*: A pair of this type can only participate in exchange, and only by donating a left lobe.*

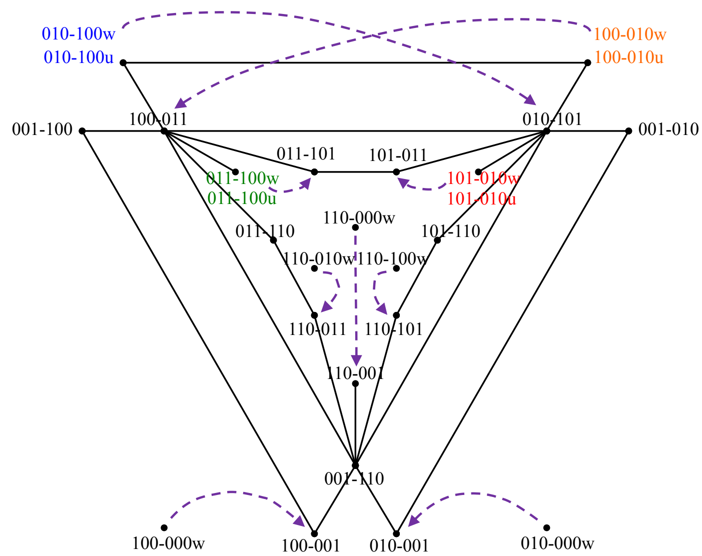



**Figure 21:** Possible transformations: Only 4 types in Category V ($010-100w$, $100-010w$, $011-100w$, $101-010w$) can both donate left lobe or right lobe in individually rational two-way exchanges. Only 5 types in Category IV ($010-000w$, $100-000w$, $110-000w$, $110-010w$, $110-100w$) can donate right lobe but not left lobe in an individually rational two-way exchange.

*Cat. IV.* $X_3 = 0, Y_3 = 0$ *&* $X > Y$*: A pair of this type can only participate in exchange, and only by donating a right lobe when they are willing (see Figure 21).*

*Cat. V.* $Y_3 = 0$ *&* $X \not\geq Y$ *&* $X \not\leq Y_1Y_21$ ($010-100$, $100-010$, $011-100$, $101-010$)*: A pair of this type can only participate in exchange, either by donating a left lobe or a right lobe when they are willing (see Figure 21).*

*Cat. VI.* $X < Y_1Y_21$ *&* $X \not\geq Y$ *&* $X \not\leq Y$*: A pair of this type can participate in exchange by donating a left lobe, or receive a direct right-lobe transplant when they are willing.*

### 3.2.3 An Incentive-Compatible and Pareto-Efficient Mechanism

We explore mechanisms that are dominant-strategy incentive compatible in the revelation of willingness types.

Formally, a *mechanism* is a function that maps each willingness type profile to a matching. Given a mechanism $\varphi$ and willingness type profile $R$, let $\varphi[R]$ denote the matching chosen by $\varphi$ under $R$, and $\varphi[R](i)$ denote the resulting outcome for each pair $i \in I$.

A mechanism $\varphi$ is *incentive compatible* if it constitutes a (weakly) dominant strategy for each pair to truthfully reveal its willingness type. That is, for any pair $i \in I$ with willingness type $t \in \{w, u\}$, any $s \in \{w, u\} \setminus \{t\}$, and any willingness type profile for other pairs $R_{-i}$,

$$\varphi[R_i^t, R_{-i}](i)\ R_i^t\ \varphi[R_i^s, R_{-i}](i).$$

When right-lobe transplantation is feasible, Pareto efficiency no longer implies transplant maximality. Moreover, in general, there exists no transplant maximal mechanism that is incentive-compatible.

**Proposition 5** *There is no incentive-compatible mechanism that maximizes*

1. *the number of transplants, or even*
2. *the number of left-lobe transplants.*

The proof of this result follows from the following example:

**Example 5** *Consider a set $I = \{i_1, i_2, i_3, i_4\}$ of patients and their types $\tau$ given as:*

$$\tau(i_1) = 101 - 011, \qquad \tau(i_2) = 100 - 011,$$
$$\tau(i_3) = 011 - 100, \qquad \tau(i_4) = 011 - 100.$$

*Suppose $i_3$ and $i_4$ are both willing.*
*Any left-lobe-donation- or total-transplant-maximizing matching (two of which can be obtained by swapping $i_3$ and $i_4$ with each other) generates two exchanges. Consider these two matchings:*

$$\mu = \Big\{\{i_1, i_3\}, \{i_2, i_4\}\Big\} \qquad \mu' = \Big\{\{i_1, i_4\}, \{i_2, i_3\}\Big\}$$

*While the donors of pairs $i_3$ and $i_4$ can only donate their right lobes to the patient of pair $i_1$, they can also more preferably donate their left lobes to the patient of pair $i_2$. Any (probabilistic) mechanism that chooses a matching with the maximum number of transplants or the maximum number of left-lobe transplants chooses at least one of these two matchings in its support. W.l.o.g., suppose $\mu$ is that matching. Then $i_3$ has an incentive to announce its type as unwilling by revealing $R'_{i_3} = R^u_{i_3}$, as the mechanism will then choose $\mu'$, which is the unique left-lobe-donation- and total-transplant-maximizing matching in this case, with probability 1. Hence, there is no incentive-compatible mechanism that maximizes the total number of transplants or left-lobe transplants.*

Therefore, we propose a mechanism that is Pareto-efficient and incentive-compatible, albeit not necessarily maximal. This mechanism is built upon a sequential algorithm akin to the one utilized for left-lobe-only exchanges. We attain incentive compatibility by gradually transforming willing pairs only after their left-lobe transplant prospects are fully exhausted. As a result, a willing pair has no incentive to falsify its type as unwilling to secure a left-lobe donation over a right-lobe one. Moreover, employing a fixed priority order in each step to manage exchanges involving different pair categories ensures that attempting to manipulate it in reverse—where an unwilling pair falsely claims willingness—yields no benefit either.

For each category of pairs specified in Lemma 2, we initially observe whether incentive compatibility is pertinent and its implications if so. These observations will then inform the formulation of a Pareto-efficient and incentive-compatible mechanism.

*Cat. 0.* They cannot participate in any direct transplant or exchange; they remain without a transplant. There are no incentive issues related to them.

*Cat. I.* They can only participate in direct left-lobe transplant; we assign them a direct left-lobe transplant at the beginning. There are no incentive issues related to them.

*Cat. II.* They can only participate in a direct right-lobe transplant; if they are willing, we transform them at the beginning and assign them a direct right-lobe transplant. There are no incentive issues related to them.

*Cat. III.* They can participate only in an exchange and only via left-lobe donation. There are no incentive issues related to them.

*Cat. IV.* They can participate only in an exchange and only via right-lobe donation; if they are willing, we transform them to donate a right lobe at the beginning. There are no incentive issues related to them.

*Cat. V.* They can participate only in an exchange, and either via left-lobe or right-lobe donation. We need to take their incentives into account and **gradually transform** them to donate a right lobe if they are willing after their left-lobe donation prospects are fully exhausted.

*Cat. VI.* They can participate in an exchange only via left-lobe donation or in a direct right-lobe transplant. We need to take their incentives into account. If they are willing and still unmatched until the end of the algorithm, we transform them at the end to participate in a direct right-lobe transplant.

The following mechanism is from Ergin, Sönmez, and Ünver (2018). It builds on the same insight from left-lobe-only exchanges to clear 0-waste, 1-waste, and then 2-waste exchanges in this order for efficiency, while integrating with our above-outlined strategy for sustaining incentive compatibility.

**Left-or-Right-Lobe Two-Way Exchange Mechanism.**

Fix a priority order over pairs.

**Step 0.** Direct transplant each Category I and Category II type $w$ pair.

**Step 1.** Transform Category IV type $w$ pairs.

Clear 0-waste exchanges following the given priority order.

At least one of the Category V types $010-100$ and $100-010$ is fully depleted.

Assume without loss of generality, type $100-010$ pairs are depleted (when type $010-100$ pairs are depleted, the algorithm is symmetrically defined):

**Step 2.** **a.** Clear all remaining exchanges of Category V type $010 - 100$ pairs (which are all 1-waste).

**b.** Transform any remaining Category V type $010 - 100w$ pairs.
No exchange remains for pairs of the Category V type $011 - 100$.
Clear the newly formed exchanges of Category V type $101 - 010$ pairs (which are 0-waste).

**c.** Transform the remaining Category V $w$ pairs, which are of types $011 - 100w$ and type $101 - 010w$.
Clear the newly formed 0-waste exchanges (which are only between auxiliary types $011 - 101$ and $101 - 011$).

**d.** Clear 1-waste exchanges following the given priority order.

**Step 3.** Optimally clear 2-waste exchanges.

**Step 4.** Direct transplant each remaining Category VI type $w$ pair.

■

We depict the steps 1-3 of the algorithm in Figure 22.

We conclude the formal analysis in this section with the following result:

**Theorem 11 (Ergin, Sönmez, and Ünver, 2018)** *In a liver exchange environment with two sizes when only two-way exchanges are allowed, the left-or-right-lobe sequential two-way exchange mechanism is individually rational, Pareto-efficient, and incentive compatible.*

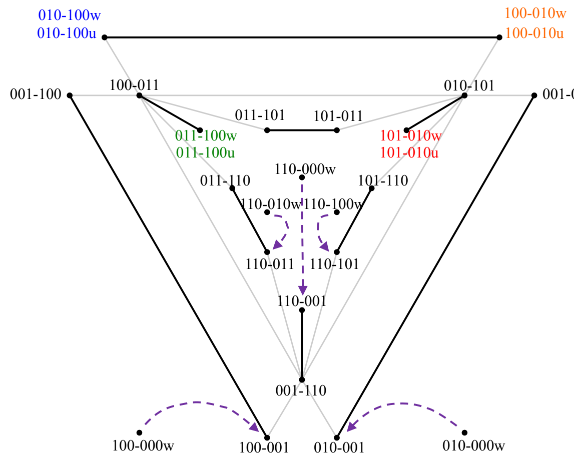
## 3.3 Liver Exchange Programs

While liver exchange has been practiced since 2002, the utilization of matching theory and market design in real-life implementation is relatively new. In this section, we discuss liver exchange programs in South Korea, India, and the US, which evolved without involvement from design economists, as well as a recent liver exchange program designed and managed by economists.

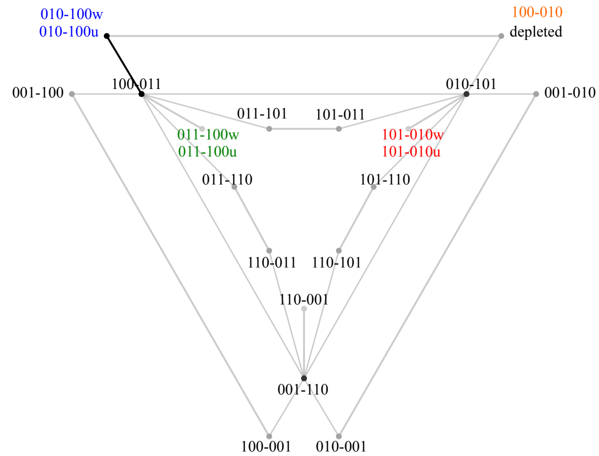
### 3.3.1 Liver Exchange Programs in South Korea, India and the US

The ASAN Center in South Korea conducted the first liver exchange in the early 2000s (Hwang et al., 2010). With the exception of two three-way exchanges, all reported liver exchanges globally had been two-way prior to 2022. The first three-way exchange took place in Pakistan (Salman, Arsalan, and Dar, 2023), and a second was conducted in Gurugram, India, at the Medanta Institute (Soin et al., 2023). The latter exchange program also reported 88 liver exchange transplants in 44 two-way exchanges conducted between 2013 and 2022 within the same paper.

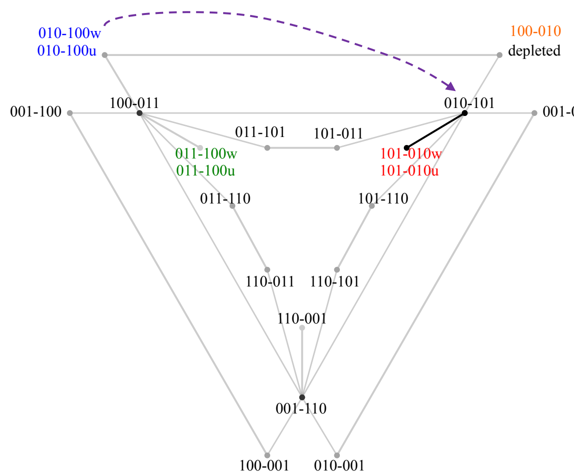
Agrawal, Gupta, and Saigal (2023) summarizes the efficacy of long-standing programs apart from Medanta. They discuss three distinct programs: Max Saket Hospital in New Delhi, India (Agrawal et al., 2022), which conducted 34 liver exchange



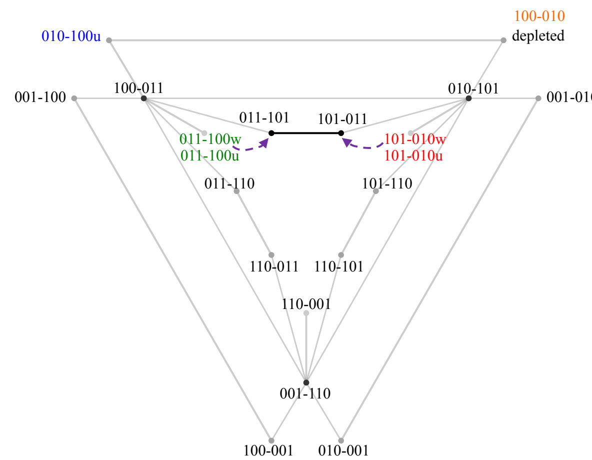

Step 1: Transform all Cat. IV $w$ and clear 0-waste exchanges



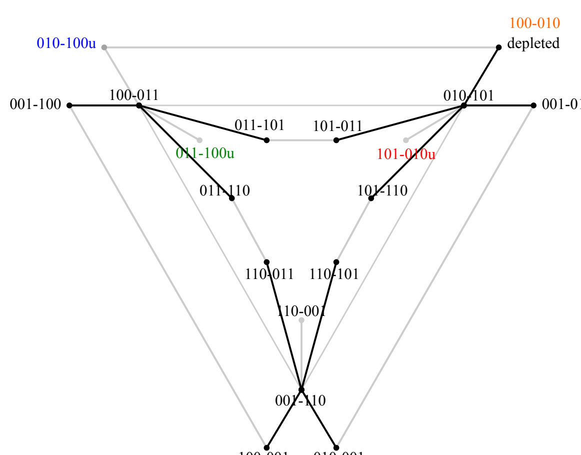

Step 2a: Clear remaining exchanges of Cat. V $010 - 100$ (all 1-waste)



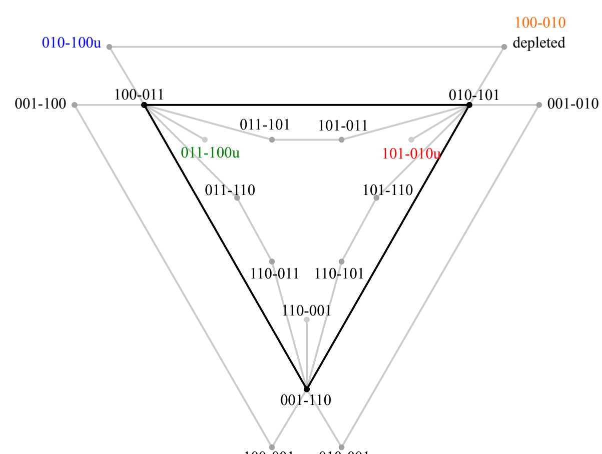

Step 2b: Transform remaining Cat. V $010 - 100w$ and clear new exchanges of Cat. V $101 - 010$ (all 0-waste)




Step 2c: Transform remaining Cat. V $w$ and
clear new 0-waste exchanges




Step 2d: Clear 1-waste exchanges




Step 3: Clear 2-waste exchanges

**Figure 22:** Steps 1-3 of the left or right-lobe two-way exchange mechanism assuming after Step 1 all Cat. V type $100 - 010$ pairs are depleted.

transplants (constituting 1.45% of all LDLTs in the center) between 2012 and 2021; the ASAN Center's program in Seoul, South Korea (Hwang et al., 2010; Jung et al., 2014), conducted 26 liver exchange transplants (1.2% of all LDLTs in the center) between 2003 and 2011; and the University of Pittsburgh Medical Center (UPMC) in the US (Gunabushanam et al., 2022), which conducted 20 liver exchange transplants between 2019 and 2021 (8.3% of all LDLTs in the center). Importantly, seven of the liver exchanges were through non-directed donor chains for the UPMC program, thus increasing its efficacy.

In the US, a pilot national liver paired exchange program was initiated through UNOS in 2021, encompassing 17 transplant centers. Unfortunately, the program ceased operation in 2023 without conducting a single exchange. Their efforts were exclusively limited to finding and conducting two-way exchanges.

Thus, possibly due to the lack of involvement from design economists until recently, the overall global landscape of liver exchange practices has been bleak compared to kidney exchange. However, a recently established program, with the assistance of the authors of this chapter, challenges this trend.

### 3.3.2 Banu Bedestenci Sönmez Liver Paired Exchange System at Malatya İnönü University, Turkey

The authors of this chapter, in collaboration with the Liver Transplant Institute team at Malatya İnönü University in Turkey, under the leadership of Dr. Sezai Yilmaz, a professor of transplantation surgery, established a liver exchange program in 2022. This institute ranks globally as the second highest in performing LDLTs annually, following the ASAN Transplant Center in South Korea.

The principles of this system were detailed in Yilmaz et al. (2023b). A significant departure from the assumptions outlined in Ergin, Sönmez, and Ünver (2018, 2020) is the routine conduct of multi-way exchanges, enabled by the center's capacity to facilitate up to six transplants simultaneously. As a prelude to the exchange program, in 2019 they executed five simultaneous (non-exchange) direct transplants, a world-first rehearsal (Yilmaz et al., 2023a).

As of March 2024, the program has performed 2 six-way exchanges, 2 five-way exchanges, 4 four-way exchanges, 10 three-way exchanges, and 14 two-way exchanges by March 2024, leading to a total of 96 transplants. In 2023 alone, this program facilitated 64 LDLTs, constituting 27.7% of the LDLTs performed in the Institute.[31] The only other reported exchanges larger than two-way in the world are the previously mentioned two three-way exchanges.

The design of the liver exchange system at the Institute relies on several center-

[31] Refer to https://canlikaracigernakli.inonu.edu.tr/en for system details, retrieved on 2/19/2024.

specific factors, such as their unique capacity to carry out large exchanges. Historically, donor outcomes from right-lobe and left-lobe donations have shown similarity at the Institute, with over 3600 LDLTs performed until March 2024. Consequently, this team prefers right-lobe transplants for adult patients, as left-lobe grafts often present more challenging anatomical variations during surgery. Their patient demographic also includes a substantial number of pediatric patients for whom they utilize left-lobe or segment 2-3 transplants. As such, the center predominantly carries out right-lobe transplants in exchanges involving adult patients and left-lobe or segments 2-3 transplants for pediatric patients. Nevertheless, the fundamental principle of incorporating compatible patient-donor pairs into the system is upheld, provided they are willing, and the patient receives at least an equivalent graft to what the donor could donate, often receiving a superior graft. Superior ABO-identical grafts are often provided instead of ABO-compatible but non-identical grafts from their donors, as non-identical ABO grafts may more often induce antibody rejection in the long term. When there is no benefit from ABO match, better-sized grafts are transplanted to the patients than their compatible donors can provide.

In conclusion, exploiting size incompatibility plays a pivotal role alongside the involvement of compatible pairs in the success of this liver exchange program. These factors, combined with the capability to execute larger exchanges, have led to an unprecedented efficacy of liver exchange at the Institute.

# 4 Cadet-Branch Matching in the US Army

In the last two decades, the branch assignment processes for cadets at the United States Military Academy (USMA) and Reserve Officer Training Corps (ROTC) have undergone a series of reforms. This section examines the shortcomings of the USMA-2006 and USMA-2020 mechanisms, each adopted by the Army to accommodate a new objective, and discusses the partnership between the Army and market designers that led to the adoption of a new mechanism, DPCO, for the Class of 2021. The primary focus will be on the theoretical and operational aspects of these mechanisms, highlighting the challenges encountered and how DPCO uniquely addressed these challenges, leading to its adoption. Building on Sönmez and Switzer (2013) and Sönmez (2013), this section largely follows Greenberg, Pathak, and Sönmez (2023).

## 4.1 Background

Each year, the US Army employs centralized matching systems to allocate thousands of graduating cadets from the USMA at West Point and the Reserve Officer Training Corps (ROTC) to their initial military occupation, also referred to as a *branch*.

These branch assignments hold significant consequences for the cadets' career progression. Prior to the Class of 2006, cadets were assigned positions within Army branches through a *priority mechanism* based on a cadet performance ranking known as the *order of merit list* (*OML*). Under this mechanism, cadets would submit their preferences for branches, and the cadet with the highest *order of merit score* (*OMS*) would be assigned their most preferred branch, followed by the second-highest OMS cadet receiving their preferred branch among those with remaining positions, and so forth. This mechanism underscored the importance of a meritocratic hierarchy within the Army.

### 4.1.1 BRADSO Program and the 2006 Branching Reform

In response to declining junior officer retention rates during the late 1990s and early 2000s, the Army initiated a series of retention incentives for cadets at USMA and ROTC starting in 2006.

One of the most popular incentives, which involved overhauling the branching mechanism, was the *branch of choice (BRADSO)* program.

Under this program, cadets receive higher priority for a portion of positions within any given branch if they express willingness to extend their Active Duty Service Obligation (ADSO) by three years with that branch. Consequently, the extended ADSO entails eight years of obligatory service, compared to the standard five-year ADSO term. We refer to the ADSO extension as the *price* of the branch.

The *message space* of the new mechanism was also expanded by requesting cadets to report the set of branches for which they are willing to pay the increased price in exchange for receiving increased priority at a fraction of its positions.

**The USMA-2006 Mechanism.** Under the *USMA-2006 mechanism*, branch assignments follow a process similar to the previous OML-induced priority mechanism, with one significant difference:

Once the base-price positions are filled within any branch, priority is given to cadets who have indicated willingness to pay the increased price for the remaining flexible-price positions.

The prices are subsequently determined as follows:

- Cadets who receive a base-price position are charged the base price.
- Cadets who receive a flexible-price position are charged:
  - the base price if they have not indicated willingness to pay the increased price for their assigned branch, and
  - the increased price if they have indicated willingness to pay the increased price for their assigned branch.

**Shortcomings of the USMA-2006 Mechanism.** Two aspects of the USMA-2006 mechanism were problematic.

First, cadets were asked if they were willing to pay more for a branch without considering other options. For example, a cadet couldn't say they'd pay more for their top choice branch if their only other option was a lower choice branch, but not if it was their second choice.

Second, cadets who express willingness to pay the increased price for a branch are charged the increased price upon receiving one of its flexible-price positions, even if they would have received the same position at a base price without expressing such willingness.

These aspects, in turn, lead to several shortcomings of the USMA-2006 mechanism, including the following two:

- *Detectable priority reversal*: A cadet may be assigned a position at the increased price, while a lower-ranked cadet, based on the OML, receives a position at the same branch at the base price.
- *Failure of incentive compatibility*: A cadet may benefit from concealing their willingness to pay the increased price (failure of BRADSO incentive compatibility or BRADSO-IC) or from misrepresenting their branch preferences.

The root causes of failures under the USMA-2006 Mechanism were twofold: (1) the message space lacks sufficient richness to capture cadet preferences over branch-price pairs, and (2) the assignment's two elements—the branch assignment and the price assignment—are determined sequentially rather than jointly.

**The Initial Proposal of the Cumulative Offer Mechanism.** Both root causes of the failures can be addressed through foundational research on *matching with contracts* (Hatfield and Milgrom, 2005), and in particular, relying on the *cumulative offer mechanism*, covered in Chapter 9 of this handbook.

Hence, to address the root causes of the failures, Sönmez and Switzer (2013) proposed a refinement of the cumulative offer mechanism, implemented with branch choice rules that capture the Army's policy objectives. The proposed mechanism operates as a direct mechanism, wherein cadets express their preferences over branch-price pairs. Initially, the Army perceived this message space as overly complex and deemed the adoption of a new mechanism with a more intricate message space unnecessary for three primary reasons:

1. BRADSO-IC failures and detectable priority reversals have been infrequent in practice.
2. Any BRADSO-IC failure or detectable priority reversal can be manually rec-

tified ex-post, as each instance involves a cadet unnecessarily paying the increased price at their assigned branch.

3. Although there may be additional priority reversals that cannot be manually corrected ex-post, their verification relies on cadet preferences over branch-price pairs, information that is unavailable under the existing USMA-2006 message space.

In summary, any failure of the USMA-2006 mechanism can be rectified manually afterward or may not be verifiable with the available data.

#### 4.1.2 USMA-2020 Mechanism

In 2012, the Army introduced a Talent-Based Branching (TBB) program to create a "talent market" where additional information about each cadet influences their priority at a branch. Under this program, branches assess cadets into three tiers: High, Medium, and Low. These ratings were initially part of a pilot initiative for several years.

For the Class of 2020, the Army incorporated these ratings into the branching process. Priorities at each branch were determined first by the tier and then by the OML within the tier. Additionally, changes were made to the BRADSO policy: cadets willing to pay the increased price now received higher priority within their tier only.

Given the decision to incorporate cadet ratings into the branching process within a tight timeline, the Army opted to maintain the existing message space for the new mechanism, as utilized in previous years. Utilizing an adjusted priority order that took into account both TBB ratings and willingness to pay increased prices, the new mechanism employed the *individual-proposing deferred acceptance* (*DA*) *algorithm* (Gale and Shapley, 1962) to determine branch assignments.

Prices were then determined as follows: Constrained by a maximum number of flexible-price positions at each branch and following the reverse-priority order of branches, cadets indicating willingness to pay the increased price at their assigned branch were charged accordingly, while remaining matched cadets were charged the base price.

#### 4.1.3 Army's Partnership with Market Designers

Alongside inheriting the limitations of the USMA-2006 mechanism, the USMA-2020 mechanism introduced fresh challenges. Although the number of flexible-price positions at each branch remained at 25% of the total capacity, a priority upgrade due to increased-price willingness was applied to all positions.

This design choice introduced a new type of incentive compatibility failure known as *strategic BRADSO*. Previously, indicating a willingness to pay the increased price could potentially harm cadets due to BRADSO-IC failures, but now it could also benefit them with a costless priority upgrade due to strategic BRADSO. Consequently, the mechanism became highly complex and more prone to widespread failures, including priority reversals that cannot be manually corrected ex-post.

The USMA leadership promptly recognized the potential for detectable priority reversals under the USMA-2020 mechanism, stemming from either the failure of BRADSO-IC or the presence of strategic BRADSO. A significant concern arose regarding the erosion of cadets' trust in the Army's branching process. Upon implementation for the Class of 2020, there was a noticeable increase in the number of cadets adversely affected compared to those impacted by USMA-2006 (Greenberg, Pathak, and Sönmez, 2023).

In response to these failures, the Army revisited the earlier reform proposal in Sönmez and Switzer (2013) and Sönmez (2013). This case serves as a reminder that reform is often triggered not by a superior alternative, but by a glaringly deficient existing institution (Sönmez, 2023). A partnership was forged with market design experts Parag Pathak and Tayfun Sönmez, with Major Kyle Greenberg spearheading the reform efforts at USMA.

In the rest of this section, we present the formal model and analysis that led to the new branching mechanism adopted by the Army for both USMA and ROTC, starting with the Class of 2021.

## 4.2 Formal Model

We define $I$ as the set of *cadets*. Let $T = \{t^0, t^+\}$ represent the possible *contractual terms* to secure a position, where $t^0$ denotes the base price and $t^+$ denotes the increased price. $B$ is a set of *branches*, where each branch $b \in B$ has a total of $q_b$ positions, with $q_b^f \leq q_b$ denoting the maximum number of positions available at the increased price $t^+$. These positions are referred to as *flexible-price positions*.

The preference relation of each cadet $i \in I$, denoted by $\succ_i$, is a linear order on $(B \times T) \cup \{\emptyset\}$. For every cadet $i \in I$ and branch $b \in B$, we assume that $(b, t^0) \succ_i (b, t^+)$, indicating that a position at the base price is always preferred over the same position at the increased price. The resulting cadet preferences are represented by $\mathcal{P}_i$.

Each branch $b$ has a *baseline priority order* $\pi_b$, which is a linear order on $I$. The set of branch baseline priorities is denoted by $\Pi$.

**The Price Responsiveness Policy.** The Army's branching system incorporates an embedded price responsiveness policy, akin to a marginal rate of substitution, which enhances the priorities of cadets willing to pay the increased price by committing to longer service periods.

For a given branch $b \in B$ and baseline priority order $\pi_b \in \Pi$, the relation $\omega_b$ is a linear order on $I \times T$ with two key properties: it reflects the baseline priority order $\pi_b$ for any fixed contractual term, and it is positively monotonic in contractual term for any given cadet. Formally,

1. for each $i, j \in I$ and $t \in T$, $\quad (i,t)\ \omega_b\ (j,t) \iff i\ \pi_b\ j$, and
2. for each $i \in I$, $\quad (i,t^+)\ \omega_b\ (i,t^0)$.

This linear order identifies the priority upgrade gained for the flexible-price positions by paying the increased cost. Let $\Omega_b(\pi_b)$ be the set of resulting price responsiveness policies.

### 4.2.1 Outcome and Mechanism

A *contract* is a triple $x \equiv \big(\mathbf{i}(x), \mathbf{b}(x), \mathbf{t}(x)\big) \in I \times B \times T$, indicating a position for cadet $\mathbf{i}(x)$ at branch $\mathbf{b}(x)$ at price $\mathbf{t}(x)$. Let $\mathcal{X} = I \times B \times T$ be the set of all contracts. Let $\mathcal{X}_i = \{x \in \mathcal{X} : \mathbf{i}(x) = i\}$ be the set of contracts that involve cadet $i$. Similarly, let $\mathcal{X}_b = \{x \in \mathcal{X} : \mathbf{b}(x) = b\}$ be the set of contracts that involve branch $b$.

An *allocation* is a set of contracts $X \subseteq \mathcal{X}$, such that:

1. For any $i \in I$, $\Big|\{x \in X : \mathbf{i}(x) = i\}\Big| \leq 1$.
2. For any $b \in B$, $\Big|\{x \in X : \mathbf{b}(x) = b\}\Big| \leq q_b$.
3. For any $b \in B$, $\Big|\{x \in X : \mathbf{b}(x) = b \text{ and } \mathbf{t}(x) = t^+\}\Big| \leq q_b^f$.

Let $\mathcal{A} \subseteq 2^{\mathcal{X}}$ be the set of allocations.

For an allocation $X \in \mathcal{A}$ and cadet $i \in I$, the *assignment* $X_i$ of cadet $i$ under allocation $X$ is defined as:

$$X_i = \begin{cases} (b,t) & \text{if } (i,b,t) \in X \\ \emptyset & \text{if } X \cap \mathcal{X}_i = \emptyset \end{cases}.$$

With a slight abuse of notation, $\mathbf{b}(X_i)$ indicates the branch of assignment $X_i$. A cadet $i \in I$ is unmatched under allocation $X \in \mathcal{A}$ if $X_i = \emptyset$.

A *mechanism* consists of a *message space* $\mathcal{S}_i$ for each cadet $i \in I$ along with an *outcome function* $\varphi : \bigtimes_{i \in I} \mathcal{S}_i \to \mathcal{A}$ that selects an allocation for each message profile. Let $\mathcal{S} = \bigtimes_{i \in I} \mathcal{S}_i$ be the set of message profiles.

A mechanism $(\mathcal{S}, \varphi)$ is a *direct mechanism* if $\mathcal{S}_i = \mathcal{P}_i$ for each $i \in I$.

### 4.2.2 The Army's Policy Objectives as Formal Axioms

Greenberg, Pathak, and Sönmez (2023) formulate Army's policy objectives as technical axioms and characterize the unique direct mechanism that satisfies all.

All but one of these axioms are defined both for allocations and mechanisms. In those cases, as before, a mechanism satisfies the axiom if its outcome satisfies the axiom for all message profiles.

An allocation $X \in \mathcal{A}$ satisfies *individual rationality* if, for any $i \in I$,

$$X_i \succ_i \emptyset.$$

Under individual rationality, no cadet is assigned a branch-price pair that they find unacceptable.

An allocation $X \in \mathcal{A}$ satisfies *non-wastefulness* if, for any $b \in B$ and $i \in I$,

$$\left|\left\{x \in X : \mathbf{b}(x) = b\right\}\right| < q_b \text{ and } X_i = \emptyset \qquad \Longrightarrow \qquad \emptyset \succ_i (b, t^0).$$

Under non-wastefulness, no position at a branch is left idle while a cadet is left unassigned, unless they would rather remain unassigned than receive the idle position at its base price.

An allocation $X \in \mathcal{A}$ satisfies *no priority reversals* if, for any $i, j \in I$, and $b \in B$:

$$\mathbf{b}(X_j) = b \text{ and } X_j \succ_i X_i \qquad \Longrightarrow \qquad j \; \pi_b \; i.$$

Under no priority reversals, no cadet $i$ prefers the branch-price package $(b, t)$ of another cadet $j$ to their own assignment, despite having a higher baseline priority for branch $b$.

We next present two auxiliary definitions highlighting our next axiom's intuition.

Given an allocation $X \in \mathcal{A}$ and a cadet $i \in I$ with $\mathbf{t}(X_i) = t^+$, a cadet $j \in I \setminus \{i\}$ has a *legitimate claim for a price-reduced version of cadet i's assignment* $X_i$ if:

1. $\big(\mathbf{b}(X_i), t^0\big) \succ_j X_j$, and
2. $(j, t^0) \;\; \omega_{\mathbf{b}(X_i)} \;\; (i, t^+)$.

Here, cadet $j$'s request for a position at branch $\mathbf{b}(X_i)$ at the base price $t^0$ is justified because the price responsiveness policy $\omega_{\mathbf{b}(X_i)}$ maintains their priority for a position at branch $\mathbf{b}(X_i)$ over cadet $i$, even if cadet $i$ opts to pay the increased price.

Given an allocation $X \in \mathcal{A}$ and a cadet $i \in I$ with $\mathbf{t}(X_i) = t^0$, a cadet $j \in I \setminus \{i\}$ has a *legitimate claim for a price-increased version of cadet i's assignment* $X_i$ if,

1. $\big(\mathbf{b}(X_i), t^+\big) \succ_j X_j$,
2. $(j, t^+) \;\; \omega_{\mathbf{b}(X_i)} \;\; (i, t^0)$, and
3. $\left|\left\{k \in I \; : \; \big(k, \mathbf{b}(X_i), t^+\big) \in X_{\mathbf{b}(X_i)}\right\}\right| < q^f_{\mathbf{b}(X_i)}$.

Here, cadet $j$'s request for a position at branch $\mathbf{b}(X_i)$ at the increased price $t^+$ is justified. This is because, even if cadet $i$ holds a higher baseline priority at branch $\mathbf{b}(X_i)$, the price responsiveness policy $\omega_{\mathbf{b}(X_i)}$ overrides this priority in favor of cadet $j$ provided that cadet $j$ offers a higher price than cadet $i$. Condition 3 is crucial here; otherwise, the assignment of cadet $i$'s position to cadet $j$ at the increased price $t^+$ would not be feasible due to the upper cap on the number of increased-price positions.

The next axiom formulates the idea that, the assignments respect the Army's price responsiveness policy.

An allocation $X \in \mathcal{A}$ satisfies *enforcement of the price responsiveness policy* if, no cadet $j \in I$ has a legitimate claim for either a price-reduced version or a price-increased version of the assignment $X_i$ of another cadet $i \in I \setminus \{j\}$.

Capturing the Army's objective to incite cadets' trust in the Army's branching process, the last axiom is the *strategy-proofness*: Truthful preference revelation is a dominant strategy for each cadet.

## 4.3 Army's New Mechanism: Dual-Price Cumulative Offer Mechanism

The *Dual-Price Cumulative Offer (DPCO) mechanism* is a direct mechanism based on the cumulative offer procedure (Hatfield and Milgrom, 2005), complemented by the following choice rule.

**Dual-Price Choice Rule $C_b^{DP}$.** Given a branch $b \in B$ and a set of contracts $X \in \mathcal{X}_b$, select (up to) $q_b$ contracts with distinct cadets in two steps as follows:

**Step 1.** For the base-price positions, exclusively select contracts at the base price with the highest-priority cadets according to their baseline priorities.

**Step 2.** For the flexible-price positions, select the highest-priority remaining contracts based on the price responsiveness policy $\omega_b$.

■

**Dual-Price Cumulative Offer (DPCO) Mechanism.**

Fix any linear order of cadets, such as the OML.[32].

**Step 0.** No contract is on hold initially.

**Step k.** (k > 0)

- The highest-OMS cadet currently without a contract on hold, whom we refer to as $i_k$, offers their most-preferred previously-unrejected contract $x_k$ to the branch of the contract $\mathbf{b}(x_k)$, and
- considering all offers $X_k$ it has received up to (and including) Step k, branch $\mathbf{b}(x_k)$ holds the contracts in $C^{DP}_{\mathbf{b}(x_k)}(X_k)$, and rejects all others.

We conclude the procedure if there are no cadets remaining with acceptable contracts that haven't been rejected or if no contracts are rejected, indicating that all contracts on hold are finalized. Otherwise, we proceed to Step $k+1$.

■

## 4.4 The Characterization Result

The following result provides justification for the Army's reform of its branching system, in partnership with market design economists, at both USMA and ROTC, commencing with the Class of 2021.

[32]The outcome remains independent of the choice of this tie-breaker, and any linear order would suffice (Kominers and Sönmez, 2016)

**Theorem 12 (Greenberg, Pathak, and Sönmez, 2023)** *Fix a profile of baseline priority orders* $(\pi_b)_{b \in B} \in \Pi$ *and a profile of price responsiveness policies* $\left(\omega_b\right)_{b \in B} \in \bigtimes_{b \in B} \Omega_b(\pi_b)$*. A direct mechanism satisfies individual rationality, non-wastefulness, enforcement of the price responsiveness policy, no priority reversals, and strategy-proofness if, and only if, it is the DPCO mechanism.*

As presented in Greenberg, Pathak, and Sönmez (2023), the entire analysis, including the characterization theorem, extends to multiple prices.

### 4.4.1 Broader Implications of Analysis

Zhou and Wang (2021) explores public high school admissions in China under the *ZX Policy* (Ze Xiao). In this price responsiveness policy, a fraction of the seats are available with an increased tuition. Baseline priorities are based on scores on a centralized exam. The higher-tuition contract increases this score by a fixed amount for the ZX-eligible seats. Shanghai and Tianjin have a single ZX tuition level, making these applications completely analogous to the Army's problem. In some cities, there were multiple tuition levels where higher tuition levels resulted in higher adjustments to student scores. This policy was discontinued after 2015.

## 4.5 Broader Implications and Proof-of-Concept for Minimalist Market Design

The Army deemed the design a success and adopted it for ROTC ahead of schedule. Greenberg, Pathak, and Sönmez (2023) argue that this serves as a proof-of-concept for a new institutional design paradigm called *minimalist market design* (Sönmez, 2023). This institution design paradigm may be especially valuable in the following contexts:

- When the need for change is not clearly established, and an outsider seeks to initiate reform.
- When the "intended" institution is apparent, but identifying it requires formalism and technical expertise.
- When the institution's mission cannot be fully described by a single-dimensional objective function.

On a broader scale, the Army's economist-guided branching reform underscores the significance of fundamental and customized theories in policy-oriented economic research, especially for the case of aspired design.

# 5 Affirmative Action in India

In Section 4, we explored the challenge faced by Army officials in integrating BRADSO incentives with the branching system. They needed to create a mechanism to allocate positions based on two criteria. While they managed to implement policies based on one criterion using a priority mechanism, incorporating a second proved challenging.

These challenges aren't confined to the military domain. Complex systems like the Army's branching structure demand expertise in market design. Therefore, academic market designers should broaden their focus beyond economic principles like preference utilitarianism and explore various normative criteria.

As underscored by Li, 2017, market design must consider ethics because policymakers often struggle to clearly articulate their ethical requirements. Despite their understanding of the environment, they may lack precision in expressing their ethical needs. Moreover, even if policymakers can articulate their ethical requirements, they may lack the expertise to design mechanisms that align with these principles.

Next, we will explore how the involvement of design economists could have helped the judiciary avoid costly crises over several decades in India. Despite lacking training in formal methods, Indian justices have significantly influenced the normative principles underpinning the country's affirmative action policies. They are often instrumental in designing the mechanisms to implement these principles.

India is home to the world's most extensive affirmative action program, encompassing recruitment for public sector jobs, admissions to public universities, and elections for legislative seats. This program is explicitly outlined and governed by the country's Constitution, shaped by congressional amendments, and frequently subject to legal challenges resolved by the Supreme Court. Our objective in this section is to illuminate the key components of this program through the lens of matching theory, providing a direct application of reserve systems discussed in Chapter 1, Section 4.2. Recently, Sönmez and Yenmez, 2022a formalized these principles in implementing the most critical Indian affirmative action policies using the framework of reserve systems.[33]

[33] From the market design side, Baswana et al. (2019) reported the design and implementation of a matching mechanism for placement to the most prestigious schools in India incorporating merit and affirmative action-based admissions. Some other works related to or inspired by affirmative action policies in India exist. Echenique and Yenmez (2015) gave affirmative action policies in India as an application of controlled choice rules in matching theory. Aygün and Turhan (2020) considered a centralized matching model in which students may have heterogeneous preferences over whether they receive a position at a college through affirmative action or regular admissions. They used Indian affirmative action as an example of why this could happen. Using an empirical approach, Thakur (2021) tested the change of the impact of the Indian Civil Service state assignment mechanism on the population of individuals affected. Finally, Evren and Khanna (2024) studied a two-dimensional dynamic apportionment problem to design a roster system. Single-dimensional roster point systems are used in Indian affirmative action, and they pinpoint which category of individuals are qualified for the next arriving position. On the other hand, recent legal developments in India suggest that hiring decisions in teaching institutions should respect affirmative action for department-wide affirmative action constraints and institution-wide affirmative action constraints.

In this section, we summarize the evolution of this system through the lens provided by Sönmez and Yenmez, 2022a. Given the complexity of the principles laid down in law, the axiomatic framework introduced by Sönmez and Yenmez, 2022a is invaluable. Through this framework, certain normative and legal principles based on the country's laws are translated into fairness axioms in matching theory (refer to Section 4 of Chapter 1 for the basis of these axioms). We'll demonstrate that, akin to the application of the Army's branching system presented earlier, minimalist market design (Sönmez, 2023) is particularly relevant in this context.

## 5.1 Vertical and Horizontal Reservations

In this subsection, we discuss the historical institutional background governing affirmative action in India. Throughout the section, we follow the terminology of Sönmez and Yenmez (2022a) and refer to a reserve system as a "choice rule".

The affirmative action program in India, commonly known as the "reservation system," is sanctioned by the 1950 Constitution of the Union of India. It provides affirmative action for various disadvantaged groups in hiring for government jobs, admissions to public universities, and elections for legislative seats. For each disadvantaged group, a fraction of positions is *reserved*.

Higher-level affirmative action provisions, called *vertical reservations* (VR) mainly include four constitutionally designated groups:

- *Scheduled Castes* (*SC*) and *Scheduled Tribes* (*ST*) are the original beneficiaries.
- *Other Backward Classes* (*OBC*) and *Economically Weaker Section* (*EWS*) are recognized these protections through constitutional amendments.

Lower-level affirmative action provisions, called *horizontal reservations* (HR), are provided for other disadvantaged groups such as persons with disabilities, women, ex-servicemen, etc.

Several Supreme Court of India (SCI) cases and their decisions played an important historical, institutional, and analytical role in shaping these policies.

### 5.1.1 Stand-Alone Implementation of VR Policy: Indra Sawhney (1992)

The concepts of vertical and horizontal reservations were introduced in the landmark Supreme Court judgment *Indra Sawhney vs. Union of India (1992)* (Supreme Court of India, 1992), also known as the *Mandal Commission Case*.

> *"A little clarification is in order at this juncture: all reservations are not of the same nature. There are two types of reservations, which may, for the sake of convenience, be referred to as vertical reservations and horizontal reservations. The reservation in favour of scheduled castes, scheduled tribes and other backward classes [under Article 16(4)] may be called vertical reservations whereas reservations in favour of physically*

*handicapped [under clause (1) of Article 16] can be referred to as horizontal reservations. Horizontal reservations cut across the vertical reservations – what is called interlocking reservations."*

*Vertical reservations* (or *VR protections*) correspond to provisions sanctioned under Article 16(4) of the Constitution (Central Government Act, 1949). As a reparatory and compensatory mechanism, they were originally intended for historically discriminated groups such as SC, ST, and OBC. With a controversial constitutional amendment, they are also offered for EWS since 2019.

The decision announced that the reservations were to be earmarked in the form of a *set aside* protection: Positions secured based on merit do not count against VR-protected positions.

*"It may well happen that some members belonging to, say Scheduled Castes get selected in the open competition field on the basis of their own merit; they will not be counted against the quota reserved for Scheduled Castes; they will be treated as open competition candidates."*

Together, individuals who do not belong to any VR-protected category make up the *general category* individuals. Positions that are not VR protected are referred to as *open category* (or *open*) positions. Any individual is eligible for these open positions.

*Horizontal reservations* (*HR protections*), as sanctioned under Article 16(1) of the Constitution, aim to provide a minimum guarantee for disadvantaged groups such as persons with disabilities, women, ex-servicemen, and others. Implementation of this secondary affirmative action policy ensures that positions secured on merit still count against HR-protected positions.

There is one other key difference between VR and HR protections. VR policy is implemented as "hard reserves"; they cannot be awarded to individuals who are not members of the protected group. On the contrary, HR policy is implemented as "soft reserves"; they simply provide preferential treatment to members of the protected group. However, if there are insufficient members of the target group, HR-protected positions can be awarded to other individuals. Refer to Section 4.2.3 of Chapter 1 of this handbook for a discussion on hard reserves and soft reserves.

In India, individuals can belong to at most one *Vertical Reservation* (*VR*)-protected category. When HR protections are absent, implementing VR protections is straightforward using the Over-and-Above Choice Rule that was discussed in Chapter 1, Section 4. We restate this rule:

**Over-and-Above Choice Rule.**

**Step 1.** Allocate open positions to the highest merit-ranking individuals.

**Step 2.** For each VR-protected group, allocate the reserved positions to the highest merit-ranking members of the group who remain unassigned.

■

### 5.1.2 Joint Implementation of VR and HR Policies: Anil Kumar Gupta (1995)

Most applications involve both HR and VR protections. Throughout India, 4-5% of positions are HR protected for persons with disabilities. Moreover, in many states, 30-35% are HR protected for women. However, since HR-protected groups can overlap with VR-protected groups, concurrent implementation poses challenges.

To address this issue, the Supreme Court's judgment in Anil Kumar Gupta vs. State of U.P. (1995) (Supreme Court of India, 1995) introduced the following choice rule, mandating it throughout India:

**SCI-AKG Choice Rule.**

**Step 1a.** Provisionally allocate open positions to the highest merit-ranking individuals.

**Step 1b.** Utilizing individuals who are not VR-protected, make any needed adjustments to provisional awardees in Step 1a to accommodate open-category HR protections.[34]

**Step 2a.** For each VR-protected group, provisionally allocate the reserved positions to the highest merit-ranking members of the group who remain unassigned.

**Step 2b.** Make any needed adjustments to provisional awardees in Step 2a to accommodate HR protections within VR-protected positions.

■

Observe that, without HR protections, Steps 1b and 2b become redundant, and the SCI-AKG choice rule becomes equivalent to the Over-and-Above choice rule. Unlike the Over-and-Above choice rule, however, the SCI-AKG choice rule has a critical flaw that has introduced two key anomalies, sparking thousands of litigations in India over the next 25 years.

Before presenting the failures of the SCI-AKG choice rule, we demonstrate how it works with an example.

**Example 6 (The Mechanics of the SCI-AKG Choice Rule)** *Suppose there is a single VR-protected reserve category besides the open category, and there are 300 positions in a government department, 240 of which are reserved for the open category and 60 reserved for the reserve category. Also, within each category, there is a minimum guarantee of 50% for women (120 seats reserved within open category and 30 within the reserve category).*

*There are 300 general-category men ($m^G$) and 100 general-category women ($w^G$) who are not eligible for the reserve category positions. Additionally, there are 100 reserve-category men ($m^R$) and 100 reserve-category women ($w^R$).*

[34] We will be more precise about how this is done below in Example 6.

*Suppose the merit exam scores — used as the admission criterion for this government job — are distributed uniformly between 1 and 100 so that the same distribution is obtained in each group, i.e., given a score, there are 3 general-category men with that score and 1 from each other group.*

*We use several figures to demonstrate the iterative procedure of the SCI-AKG choice rule. Before the assignment starts, Figure 23 summarizes the setup.*

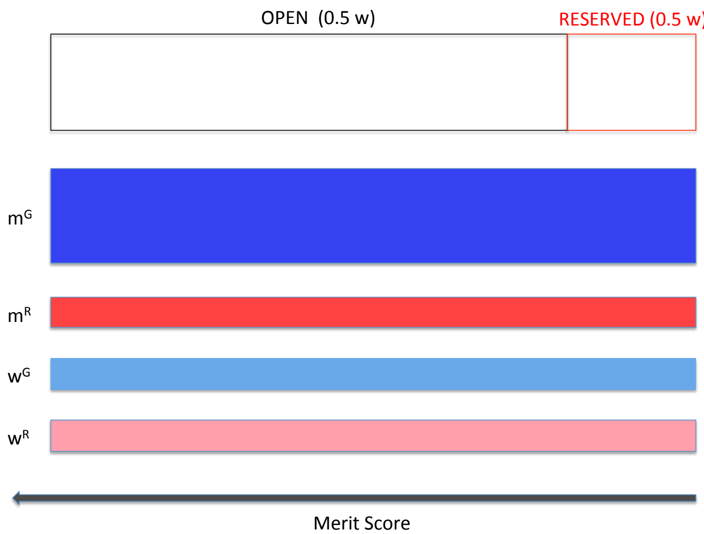



**Figure 23:** Setup of Examples 6 and 7.

*SCI-AKG choice rule allocates positions through the following steps:*

***Step 1a.*** *Open category positions are provisionally assigned to the highest-merit individuals, resulting in 120 positions being assigned to members of the group $m^G$ and 40 positions being assigned to each of the groups $m^R$, $w^G$, and $w^R$ (See Figure 24).*

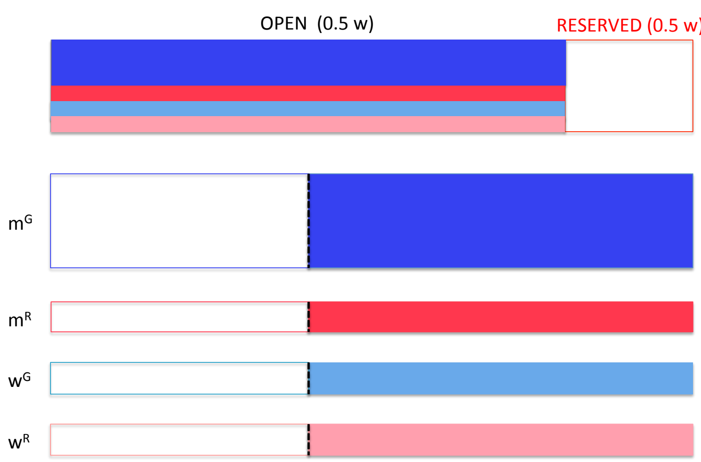



**Figure 24:** Step 1a of SCI-AKG choice rule in Example 6.

*__Step 1b.__ In Step 1a, only 80 women are admitted to the open category, although there is a minimum guarantee of 120 seats for women within the open category. Thus, as part of the HR policy "adjustment", 40 of the lowest-scoring men are removed from open category seats: 30 general-category men and 10 reserve-category men are displaced (see Figure 25).*

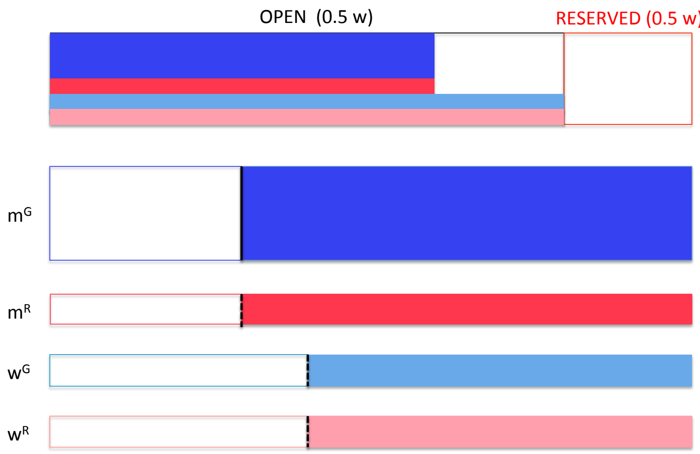



**Figure 25:** Step 1b - removal of "extra" men in the SCI-AKG Rule in Example 6 as part of open-category HR adjustment.

*To complete the open-category HR policy adjustment process, 40 general-category women with the next highest scores are admitted to fill these vacated positions, resulting in a total of 80 open category positions being awarded to general-category women (see Figure 26).*

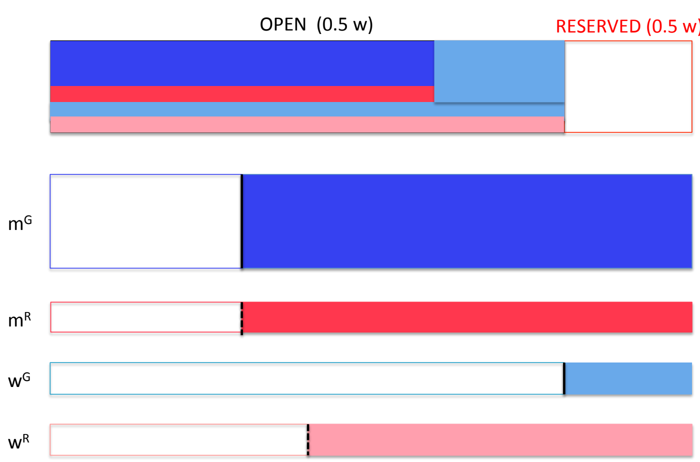



**Figure 26:** Step 1b - replacement of women to vacated positions in the SCI-AKG Rule in Example 6 as part of open-category HR adjustment.

*__Step 2.__ Afterwards, reserve category positions are provisionally assigned to the remain-*

*ing highest-scoring members of the reserve category. The 10 reserve category men who were displaced in Step 1b from their open category seats to fulfill the HR protections for women all have higher scores than all remaining reserve-category women. In Step 2a, each of these men receives a position, followed by 25 reserve-category men and 25 reserve-category women. As there is a minimum guarantee of 30 positions for women within the reserve category, adjustments are necessary. In Step 2b, the 5 lowest-merit reserve-category men are displaced and replaced with the remaining highest-merit reserve-category women. Eventually, 30 reserve-category women and 30 reserve-category men are chosen for the reserve-category positions (see Figure 27 for the final assignment).*

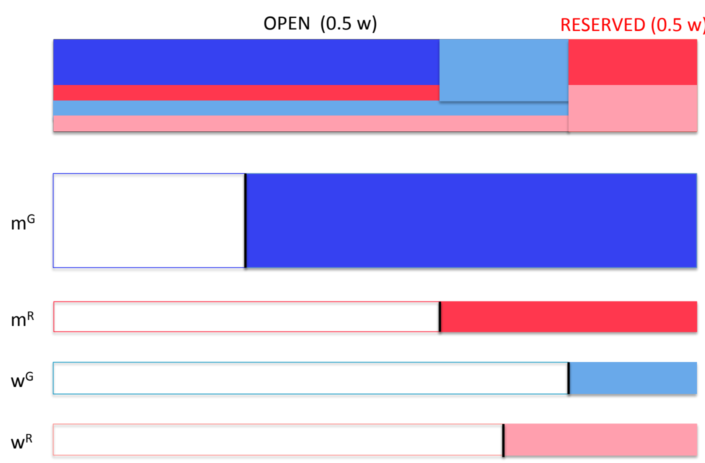



**Figure 27:** Final outcome of the SCI-AKG Rule after Step 2 in Example 6.

*The outcome of the SCI-AKG choice rule is such that, subject to awarding positions to the highest-scoring individuals within each group:*

- *90 out of 300 general-category men,*
- *60 out of 100 reserve-category men,*
- *80 out of 100 general-category women, and*
- *70 out of 100 reserve-category women*

*receive positions.*

### 5.1.3 Limitations of the SCI-AKG Choice Rule and Their Adverse Implications

Example 6 illustrates two related shortcomings of the SCI-AKG choice rule.

In Step 1b, as we adjust the provisional assignment from Step 1a to respect the HR protections for women in the open category, the reserve category women are not considered for the vacated seats. Some may be more deserving than the general-category women who were assigned to the vacated seats. This is especially problematic if some of these more deserving reserve-category women do not receive a VR-protected position either in Step 2. This is exactly what has happened in this example. As seen in Figure 27, the lowest-merit individual accepted from each group, $w^G$, has a lower cutoff than $w^R$. This creates justified envy for the 10 reserve-category women who are not matched but each have a score that is at least as high as the lowest-score general-category woman accepted (with 9 having strictly higher scores). In essence, low-privilege women lose positions to higher-privilege women despite having higher merit scores, directly conflicting with the essence of affirmative action.

Due to the justified envy, the SCI-AKG choice rule also incentivizes adversely affected reserve-category women to conceal their VR-protection status. They would be better off if they never revealed their reserve category.

Formally, SCI-AKG choice rule violates the following two properties, often resulting in outcomes that contradict the philosophy of affirmative action.

*No Justified Envy* (Sönmez and Yenmez, 2022a): A higher-merit-ranking individual cannot lose a position to a lower-merit-ranking individual, unless the latter is of strictly lower privilege.

*Incentive Compatibility* (Aygün and Bó, 2021): An individual never loses a position solely due to declaring their reserve-eligible attributes.

**Root Cause of the Failures of the SCI-AKG Choice Rule.** Observe that, the key issue with the SCI-AKG choice rule lies in its Step 1b. VR-protected individuals are not considered for open positions when adjustments for HR policy are made. Thus, the root cause of the failures is the denial of VR-protected individuals of their open category HR protections if they claim their VR protections.

In practice, these problems with the SCI-AKG choice rule had a profound effect in the country, leading to thousands of lawsuits with a waste of resources in litigation efforts and an effective halt in hiring practices in some places. We summarize some of these judicial cases filed in the decades following 1995 to demonstrate the severity of the problems due to failures of no justified envy and incentive compatibility.

**Litigations Related to the Axiom of No Justified Envy.** In numerous cases, public institutions resisted adopting the Supreme Court-mandated procedure and allowed reserve-category candidates to benefit from open-category HR protections. This resistance often led to litigation from lower merit-ranking general-category candidates who were not selected.

*1. Rajeshwari vs State (2013)* (Rajasthan High Court, 2013): A large-scale litigation with 120 petitions against the State of Rajasthan ensued when the State allowed reserve-category women to benefit from open-category HR protections. The High Court ruled that the State was at fault and ordered the State to adopt the Supreme Court-mandated procedure.

*2. Ashish Kumar Pandey and Others vs State (2016)* (Allahabad High Court, 2016): This case resembled Rajeshwari vs State (2013), with 25 petitioners litigating against the State of Uttar Pradesh for allowing reserve-category women to benefit from open-category HR protections. This case was polarizing, as the counsel for petitioners argued that the error was intentional, stating,

> *"The action of the Board is not only motivated but purports to take forward the unwritten agenda of the State Government to accommodate as many numbers of OBC/SC candidates in the open category."*

The judge of the case ruled that the State must correct their erroneous application of HR protections, emphasizing that the State played foul:

> *"There is merit in the submission of the learned counsel for the petitioners that the conduct of the members of the Board appears not only mischievous but motivated to achieve a calculated agenda by deliberately keeping meritorious candidates out of the select list [. . .] I am constrained to hold that both the State and the Board have played fraud on the principles enshrined in the Constitution with regard to public appointment."*

The State appealed the judgment and lost the appeal as well.

*3. Smt. Megha Shetty vs State (2013)* (Rajasthan High Court - Jodhpur, 2013): This case was similar to the earlier ones, with a general-category petitioner litigating against the State for allowing reserve-category women to benefit from open category HR protections. Unlike the earlier cases, this case was dismissed at the High Court. The petitioner appealed the decision, bringing the case to a larger bench of the High Court. The appeal was also dismissed. The judges had difficulty entertaining the possibility that a procedure mandated by the Supreme Court could possibly allow for justified envy:

> *"The outstanding and important feature to be noticed is that it is not the case of the appellant-petitioner that she has obtained more marks than those 8 OBC (Woman) candidates..."*

In numerous other cases, a public institution that used the Supreme Court-mandated procedure faced litigation from reserve-category candidates who were not selected despite having higher merit scores than their general-category counterparts who were selected.

*4. Asha Ramnath Gholap (2016)* (Bombay High Court, 2016): Following the law, the State used the Supreme Court-mandated choice rule, resulting in an instance of justified envy. A reserve-category petitioner brought the case to the High Court. The judges granted the petition, stating that a candidate cannot be denied an open-category position based on their reserve-category membership.

**Litigations Related to the Failure of Incentive Compatibility.** Some cases were examples of wrongful implementation and possible misconduct. While applicants are entitled to declare their social categories or traits, they are not required to. Since the SCI-AKG choice rule is not incentive-compatible, withholding this information may make sense.

*5. Shilpa Sahebrao Kadam (2019)* (Bombay High Court, 2019a): Several candidates withheld their reserve-category memberships to take advantage of the open-category HR protections. Authorities requested personal information to identify their reserve-category memberships and evaluated their applications as if these candidates claimed their VR protections. The candidates were all denied positions despite having higher merit scores than their general-category counterparts who were selected due to open-category HR protections. So they went to court.

The petitioners lost the case despite the "faulty" implementation. Indeed, the faulty implementation seems to be "systematic" and "intentional" as revealed by the court proceedings.

> *"According to Respondent - Maharashtra Public Service Commission, in view of the Circular dated 13.08.2014, only the candidates belonging to open (Non-reserved) category can be considered for open horizontally reserved posts meaning thereby, the reserve-category candidates cannot be considered for open horizontally reserved post. Reference is made to a communication issued by the Additional Chief Secretary (Service) of the State of Maharashtra dated 26.07.2017, whereunder it is prescribed that a female candidate belonging to any reserve-category, even if tenders application form seeking employment as an open category candidate, the name of such candidate shall not be recommended for employment against an open category seat."*

*6. Smt. Tejaswini Raghunath Galande (2019)* (Bombay High Court, 2019b): The petitioner declared her reserve-category membership despite the lack of VR-protected positions for her category. She lost access to open-category HR protections, resulting in an instance of justified envy. Prior to bringing her case to the High Court, she filed a petition to a lower court. Her case was dismissed. She appealed at the High Court, which in turn was granted. There are, however, similar petitions which have been dismissed.

### 5.1.4 Addressing the Failures: Two-Step Minimum Guarantee Choice Rule

Since the root cause of the crisis from SCI-AKG choice rule is the denial of VR-protected individuals' open-category HR protections, a resolution lies in removing this restriction. That is, in Step 1b of the procedure, all individuals are considered the adjustments for open-category HR protections rather than only the members of the general category who are ineligible for VR protections. For reasons that will be clear in Section 5.3, we refer to this modified choice rule as *two-step minimum guarantee* (2SMG) choice rule.

Using the same setup as in Example 6 (see Figure 23), we next present how 2SMG choice rule works.

**Example 7 (Mechanics of the 2SMG Choice Rule)**
***Step 1a.*** *Open category positions are provisionally assigned to the highest-merit individuals, resulting in 120 positions being assigned to members of the group $m^G$ and 40 positions being assigned to each of the groups $m^R$, $w^G$, and $w^R$ (see Figure 28).*

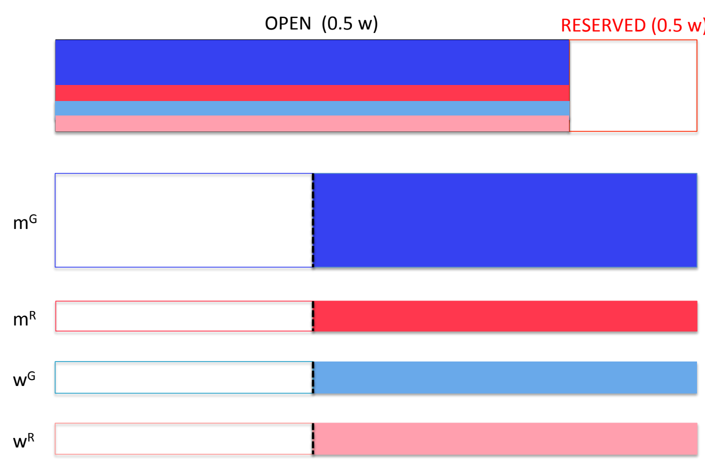



**Figure 28:** Step 1a of 2SMG Choice Rule in Example 7.

***Step 1b.*** *In Step 1a, only 80 women are admitted to the open category, although there is a minimum guarantee of 120 seats for women within the open category. Thus, as part of the HR policy "adjustment", 40 of the lowest-scoring men are removed from open category seats: 30 general-category men and 10 reserve-category men are displaced (see Figure 29).*

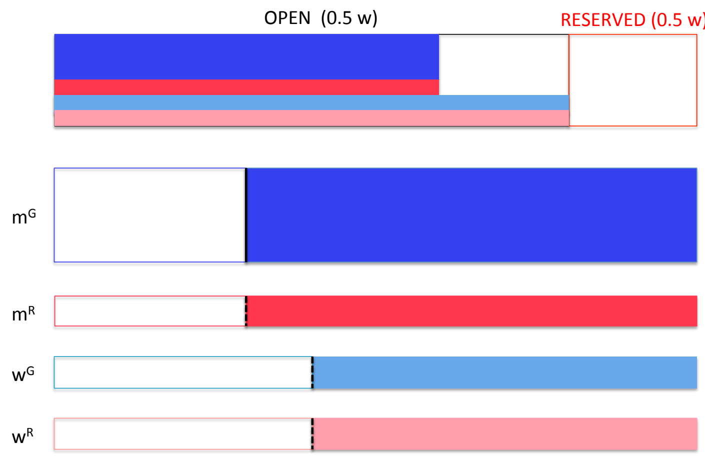



**Figure 29:** Step 1b - replacement of women to vacated positions in the 2SMG choice rule in Example 7 as part of open-category HR adjustment.

*This is where the procedure differs from the SCI-AKG choice rule: To complete the open-category HR policy adjustment process, 40 women with the next highest scores—20 each from general and reserve categories—are admitted to fill these vacated positions, resulting in a total of 60 open category positions being awarded to general-category women and 60 to reserve-category women (see Figure 30).*

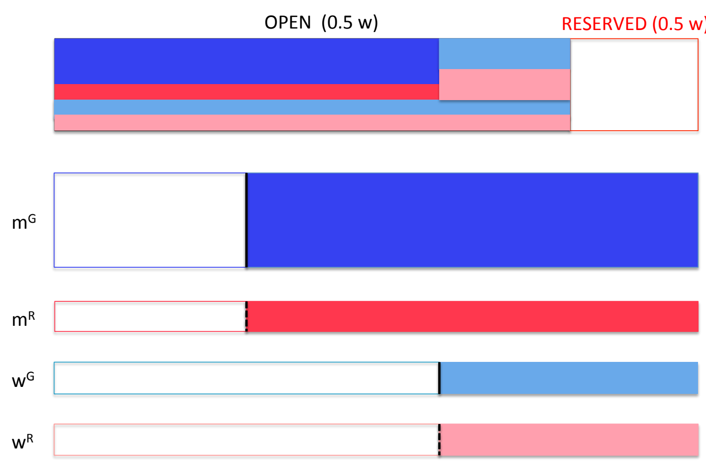



**Figure 30:** Step 1b - replacement of women to vacated positions in the 2SMG choice rule in Example 7 as part of open-category HR adjustment.

***Step 2.*** *Afterwards, reserve category positions are provisionally assigned to the remaining highest-scoring members of the reserve category. In Step 2a, 40 reserve-category men are provisionally placed as they have the highest merit scores among the remaining reserve-category individuals, followed by 10 reserve-category men and 10 women for the remaining 20 seats.*

*Since this assignment violates the minimum guarantee of 30 positions for reserve-category women, 20 reserve-category men with the lowest merit scores are displaced in Step 2b and replaced with the remaining 20 reserve-category women with the highest merit scores. Thus, in the end, 30 reserve-category women and 30 reserve-category men are placed in Step 2 (see Figure 31).*

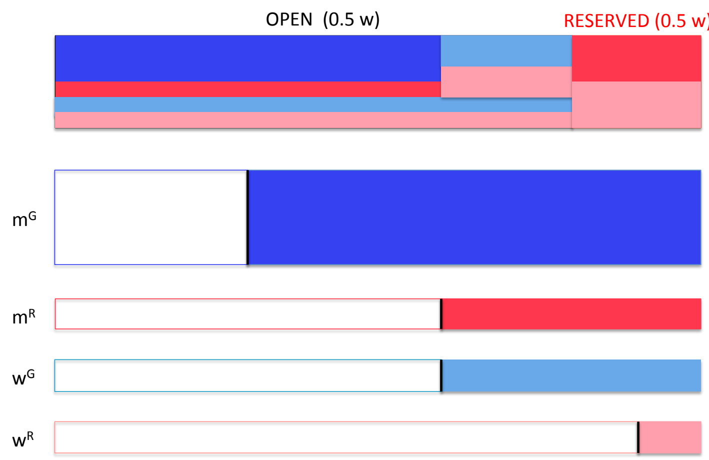



**Figure 31:** Final outcome of 2SMG Choice Rule after Step 2 in Example 7.

*The outcome of 2SMG choice rule is such that, subject to awarding positions to the highest-scoring individuals within each group:*

- *90 out of 300 general-category men,*
- *60 out of 100 reserve-category men,*
- *60 out of 100 general-category women, and*
- *90 out of 100 reserve-category women*

*receive positions.*

Directly addressing the root cause of the failures of the SCI-AKG choice rule, it is easy to observe that the 2SMG choice rule satisfies the axioms of no justified envy and incentive compatibility (Sönmez and Yenmez, 2022a). In Example 7, eligible for open-category HR protections, not only are the reserve-category women no longer disadvantaged compared to general category women, but they also benefit from positive discrimination as intended by the Constitution.

#### 5.1.5 Constitutional Resolution: Saurav Yadav vs. State of Uttar Pradesh (2020)

In a March 2019 working paper, Sönmez and Yenmez, 2019—later published in Sönmez and Yenmez, 2022a—documented the failures of the SCI-AKG choice rule and linked its shortcomings to its failure of the *no justified envy* axiom. They proposed the

2SMG choice rule as a solution. Despite causing significant disruption in the country for a quarter of a century, the Supreme Court had not addressed the failure of the SCI-AKG choice rule before the circulation of this paper. However, this changed in December 2020 with a landmark Supreme Court judgment. Using arguments similar to those presented by Sönmez and Yenmez (2019, 2022a), the justices arrived at the same conclusions in Saurav Yadav vs. State of Uttar Pradesh (2020) (Supreme Court of India, 2020) as Sönmez and Yenmez had earlier. Specifically:

- The *no justified envy* axiom—first time formulated in a Supreme Court judgment—is now mandated for all choice rules used in India.
- The long-standing SCI-AKG choice rule was rescinded due to its failure to satisfy the *no justified envy* axiom.
- As a possible replacement for the SCI-AKG choice rule, the 2SMG choice rule is endorsed, although it is not explicitly mandated.
- The implementation of VR protections in the presence of HR protections has gained clarity to a degree previously unavailable.

Sönmez, 2023 interprets the parallel with Sönmez and Yenmez, 2019, 2022a as external validation for minimalist market design, briefly discussed in Section 4.5.

We proceed with our section, conducting a formal analysis as an application of matching theory to elaborate on the broader implications of this reform in India.

## 5.2 Formal Model

There are $q$ identical positions that need allocation to a set of individuals $I$. Each individual $i \in I$ requires one position and endowed with a distinct merit score $\sigma(i) \in \mathbb{R}_+$. While individuals with higher merit scores naturally have stronger claims for a position in the absence of affirmative action (AA) policies, various groups benefit from two types of AA policies: *Vertical Reservations* (*VR*) and *Horizontal Reservations* (*HR*).

The VR policy is managed through a system of category membership. Let $R$ be the set of reserve-eligible categories, and $g \notin R$ denote the general category for those ineligible for VR protections. The function $\rho : I \to R \cup \{\emptyset\}$ represents the (reserve-eligible) category membership function. Each individual belongs to a single category in $R \cup \{g\}$, so that $\rho(i) = c$ indicates that individual $i$ belongs to the reserve-eligible category $c \in R$, and $\rho(i) = \emptyset$ indicates that individual $i$ belongs to the general category $g$.

Let $q^c$ denote the number of *category-c* positions set aside for members of a reserve-eligible category $c \in R$. For any reserve-eligible category $c \in R$, an individual $i \in I$ is *eligible for category-c positions* if $\rho(i) = c$. Let $q^o = q - \sum_{c \in R} q^c$ denote the number of the *open category* (or *category-o*) positions. All individuals are eligible for open category positions. Let $V = R \cup \{o\}$ denote the set of *vertical categories* for positions.

For each vertical category $v \in V$, let $I^v \subseteq I$ denote the set of individuals who are eligible for positions in category $v$. Therefore, $I^o = I$, as all individuals are eligible for the open category, and $I^c = \{i \in I : \rho(i) = c\}$ for any $c \in R$.

Given a category $v \in V$, a *category-$v$ choice rule* (or *single-category choice rule for category $v$*) is a function $\mathcal{C}^v : 2^I \to 2^{I^v}$ such that, for any $J \subseteq I$,

$$\mathcal{C}^v(J) \subseteq J \cap I^v \quad \text{and} \quad |\mathcal{C}^v(J)| \leq q^v.$$

A (multi-category) *choice rule* is a function $\mathcal{C} = \left(\mathcal{C}^\nu\right)_{\nu \in V} : 2^I \to \bigtimes_{\nu \in V} 2^{I^\nu}$ such that, for any set of individuals $J \subseteq I$,

1. for any vertical category $v \in V$,

$$\mathcal{C}^v(J) \subseteq J \cap I^v \quad \text{and} \quad |\mathcal{C}^v(J)| \leq q^v,$$

2. for any two distinct vertical categories $v, v' \in V$,

$$\mathcal{C}^v(J) \cap \mathcal{C}^{v'}(J) = \emptyset.$$

For any choice rule $\mathcal{C} = \left(\mathcal{C}^\nu\right)_{\nu \in V}$, the resulting *aggregate choice rule* $\widehat{\mathcal{C}} : 2^I \to 2^I$ is given as

$$\widehat{\mathcal{C}}(J) = \bigcup_{\nu \in V} \mathcal{C}^\nu(J) \qquad \text{for any } J \subseteq I.$$

For any set of candidates, a choice rule $\mathcal{C}$ indicates which ones receive positions and from which categories. The resulting set of candidates selected by the choice rule $\mathcal{C}$ is given by the aggregate choice rule $\widehat{\mathcal{C}}$.

The HR policy is managed through a system of trait ownership. Let $T$ be a set of traits associated with HR protections. Let $\tau : I \to 2^T$ be the trait function that identifies each individual's traits. For any vertical category $v \in V$ and trait $t \in T$, let $q_t^v$ be the minimum number of category-$v$ positions that must be awarded to eligible individuals with trait $t$. We refer to these positions as the category-$v$ HR-protected positions for trait $t$.

While each individual is a member of a single category in $R \cup \{g\}$ in India, they may have multiple traits. We refer to HR policies where an individual can have at most one trait as *non-overlapping HR protections*, and those where an individual can have multiple traits as *overlapping HR protections*.

We start with the more basic case of non-overlapping HR protections as it is simpler, and more importantly, the court rulings are presented for this case. The failure of the SCI-AKG choice rule is already prominent in this case. Although the SCI-AKG choice rule is not well-defined for the more general case, in Section 5.4, we will present an analysis of this case as well. We will show that there is a choice rule in this case as well that uniquely satisfies the Supreme Court's mandates in Saurav Yadav (2020) (Supreme Court of India, 2020).

## 5.3 Analysis for Non-Overlapping HR Protections

HR protections are provided within each vertical category on a *minimum guarantee* basis. This means that positions obtained without invoking any HR protection still accommodate the HR protections. Institutions in India typically implement the HR policy through the kind of "adjustments" illustrated in Examples 6 and 7. However, there is a more direct way to implement this policy.

Given any vertical category $v \in V$, HR protections within category-$v$ can be implemented with the following choice rule, first formalized in Hafalir, Yenmez, and Yildirim (2013).

**Minimum Guarantee Choice Rule $\mathcal{C}^v_{mg}$.**
For any set of individuals $J \subseteq I^v$,

**Step 1.** For each trait $t \in T$, assign HR-protected positions to highest merit-score individuals in $J$ who have trait $t$.

**Step 2.** For positions unfilled in Step 1 (open or HR-protected), choose the highest merit-score individuals in $J$ who are still unassigned.

■

Since VR protections are implemented on an over-and-above basis and HR protections are implemented within each vertical category on a minimum guarantee basis, Sönmez and Yenmez (2019, 2022a) proposed a two-step implementation of the minimum guarantee choice rule for the joint implementation of the two policies. This involves first applying it for the open category, and subsequently for each reserve-eligible category.

**2-Step Minimum Guarantee Choice Rule (2SMG) $\mathcal{C}^{2s}_{mg} = \left(\mathcal{C}^{2s,\nu}_{mg}\right)_{\nu \in V}$.**
Given a set of individuals $J \subseteq I$,

$$\textbf{Step 1.} \quad \mathcal{C}^{2s,o}_{mg}(J) = \mathcal{C}^{o}_{mg}(J),$$
$$\textbf{Step 2.} \quad \mathcal{C}^{2s,c}_{mg}(J) = \mathcal{C}^{c}_{mg}\Big(\big(J \setminus \mathcal{C}^{o}_{mg}(J)\big) \cap I^c\Big) \quad \text{for any } c \in R.$$

■

Importantly, when HR protections are non-overlapping, this procedure is equivalent to the choice rule presented in Section 5.1.4, which is derived as a minimalist amendment of the SCI-AKG choice rule by ensuring everyone's eligibility for HR policy adjustments in Step 1b (Sönmez and Yenmez, 2022a). By processing the HR-protected positions first within each category, the current formulation becomes more direct, eliminating any need for adjustments.

Saurav Yadav (2020) (Supreme Court of India, 2020) marks the end of an era where a three-judge bench of the Supreme Court brought an end to the 25-year tenure of the AKG-SCI choice rule and endorsed the 2SMG choice rule, the formulation defined through the adjustment process. This formulation of the 2SMG choice rule first appeared in Indian court rulings with the judgment *Tamannaben Ashokbhai Desai (2020)* (Gujarat High Court, 2020) where it became mandated for the State of Gujarat.

Another key mandate in *Saurav Yadav (2020)* is the enforcement of the axiom of *no justified envy,* in case a choice rule that differs from 2SMG is adopted by a public institution. While this axiom was always implicitly expected in the country, it was never explicitly formulated prior to this judgment, thus causing much of the crisis on joint implementation of VR and HR policies. Therefore, not only Sönmez and Yenmez (2019, 2022a) correctly anticipated the "proper" amendment of the SCI-AKG choice rule, but also the intended social justice axiom in the spirit of the country's affirmative action policies.

Perhaps due to the aftermath of the AKG-SCI choice rule's enforcement, the Supreme Court justices have merely endorsed the 2SMG choice rule and refrained from enforcing it. However, one misleading aspect of this "seemingly" more flexible guidance on selecting an allocation mechanism exists.

**Desiderata Mandated under Saurav Yadav (2020).** As we have emphasized, accurately identifying the root cause of the crisis as the failure of *no justified envy* under the rescinded SCI-AKG choice rule, the justices have mandated this important axiom under *Saurav Yadav (2020)*. On top of this axiom, three additional desiderata are also mandated with this landmark judgment.

We next formulate all four mandates as rigorous axioms.

A choice rule $\mathcal{C} = \left(\mathcal{C}^v\right)_{v \in V}$ is *non-wasteful* if, for every $J \subseteq I,\ v \in V$, and $j \in J$,

$$j \notin \widehat{\mathcal{C}}(J) \ \text{ and } \ |\mathcal{C}^v(J)| < q^v \quad \implies \quad j \notin I^v.$$

A position can remain idle at any category $v \in V$ only if none of the individuals who remain unassigned is eligible for a category-$v$ position. This mild efficiency axiom has been mandated in India since *Indra Sawhney (1992)*.

The following auxiliary concept simplifies the formulation of our next three axioms.

Given a category $v \in V$ and set of individuals $J \subseteq I^v$, let $n^v(J)$ denote the maximum number of HR-protected positions that can be *honored*; i.e., awarded to their intended beneficiaries. For the case of non-overlapping HR protections, for any category $v \in V$, the *HR-maximality function* $n^v : 2^{I^v} \to \mathbb{Z}_+$ is given as follows:

For any $J \subseteq I^v$,

$$n^v(J) = \sum_{t \in T} \min \Big\{ \big|\{i \in J \ : \ t \in \tau(i)\}\big|, \, q_t^v \Big\}.$$

A choice rule $\mathcal{C} = \left(\mathcal{C}^\nu\right)_{\nu \in V}$ *maximally accommodates HR protections* if, for each $J \subseteq I$, $v \in V,$ and $j \in \left(J \cap I^v\right) \setminus \widehat{\mathcal{C}}(J)$,

$$n^v\big(\mathcal{C}^v(J) \cup \{j\}\big) \not> n^v\big(\mathcal{C}^v(J)\big).$$

An individual cannot remain unassigned if they can increase the number of HR-protected positions honored at some category for which they are eligible. It became mandated in India with *Saurav Yadav (2020)* in this form. It fails under the rescinded SCI-AKG choice rule because VR-protected individuals had been deemed ineligible for HR protections within open positions under this rule.

A choice rule $\mathcal{C} = \left(\mathcal{C}^\nu\right)_{\nu \in V}$ satisfies *no justified envy* if, for each $J \subseteq I, \ v \in V,$ $i \in \mathcal{C}^v(J)$, and $j \in \left(J \cap I^v\right) \setminus \widehat{\mathcal{C}}(J)$,

$$\text{either} \quad \sigma(i) > \sigma(j) \quad \text{or} \quad n^v\big(\mathcal{C}^v(J)\big) > n^v\Big(\big(\mathcal{C}^v(J) \setminus \{i\}\big) \cup \{j\}\Big).$$

At any category $v \in V$, a lower merit-ranking individual $i \in I^v$ can receive a position at the expense of a higher merit-ranking individual $j \in I^v$ who remains unassigned only if replacing $i$ with $j$ decreases the number of HR-protected positions that are honored at category $v$. The mandate of this axiom is the main message of *Saurav Yadav (2020)*. In India, this axiom is widely referred to as the *principle of merit* when $v = o$, and as the *principle of inter se merit* when $v \in R$.

A choice rule $\mathcal{C} = \left(\mathcal{C}^\nu\right)_{\nu \in V}$ *complies with VR protections* if, for each set of individuals $J \subseteq I$ and reserve-eligible category $c \in R$, whenever $i \in \mathcal{C}^c(J)$ (and hence $i \notin \mathcal{C}^o(J)$) the following three conditions hold:

1. $|\mathcal{C}^o(J)| = q^o$,
2. for each $j \in \mathcal{C}^o(J)$,

$$\text{either} \quad \sigma(j) > \sigma(i) \quad \text{or} \quad n^o\big(\mathcal{C}^o(J)\big) > n^o\big((\mathcal{C}^o(J) \setminus \{j\}) \cup \{i\}\big), \text{ and}$$

3. $n^o\big(\mathcal{C}^o(J) \cup \{i\}\big) \not> n^o\big(\mathcal{C}^o(J)\big)$.

Here, the first two conditions formulate the idea of a vertical reservation à la Indra Sawhney (1992). The third condition is a new mandate in *Saurav Yadav (2020)*, and it additionally requires that a member of a reserve-eligible category who can increase the number of HR-protected positions honored in the open category shall not be awarded a VR-protected position.

Apart from enforcing the axiom of *no justified envy* and rescinding the SCI-AKG choice rule, *Saurav Yadav (2020)* also brings much-needed clarity to a subtle aspect of the implementation of VR protections in the presence of HR protections. When the concept of vertical reservations was introduced in *Indra Sawhney (1992)*, its defining characteristics were described as follows:

> *"It may well happen that some members belonging to, say Scheduled Castes get selected in the open competition field based on their own merit; they will not be counted against the quota reserved for Scheduled Castes; they will be treated as open competition candidates."*

However, no judgment of the Supreme Court before *Saurav Yadav (2020)* explicitly formulated what it means to get selected in the open competition based on merit when there are also HR protections. To a large extent, much of the disarray about the concurrent implementation of VR and HR policies boils down to this ambiguity.[35] This vagueness is now removed under *Saurav Yadav (2020)*, where an individual who gets selected in the open competition based on merit is legally defined as one who deserves an open category position based on merit with or without invoking the HR protections. The third condition in our last axiom is an implication of this clarification.

Collectively, the mandates in *Saurav Yadav (2020)* have a very sharp policy implication.

**Theorem 13 (Sönmez and Yenmez, 2022a)** *Suppose each individual has at most one trait. A choice rule is non-wasteful, maximally accommodates HR protections, satisfies no justified envy, and complies with VR protections if, and only if, it is the 2SMG choice rule* $\mathcal{C}^{2s}_{mg}$.

Therefore, while *Saurav Yadav (2020)* has not explicitly enforced and merely endorsed the 2SMG choice rule, it has indirectly enforced this choice rule through its other mandates.

## 5.4 Overlapping HR Protections

We next extend our analysis to the general version of the problem with overlapping HR protections. In India, VR-protected groups do not overlap with each other, although they do overlap with HR-protected groups. So far, we have assumed that HR-protected groups do not overlap with each other either (i.e., individuals have at most one trait). Court cases seem to abstract away from any complications due to overlapping HR-protected groups.

However, in many field applications, HR-protected groups overlap, e.g., HR pro-

[35] A key (but underutilized) role for a market designer is bringing formalism to analytical concepts developed by layman or experts in non-technical fields.

tections for “women” and “persons with disabilities”.

A key question arises: Does a member of multiple HR-protected groups count towards minimum guarantees for all these groups or only one of them upon admission? This is unlegislated and typically left at the discretion of the central planner.

We adopt the latter *one-to-one HR matching* convention. There are several reasons for this choice. Position numbers are typically announced for category-trait pairs in practice in India, which automatically embeds this convention into the outcome of the problem. Thus, it is way more widespread in the field. From a formal perspective, this convention also provides a “clean” generalization of Theorem 13 to the general version of the problem with overlapping HR protections, thus specifying a unique choice rule that abides by the mandates of *Saurav Yadav (2020)* for this case as well.

One natural question is, why not use the 2SMG choice rule (cf. Section 5.3) for this case as well? That is, first allocate open positions with the minimum guarantee choice rule in Step 1, and next use the same choice rule for each VR-protected category in Step 2. After all, 2SMG uniquely satisfies the mandates of *Saurav Yadav (2020)* for the simpler case with non-overlapping HR protections.

This idea is not farfetched; however, there is a technical problem. When HR protections are non-overlapping, no individual is eligible for multiple HR-protected positions, thus making Step 1 of the minimum guarantee rule (cf. Section 5.3) uniquely defined. However, this is not the case when HR-protected positions overlap. Depending on how we process the traits, we may choose different sets of individuals. For example, if we process traits with a given fixed sequence, some processing orders may lead to the needless rejection of high-merit individuals. This problem is analogous to the problem with sequential reserve rules introduced in Section 4 of Chapter 1, as illustrated in Example 20 of that chapter.

Thus, processing HR protections in a “mechanical” way using a fixed sequence of traits may lead to implausible outcomes, which may depend on the processing sequence of HR-protected groups. Higher merit-score individuals can be rejected at the expense of lower merit-score individuals without increasing the overall representation of HR-protected groups. HR-protected groups may end up being unnecessarily underrepresented.

The admission of an individual with multiple traits presents a “flexibility” in accommodating HR protections; one that is lost under the minimum guarantee choice rule.

This flexibility can be utilized to obtain a more meritorious outcome. To formulate a generalization of the minimum guarantee choice rule that achieves this objective, we first need to generalize the HR-maximality function $n^v$ (cf. Section 5.3) for each vertical category $v \in V$.

### 5.4.1 Generalized HR-Maximality Function and Meritorious Horizontal Choice Rule

Consider a situation with two HR-protected positions, one for women and one for persons with disabilities. If the only individuals with these traits are a disabled man and a disabled woman, it will not be possible to honor both HR-protected positions if the HR-protected position for persons with disabilities is awarded to the disabled woman. While she could as well receive the HR-protected position for women, the disabled man cannot. Thus, the assignment of HR-protected positions has non-trivial consequences for cases of overlapping HR protections.

Given a category $v \in V$ and set of individuals $J \subseteq I^v$, recall that $n^v(J)$ denotes the maximum number of HR-protected positions that can be "honored"; ie, awarded to target beneficiaries. This expression is key for all *Saurav Yadav (2020)* axioms. For the general case with overlapping HR protections, this number can be found through several polynomial time algorithms, such as Edmonds' Blossom Algorithm (Edmonds, 1965). It requires the *maximal matching* of individuals to traits.

Given a category $v \in V$ and a set of individuals $J \subsetneq I^v$, an individual $i \in I^v \setminus J$ *increases HR utilization of* $J$ if

$$n^v(J \cup \{i\}) = n^v(J) + 1.$$

We are ready to formulate a generalization of the minimum guarantee choice rule, introduced by Sönmez and Yenmez, 2022a, that utilizes the flexibility in accommodating the HR protections under the one-to-one HR matching convention.

**Meritorious Horizontal Choice Rule $\mathcal{C}^v_{Ⓜ}$.**

**Step 1.**

**Step 1.0.** Let $I_0 = \emptyset$.

**Step 1.k.** (k $>$ 0) Assuming such an individual exists, choose the highest merit-score individual in $J \setminus I_{k-1}$ who increases the HR utilization of $I_{k-1}$. Denote this individual by $i_k$ and let

$$I_k = I_{k-1} \cup \{i_k\}.$$

Continue with Step 1.k+1.

If no such individual exists, proceed to Step 2.

**Step 2.** For unfilled positions, choose unassigned individuals with the highest merit scores until either all positions are filled or all individuals are selected.

■

The following result justifies the naming of the meritorious horizontal choice rule $\mathcal{C}^v_{Ⓜ}$.

**Proposition 6 (Sönmez and Yenmez, 2022a)** *Given a category $v \in V$, let $\mathcal{C}^v$ be any single-category choice rule that maximally accommodates HR protections. Then, for every set of individuals $J \subseteq I^v$,*

1. $|\mathcal{C}^v(J)| \leq |\mathcal{C}^v_{Ⓜ}(J)|$*, and*
2. *for every $k \leq |\mathcal{C}^v(J)|$, if $i$ is the $k$-th highest merit-score individual in $\mathcal{C}^v_{Ⓜ}(J)$ and $j$ is the $k$-th highest merit-score individual in $\mathcal{C}^v(J)$, then*

$$i = j \quad \text{or} \quad \sigma(i) > \sigma(j).$$

The next result shows that the meritorious horizontal choice rule $\mathcal{C}^v_{Ⓜ}$ is the only plausible procedure to accommodate the HR protections under the one-to-one HR matching convention.

**Theorem 14 (Sönmez and Yenmez, 2022a)** *Given a category $v \in V$, a single-category choice rule is non-wasteful, maximally accommodates HR protections, and satisfies no justified envy, if, and only if, it is the meritorious horizontal choice rule $\mathcal{C}^v_{Ⓜ}$.*

### 5.4.2 2-Step Meritorious Horizontal Choice Rule

The following choice rule, as proposed by Sönmez and Yenmez, 2022a, is a generalization of the 2SMG choice rule for the case of the model with overlapping HR protections. Instead of the minimum guarantee choice rule, it employs the meritorious horizontal choice rule in each step.

**2-Step Meritorious Horizontal Choice Rule (2SMH) $\mathcal{C}^{2s}_{Ⓜ} = (\mathcal{C}^{2s,\nu}_{Ⓜ})_{\nu \in V}$.**

Given a set of individuals $J \subseteq I$,

$$\begin{aligned} &\textbf{Step 1.} \qquad \mathcal{C}^{2s,o}_{Ⓜ}(J) = \mathcal{C}^{o}_{Ⓜ}(J) \\ &\textbf{Step 2.} \qquad \mathcal{C}^{2s,c}_{Ⓜ}(J) = \mathcal{C}^{c}_{Ⓜ}\Big(\big(J \setminus \mathcal{C}^{o}_{Ⓜ}(J)\big) \cap I^c\Big) \quad \text{for any } c \in R. \end{aligned}$$

■

Our next result establishes that the 2SMH choice rule is the only mechanism that abides by the *Saurav Yadav (2020)* axioms for the general case of the problem with overlapping HR protections.

**Theorem 15 (Sönmez and Yenmez, 2022a)** *A choice rule is non-wasteful, maximally accommodates HR protections, satisfies no justified envy, and complies with VR protections if, and only if, it is the 2SMH choice rule $\mathcal{C}^{2s}_{Ⓜ}$.*

## 5.5 Indian Affirmative Action with Multiple Institutions

In many cases, the positions to be filled are not identical. For instance, there are multiple institutions, each of which requires adherence to affirmative action laws. Applicants often have strict preferences for these institutions.

This broader version of the problem is significant in India because it covers not only the allocation of some of the most prestigious public jobs (e.g., Indian Administrative Service positions) but also the assignment of public college seats, such as the admissions process for the Indian Institutes of Technology, world renowned engineering schools in India (Baswana et al., 2019)).

The results we give here extend the prescription of the 2SMH choice rule introduced in Section 5.4 to this setting. This subsection follows Sönmez and Yenmez (2022b), and almost all definitions and results are from this paper.

Suppose a set of students $I$, a set of vertical categories $V = R \cup \{o\}$, a set of traits $T$, a reserve category eligibility function $\rho$, and a trait function $\tau$ are given as before.

Let the set of schools be denoted as $S$. Each school $s \in S$ has $q_s$ identical positions. For each $s \in S$ and $c \in R$, let $q_s^c$ denote the number of VR-protected school seats at school $s$ for members of reserve-eligible category $c$. For each $s \in S$, the number of open category seats are given as $q_s^o = q_s - \sum_{c \in R} q_s^c$. For each $s \in S$, $v \in V$, and $t \in T$, let $q_{s,t}^v$ denote the minimum number of seats guaranteed at school $s$ for category-$v$ eligible students who possess trait $t$. Let

$$q = \Big(q_s, \big(q_s^\nu, (q_{s,t}^\nu)_{t \in T}\big)_{\nu \in V}\Big)_{s \in S}$$

refer to the capacity and reserve vector at all schools.

For each student $i \in I$, let $\succ_i$ denote a strict preference relation over $S \cup \{\emptyset\}$, where $\emptyset$ refers to remaining unmatched. Let $\succeq_i$ refer to its weak preference relation. We define $\mathcal{P}_i$ as the set of all strict preferences for student $i$. Let $\mathcal{P} = \bigtimes_{i \in I} \mathcal{P}_i$ represent the set of preference profiles.

Merit scores are specific to each school. For every school $s \in S$ and student $i \in I$, let $\sigma_s(i)$ denote the score of student $i$ at school $s$. Let $\sigma_s = \big(\sigma_s(i)\big)i \in I$ represent the vector of scores relevant for school $s \in S$. For any given school $s \in S$, we assume that no two different students share the same score in $\sigma_s$. The score profile is denoted as $\sigma = (\sigma_s)_{s \in S}$.

An *Indian AA environment with multiple institutions* is denoted by $[I, S, \mathcal{P}, V, T, q, \sigma, \rho, \tau]$. In this subsection, we fix such an environment. Therefore, a problem is denoted by a preference profile $\succ = (\succ_i)_{i \in I} \in \mathcal{P}$.

An outcome of a problem in this environment is a *matching* $\mu : I \to \big(S \times V\big) \cup \{\emptyset\}$ such that $|\mu^{-1}(s, v)| \leq q_s^v$ for each $s \in S$ and $v \in V$. Given a matching $\mu$ and a student $i \in I$, define

$$\mathbf{s}\big(\mu(i)\big) = \begin{cases} s & \text{if } \mu(i) = (s, v) \\ \emptyset & \text{if } \mu(i) = \emptyset \end{cases} .$$

Let $\mathcal{M}$ be the set of matchings in this environment.

Fix a problem $\succ \in \mathcal{P}$. We have the following axioms.

A matching $\mu$ is *individually rational* if for each student $i \in I$,

$$\mathbf{s}\big(\mu(i)\big) \succeq_i \emptyset.$$

Our next four axioms represent generalizations of *Saurav Yadav (2020)*'s axioms (cf. Section 5.3) to the most general version of the problem with overlapping HR protections and heterogeneous positions.

A matching $\mu$ satisfies *non-wastefulness* if for each $s \in S$, $v \in V$, and $i \in I$,

$$s \succ_i \mu(i) \text{ and } |\mu^{-1}(s,v)| < q_s^v \quad \Longrightarrow \quad i \notin I^v.$$

Since HR protections can be overlapping, for each school $s \in S$, we rely on the generalized HR maximality function $n^v(.)$ given in Subsection 5.4, originally formulated for a single institution.

A matching $\mu$ satisfies *maximal accommodation of HR protections* if for each $s \in S$, $v \in V$, and $i \in I^v$,

$$s \succ_i \mathbf{s}\big(\mu(i)\big) \quad \Longrightarrow \quad n_s^v\Big(\mu^{-1}(s,v) \cup \{i\}\Big) \not> n_s^v\Big(\mu^{-1}(s,v)\Big).$$

A matching $\mu$ satisfies *no justified envy* if for each $i \in I$, $s \in S$, $v \in V$, and $j \in I^v$,

$$\left.\begin{array}{c} \mu(i) = (s,v) \text{ and} \\ s \succ_j \mathbf{s}\big(\mu(j)\big) \end{array}\right\} \Longrightarrow \left\{\sigma_s(i) > \sigma_s(j) \text{ or } n_s^v\Big(\mu^{-1}(s,v)\Big) > n_s^v\Big(\big(\mu^{-1}(s,v) \setminus \{i\}\big) \cup \{j\}\Big).\right.$$

A matching $\mu$ *complies with VR protections* if for each $s \in S$, $c \in R$, and $i \in I^c$, whenever $\mu(i) = (s,c)$, the following three conditions hold:

1. $|\mu^{-1}(s,o)| = q_s^o$,
2. for each $j \in I$ with $\mu(j) = (s,o)$,

$$\sigma_s(i) > \sigma_s(j) \text{ or } n_s^o\Big(\mu^{-1}(s,o)\Big) > n_s^o\Big(\big(\mu^{-1}(s,o) \setminus \{i\}\big) \cup \{j\}\Big), \text{ and}$$

3. $n_s^o\Big(\mu^{-1}(s,o) \cup \{i\}\Big) \not> n_s^o\Big(\mu^{-1}(s,o)\Big)$.

A *mechanism* is a function $\varphi : \mathcal{P} \to \mathcal{M}$ that chooses a matching for each problem of the environment.

A mechanism $\varphi$ satisfies a property defined for matchings, if for each problem $\succ \in \mathcal{P}$, $\varphi[\succ]$ satisfies this property for problem $\succ$.

The following incentive-compatibility property, a gold standard in practical applications of matching theory, is defined for mechanisms only.

A mechanism $\varphi$ is *strategy-proof* if for each $\succ \in \mathcal{P}$, $i \in I$ and $\succ_i' \in \mathcal{P}_i$,

$$\varphi[\succ_i, \succ_{-i}](i) \succeq_i \varphi[\succ_i', \succ_{-i}](i).$$

We next introduce a mechanism that relies on the generalization of the celebrated individual-proposing deferred acceptance (DA) algorithm (Gale and Shapley, 1962), defined in Section 2 of Chapter 1 when schools have choice rules.

Each school $s \in S$ uses the 2SMH choice rule with respect to its score vector $\sigma_s$ and its capacity and reserve vector $\left(q_s, \left(q_s^v, (q_{s,t}^v)_{t \in T}\right)_{v \in V}\right)$. Denote this multi-category choice rule as $\mathcal{C}_s = \left(\mathcal{C}_s^v\right)_{v \in V}$.

The following mechanism, formulated and proposed for India in Sönmez and Yenmez, 2022b, comprises implementing the individual-proposing deferred acceptance (DA) algorithm under the choice rule profile $\left(\mathcal{C}_s\right)_{s \in S}$:

**Mechanism 2SMH+DA.**

**Step 0.** At the initiation, no offers are considered rejected by any school, and no student holds an offer from any school, and thus $H_s^{(0)} = \emptyset$ for each school $s \in S$.

**Step k.** ($k \geq 1$)

*Offer stage*: Each student $i$ who does not have a held offer from the previous step offers a match to their most preferred acceptable school, which has not rejected them in a previous step. If such a school does not exist, they remain unmatched at the end of the algorithm.

*Holding and Rejection stage*: Each school $s \in S$, which has been holding offers from the set of students $H_s^{(k-1)} \subseteq I$ and receives offers from a set of students $O_s^{(k)} \subseteq I$ at the Offer Stage of Step k, holds offers of students in

$$H_s^{(k)} = \widehat{\mathcal{C}}_s\left(H_s^{(k-1)} \cup O_s^{(k)}\right)$$

and rejects the rest of the students in $H_s^{(k-1)} \cup O_s^{(k)}$.

The algorithm terminates if there are no rejections. The outcome matching of the mechanism $\mu$ is determined as follows for each $s \in S$, $v \in V$, and $i \in \mathcal{C}_s^v\left(H_s^{(k)}\right)$,

$$\mu(i) = (s, v).$$

Otherwise, the algorithm continues with Step k+1.

■

We are ready to present the two main results of this subsection:

**Theorem 16 (Sönmez and Yenmez, 2022b)** *In an Indian AA environment with multiple institutions, of all mechanisms that satisfy individual rationality, non-wastefulness, maximal* *accommodation of HR protections, no justified envy, and compliance with VR protections, the mechanism 2SMH+DA Pareto dominates any other.*

**Theorem 17 (Sönmez and Yenmez, 2022b)** *In an Indian AA environment with multiple institutions, a mechanism satisfies individual rationality, non-wastefulness, maximal accommodation of HR protections, no justified envy, compliance with VR protections, and strategy-proofness if, and only if, it is the 2SMH+DA mechanism.*

## 5.6 Extensions: Some Other Applications of Reserve Systems

There are many other applications in the matching theory literature integrating the implementation of affirmative action policies into school choice or student placement. Dur et al. (2018) documented the unintended consequences of certain reserve category processing policies in the Boston school choice system, which later influenced policy in Boston. The consequences of incorporating a reserve system for affirmative action into the Chicago school choice system were analyzed by Dur, Pathak, and Sönmez (2020). Aygün and Bó (2021) analyzed incentive problems associated with implementing affirmative action in Brazilian college admissions.

Sönmez and Yenmez (2022b) showed that the Chilean K-12 school choice system is a special case of the model presented in Section 5.5 with only an open category. In this application, there are only minimum guarantee reserves.

Pathak, Rees-Jones, and Sönmez (2020) documented and analyzed the sequential reserve system used in the H1-B visa application process by foreign workers applying for work permits in the US. A recent policy change in the order in which the reserve categories are processed led to an overall change in admission outcomes without altering the reserve quotas.

Economist-designed reserve systems guided by medical ethics principles (cf. Pathak et al., 2023) have also played a crucial role in the equitable allocation of scarce medical resources during the COVID-19 pandemic, which we discuss in Section 8.3.

# 6 Entry-Level Physician Matching Markets and Unraveling of Their Appointment Dates

One of the first documented and studied markets through the lens of matching theory is the new physician markets in the US (Roth, 1984) and the regions of the UK (Roth, 1990). Some of these markets adopted centralized mechanisms that rely on variants of the celebrated deferred acceptance (DA) algorithm of Gale and Shapley (1962) in their matching process, either before or independently of Gale and Shapley's academic work. Later, these methods were adopted in some other professional entry-level labor markets (see Roth, 2008 for further markets where variants of DA algorithms were adopted in two-sided matching markets). These markets do not explicitly involve wages in the bargaining process, as these jobs mostly consist of interim educational positions, such as internships, residencies, or fellowships, and are different in nature from many other professional job markets.

Roth (1984) documented the challenges existing in the labor market for new graduates of US medical schools during the mid-20th century, followed by a widespread market failure that led to a system overhaul. The quality of prospective doctors was uncertain in advance of their graduation, and yet, due to the "unraveling" of appointment dates, such appointments were being finalized earlier and earlier in successive years. This unraveling prompted a shift from a decentralized market to establishing a centralized clearinghouse mechanism. Remarkably, the labor market started to use a stable mechanism equivalent of the college-optimal stable mechanism a decade before it was introduced and theoretically studied by the seminal work of Gale and Shapley (1962).

Alvin Roth and his collaborators went on to extend this unraveling vs. stability hypothesis to various entry-level labor markets for skilled professionals and other interesting decentralized or semi-centralized matching institutions, with sides in full (but also rarely, partial) possession of the required "property rights" of their arrangements (see Chapter 1 for a discussion and the importance of property rights in matching).

In this section, we survey these developments and the heuristic used by Roth and Peranson (1999) for the *couples problem* in the 1990s after a second market failure of the centralized market in the National Residency Matching Program (NRMP) in the US.

## 6.1 Background

During the 1940s, a decentralized market structure led to a chaotic environment where hospitals searched for promising medical students, often extending job offers as early as two years before a student's graduation. Upon a doctor's graduation, they needed to work as a medical resident for a period of time to be certified as doctors. This premature action resulted in the phenomenon of *unraveling of appointment dates*, disrupting the orderly market process. Students were compelled to commit to positions well in advance of finalizing their interests, also hindering the assessment of the adequacy of the students by the hospitals as their final grades were not available.

Recognizing the chaos stemming from these practices, pivotal changes were instituted. In 1944, the Association of American Medical Colleges (AAMC) implemented measures delaying the release of crucial student information, such as grades, until later stages of their medical education. Despite mitigating the immediate concerns, this approach merely addressed symptoms rather than fundamentally resolving the underlying issues. A sequence of subsequent adjustments between 1945 and 1951 attempted to tackle the persisting problems did not result in much improvement, particularly regarding waitlisted students' strategies in delaying acceptance to secure preferable positions.

These shortcomings paved the way for adopting a centralized matching mechanism (Roth, 1984). This transformation introduced in the 1951-52 market, initially voluntary, improved the situation despite allowing participants to seek alternative arrangements outside the system. Concerns, however, lingered about the fairness of the assignment algorithm, prompting deliberations and the evolution of the matching process. The Boston Pool Plan, which is equivalent to the college-proposing deferred acceptance algorithm, gained an advantage due to its ability to ensure equitable matches. Over time, despite the voluntary nature of participation, this algorithm became the de facto mechanism for the residency match.

Roth argues that the success of the NRMP mechanism is due to its ability to generate stable matchings besides ending chaotic decentralized hiring. This was pivotal in diminishing incentives for doctors and hospitals to seek external contractual arrangements, reinforcing the efficacy of the centralized matching system. We next explore more carefully the relationship between the stability of a mechanism and costly unraveling.

## 6.2 Unraveling of Transactions in Matching Markets, Centralization, and The Stability Hypothesis

The evolution towards a centralized clearinghouse in the medical resident market underscores the issues caused by the congestion and coordination failures of decentralized offers in market clearing. Additionally, Roth and Xing (1994) studied how coordination problems arising from dynamic market features and challenges associated with coordinating transactions influence market interactions.

The unraveling problem was, at least anecdotally, documented in various markets. These include sorority rushes at US colleges (Mongell and Roth, 1991), the clinical psychology PhD labor market (Roth and Xing, 1997), the federal law clerkship market (Avery et al., 2001; Haruvy, Roth, and Ünver, 2006; Posner et al., 2007), the market for gastroenterology fellows (Niederle and Roth, 2003, 2004), and even the arrangement of college-level American football post-season games (Frechette, Roth, and Ünver, 2007). The last paper and Niederle and Roth (2003) provide two of the few empirical pieces of evidence of the inefficiency of unraveling. Otherwise, much of the evidence on inefficiencies caused by unraveling is based on anecdotal evidence or historical accounts of the evolutionary market forces that led to reorganization.

From a theoretical side, Li and Rosen (1998) developed a two-period matching model with monetary transfers, focusing on unraveling due to initial productivity uncertainty. Niederle, Roth, and Ünver (2013) proposed a model in labor markets where firms have a fixed identity, exploring unraveling driven by insurance against potential imbalances between firms and productive workers, without transfers. Other unraveling models consider search frictions (Damiano, Li, and Suen, 2005) and strategic complementarities (Echenique and Pereyra, 2016) where early contracting arises due to search costs.

Roth (1990) posits that implementing a stable matching through a clearinghouse might reduce the likelihood of unraveling and widespread market failure. Kagel and Roth (2000) conducted stylized laboratory experiments illustrating that introducing a stable mechanism can prevent costly unraveling. In contrast, unraveling was widespread under some of the unstable mechanisms abandoned in the field.

Yet, some unstable mechanisms survived in the regional medical matching markets of the UK, possibly due to other idiosyncratic features of the market environments and mechanisms (Roth, 1991; Ünver, 2001, 2005).

The stability hypothesis also provides only a rough intuition against unraveling. Sönmez (1999a) shows that stable matching mechanisms do not entirely prevent agents from pre-arranging matches. The basic intuition is that, while the stability of a mechanism assures that its outcome is immune to re-contracting, it does not ensure that better outcomes cannot be obtained if institutions rely on the mechanism to fill only some of their positions.

Formally, a matching mechanism $\varphi$ is *weakly manipulable via pre-arranged matches* if there is a college admissions market with responsive college preferences $[C, I, q, \succ]$ and some college-student pair $(c, i)$ such that

$$c \ \succeq_i \ \varphi[C, I, q, \succ](i)$$

and

$$\varphi\big[C, I \setminus \{i\}, (q_c - 1, q_{-c}), \succ_{-i}^{-i}\big](c) \cup \{i\} \ \succeq_c \ \varphi[C, I, q, \succ](c)$$

with at least one of the weak preference relations holding strictly.[36] Here, preference profile $\succ_{-i}^{-i}$ is obtained from the preference profile $\succ_{-i}$ of all agents except student $i$, and for the restricted market by removing student $i$ from their preferences.

In this arrangement between student $i$ and college $c$,

1. student $i$ weakly prefers college $c$ over its outcome from the mechanism, and
2. the college weakly prefers student $i$ together with its outcome from the cen-

[36] A college admissions environment with responsive college preferences was defined in Section 2 of Chapter 1 in this handbook. Here, we allow the possibility of agent sets and college capacities for colleges to change instead of fixing them as in that section. Thus, a market is denoted by the whole list of students, colleges, their capacities and the preference profile, not just by the preference profile.

tralized market (which does not include student $i$) over its outcome from the mechanism (which includes student $i$) without the outside arrangement with student $i$.

A mechanism $\varphi$ is *strongly manipulable via pre-arranged matches* if both preference relations are strict in the above definition.

The following result implies that the link between the stability of a mechanism and unraveling is not as strong as one might think.

**Theorem 18 (Sönmez, 1999a)** *In a college admissions environment with responsive college preferences, there is no matching mechanism that is stable and immune to strong manipulations via pre-arranged matches.*

## 6.3 Calls for Doctor-Proposing Deferred Acceptance Mechanism

As discussed in Section 2 of Chapter 1 of this handbook, truth-telling is neither a dominant strategy for students nor for the colleges under the college-optimal stable mechanism. Moreover, of all stable matchings, while it generates the best one for colleges, it yields the worst one for students. These theoretical results became widely known in the mid-1990s in the medical community. The National Residency Matching Program (NRMP) came under increased scrutiny from students and their advisors, who believed that it did not operate in the best interest of students and was open to the possibility of different kinds of strategic behavior (Roth and Rothblum, 1999). There were calls to change to the doctor-proposing version of deferred acceptance, which generates the doctor-optimal stable matching. Another significant reason for these calls was that truth-telling is a dominant strategy for doctors under the proposed algorithm, whereas it is not under the college-optimal stable one. Doctors would not bear the burden of strategizing under such a reform.

## 6.4 Addressing the "Couples Problem" in Medical Matching

As early as 1970s, another significant issue in medical matching pertained to the increasing number of married couples among medical school graduates. The initial hospital-optimal stable matching mechanism did not accommodate couples seeking to coordinate their job placements. This often resulted in scenarios where each doctor in the couple could be assigned to hospitals in different parts of the country, and hence, they had limited choices to apply if they wanted to avoid this outcome.

The rise in married doctor couples led to a decline in medical student participation in the NRMP during the 1970s. Consequently, the NRMP introduced a new procedure to address couple preferences. In this procedure, couples, certified by their dean, could register as a unit. One member was designated the "leading member," and both submitted separate rank-order lists of positions. The leading member underwent the match process individually, while the other member's list was adjusted to align with the "community," a group of hospitals (for example, in the same geographic area) where the leading member was matched. Initially, the NRMP determined communities, but later proposals allowed couples to specify their preferred community.

In this design, the challenge arises from couples consuming pairs of job positions, whereas the procedure only solicits preferences over individual positions. Accommodating preferences over pairs of hospitals is a challenge under stability. Even with just one couple, a stable matching may not exist. We explore these issues in more detail below.

**Stability with Couples.** Consider the college admissions market with responsive college preferences introduced in Section 2 of Chapter 1 of this handbook. Here, we modify this model to include couples.

Let $H$ be a set of *hospitals* and $D$ be a set of *doctors*. Let $S \subseteq D$ denote the set of *single doctors*. The remaining set of doctors $D \setminus S$ is even in size and made up of *couples*. Each hospital $h$ has a capacity $q_h$ and a strict responsive preference relation $\succ_h$ over the set of doctors $D$. Each single doctor $d \in S$ has a strict preference relation $\succ_d$ over hospitals $H$ and the option $\emptyset$ to remain unmatched. Each couple consists of a pair of doctors $d_1, d_2 \in D \setminus S$ and is denoted as $c = (d_1, d_2)$. Let $C$ be the set of couples. Each couple has a strict preference relation $\succ_c$ over $\big(H \cup \{\emptyset\}\big)^2$. Given strict preferences $\succ_a$ of an agent in $a \in H \cup S \cup C$, let $\succeq_a$ denote the induced weak preference relation.

In this section, we refer to the 4-tuple $[S, C, H, \succ]$ as a *many-to-one matching market with couples and responsive hospital preferences* or simply a *couples problem*.

A match for a couple $(d_1, d_2)$ is denoted as a pair $(x_1, x_2)$ so that $x_1$ is the assignment of $d_1$ and $x_2$ is the assignment of $d_2$.

A matching is appropriately defined for this market so that no hospital $h \in H$ is matched with more than $q_h$ doctors, each single doctor is matched with at most one hospital, and each couple is matched with at most two hospitals.

We consider the following *blocks* to a matching.

A single doctor $d \in S$ *blocks* a matching $\mu$ if

$$\emptyset \ \succ_d \ \mu(d).$$

A couple $c \in C$ *block* a matching $\mu$ if

$$\big(\mu(c)_1, \emptyset\big) \ \succ_c \ \mu(c), \qquad \text{or} \qquad \big(\emptyset, \mu(c)_2\big) \ \succ_c \ \mu(c), \qquad \text{or} \qquad (\emptyset, \emptyset) \ \succ_c \ \mu(c).$$

A hospital $h \in H$ *blocks* a matching $\mu$ if

$$\emptyset \succ_h d \quad \text{for some doctor } d \in \mu(c).$$

A single doctor $d \in S$ and a hospital $h \in H$ *block* a matching $\mu$ if

1. $h \succ_d \mu(d)$, and
2. $d \succ_h d'$ for some $d' \in \mu(h)$ or $\left[d \succ_h \emptyset \text{ and } |\mu(h)| < q_h\right]$.

A couple $c = (d_1, d_2) \in C$ and a hospital $h \in H$ *block* a matching $\mu$ if there exists a doctor of the couple $d_k \in \{d_1, d_2\}$ and their partner $d_\ell \in \{d_1, d_2\} \setminus \{d_k\}$ such that

1. $(x_1, x_2) \succ_c \mu(c)$ where $x_k = h$ and $x_\ell = \mu(c)_\ell$, and
2. $d_k \succ_h d$ for some $d \in \mu(h)$ or $\left[d_k \succ_h \emptyset \text{ and } |\mu(h)| < q_h\right]$.

Given a pair of (not necessarily distinct) hospitals $h_1, h_2 \in H$, a couple $c = (d_1, d_2) \in C$ and the pair of hospitals $(h_1, h_2)$ *block* a matching $\mu$ if

1. $(h_1, h_2) \succ_c \mu(c)$,

2a. $d_1 \succ_{h_1} d$ for some $d \in \mu(h_1)$ or $\left[d_1 \succ_{h_1} \emptyset \text{ and } |\mu(h_1)| < q_{h_1}\right]$, and,

2b. $d_2 \succ_{h_2} d$ for some $d \in \mu(h_2)$ or $\left[d_2 \succ_{h_2} \emptyset \text{ and } |\mu(h_2)| < q_{h_2}\right]$.

We refer to the set of agents that form a block as a *blocking coalition*.

A matching is *stable* if it is not blocked.

Roth (1984) shows that, in general, a stable matching may not exist in a many-to-one matching market with couples. The following example, based on an example in Kojima, Pathak, and Roth (2013), illustrates this possibility.

**Example 8** *Suppose there are two hospitals $h_1, h_2$ with one position each. There is one single doctor $s$ and one couple $c = (w, m)$ with the following preference profile:*

$$\succ_s: \ h_2 \ h_1 \ \emptyset$$
$$\succ_{(w,m)}: (h_1, h_2) \ (h_1, \emptyset) \ (\emptyset, \emptyset)$$
$$\succ_{h_1}: \{s\} \ \{w\} \ \emptyset$$
$$\succ_{h_2}: \{m\} \ \{s\} \ \emptyset$$

*Consider a matching $\mu$ which is not blocked by a doctor, by a couple or by a hospital. We will show that there is still a coalition that blocks $\mu$.*

*If couple's both doctors are matched, $\mu(w, m) = (h_1, h_2)$, then the single doctor $s$ is unmatched, and together with hospital $h_1$ they block $\mu$.*

*If only the couple's first doctor $w$ is matched (i.e., $\mu(w, m) = (h_1, \emptyset)$), then hospital $h_2$ is either unmatched or matched with the single doctor $s$. In this case, hospital $h_2$ and couple $c = (w, m)$ together block $\mu$.*

*If both doctors of the couple are unmatched (i.e., $\mu(w, m) = (\emptyset, \emptyset)$), then*

- *if $\mu(s) = h_2$, then couple $c = (w, m)$ together with hospitals $h_1$ and $h_2$ block $\mu$, and*
- *if $\mu(s) = h_1$ or $\mu(s) = \emptyset$, then single doctor $s$ and hospital $h_2$ block $\mu$.*

*Hence, there exists no stable matching.*

**The New NRMP Mechanism.** As this negative result rules out the possibility of constructing a mechanism that always finds a stable matching, Roth and Peranson (1999) explored heuristics to assign single doctors and couples to hospitals.[37] These heuristics are based on an alternative algorithm to find a stable matching in an opposite-sex marriage market by Roth and Vande Vate (1990) starting from an arbitrary unstable matching and matching agents in one blocking coalitions with each other in each step. This version of the heuristic adopted resembles the sequential adaptation of the deferred acceptance algorithm by McVitie and Wilson (1971). [38]

Given a couples problem $[S, C, H, \succ]$, we describe this heuristic based on the conceptual design in the online and print versions and the algorithmic schematic in the print version of Roth and Peranson (1999).[39] We refer to each single doctor and each couple as an *applicant* in this heuristic.

**Sequential NRMP Heuristic.**

We construct finite sequences of agent sets $A^0 \subsetneq A^1 \subsetneq A^2 \ldots$ and tentative matchings $\mu^0, \mu^1, \mu^2, \ldots$ such that for each $k \geq 0$, there is no blocking coalition $K \subseteq A^k$ of $\mu^k$.

**Step 0.** Set $A^0 = H$ and let $\mu^0 = \emptyset$ be the matching that leaves all agents unmatched. There is no coalition $K \subseteq A^0$ that blocks $\mu^0$.

**Step k.** (k > 0) Take an agent $i \in (S \cup C) \setminus A^{k-1}$ and define

$$A^k = A^{k-1} \cup \{i\}.$$

We construct tentative matching $\mu^k$ through a recursive stage using $\mu^{k-1}$. In

[37]In NRMP, there are other complications sometimes requiring a doctor to be matched with multiple programs, one for first-year residency and one for later years, and other complications in some residency programs' assignment requirements. For brevity, we omit these complications here and focus on complications arising from the existence of couples.

[38]Later, such algorithms were found by Klaus and Klijn (2007) in the couples problem when couple preferences satisfy a weak responsiveness property. A couple's problem is a variant of a two-sided many-to-many matching market. For many-to-many matching markets, when agents on one side (for example, couples)have substitutable preferences (for example, couples) and agents on the other side have responsive preferences (for example, hospitals), Kojima and Ünver (2008) introduced an algorithm that reaches to a pairwise stable matching starting from an arbitrary unstable matching. Pairwise stability is weaker than stability in the couples problem, as a matching is pairwise stable if it is immune to blocking by coalitions including at most two agents.

[39]The schematic did not exist in the online version reached from the American Economic Review website on 02/28/2024.

this stage, we keep track of a set of applicants $\hat{I} \subseteq S \cup C$ and a set of hospitals $\hat{H} \subseteq H$ that we dynamically update by adding and removing agents. We refer to $\hat{I}$ as the *applicant stack* and $\hat{H}$ as the *hospital stack*. These are ordered sets of agents: When an agent is added to either set, it is added to the *top*. When an agent is removed, it is removed from the top as well.
We also keep track of an interim tentative matching $\hat{\mu}$ that gets also updated so that all applicants in $\hat{I}$ are unmatched in $\hat{\mu}$. We initialize them as $\hat{I} = \{i\}$, and $\hat{H} = \emptyset$, and $\hat{\mu} = \mu^{k-1}$.

**A. Internal Stabilization Stage**:

A new tentative matching $\nu$ is formed using $\hat{\mu}$ as follows:
Remove an applicant $j$ from applicant stack $\hat{I}$. Applicant $j$ is either a single doctor or a couple. If $j$ is not in a blocking coalition of $\hat{\mu}$, then set $\nu = \hat{\mu}$ and go to the Internal Stability Check below. Otherwise, construct $\nu$ as follows:

A.1. We determine $\nu(j)$ as follows.
- If $j$ is a single doctor: $j$ is matched with the best hospital with respect to $\succ_j$ that they are blocking $\hat{\mu}$ with.
- If $j$ is a couple: Among all the blocking coalitions of $\hat{\mu}$ that $j$ is a member of, they are matched with the best option with respect to $\succ_j$, which can be two hospitals, one for each spouse, or one hospital that only one spouse is matched with, or one hospital that both spouses are matched with.

A.2 For each hospital $h \in H$: Let $D^+(h)$ be the (possibly empty) set of doctors matched with $h$ in Substep A.1. Let

$$D(h) = \max_{\succ_j} \Big\{ E \subseteq \hat{\mu}(h) \cup D^+(h) \ : \ \big|E\big| \leq q_h \Big\}.$$

Doctors in $D(h) \setminus \nu(h)$ are *rejected* by hospital $h$.

A.3. For each rejected doctor $d$ by a hospital in Substep A.2:
- If $d$ is a single doctor: Add $d$ to the applicant stack $\hat{I}$ and set $\nu(d) = \emptyset$.
- If $d$ is a spouse in a couple $c$: Add $c$ to the applicant stack $\hat{I}$. If the other spouse of $c$ is matched with some hospital $h'$ in $\hat{\mu}$, then *withdraw* the spouse's assignment from hospital $h'$. Set $\nu(c) = (\emptyset, \emptyset)$.

A.4. For each hospital $h \in H$: Let $D^-(h)$ be the set of the doctors who withdrew their assignments from $h$ in Substep A.3. If $D^-(h) \neq \emptyset$, add $h$ to the hospital stack $\hat{H}$. Set

$$\nu(h) = D(h) \setminus D^{-}(h).$$

A.5 The assignment of each remaining applicant is set to be the same in $\nu$ as their assignment in $\hat{\mu}$.

Continue with the following check:

**B. Internal Stability Check:**

B.1. If $\hat{I} \neq \emptyset$: Return to the Internal Stabilization Stage using tentative matching $\hat{\mu} := \nu$.

B.2. If $\hat{I} = \emptyset$ and $\hat{H} \neq \emptyset$: Remove a hospital $h$ from the hospital stack $\hat{H}$.

B.2.1. Each single doctor $i \in A^k$ that is preferred by $h$ to a doctor in $\nu(h)$ if $|\nu(h)| = q_h$ and is acceptable if $|\nu(h)| < q_h$ is added to the applicant stack $\hat{I}$.

B.2.2. Each couple $i \in A^k$ that has a spouse preferred by $h$ to a doctor in $\nu(h)$ if $|\nu(h)| = q_h$ and has an acceptable spouse if $|\nu(h)| < q_h$ is added to the applicant stack $\hat{I}$.

B.2.3. The tentative matching $\nu$ is updated such that each applicant added to $\hat{I}$ in Substep B.2.1 or Substep B.2.2 is left unmatched, and their matches from their assigned hospitals — if there are any — are *withdrawn* and otherwise the matches of other applicants and hospitals remain unchanged.

B.2.3. Any hospital with a withdrawn match in Substep B.2.3 is added to the hospital stack $\hat{H}$.

B.2.4. Return to Internal Stability Check.

B.3. If $\hat{I} = \emptyset$ and $\hat{H} = \emptyset$:

- If there is a blocking coalition $K \subseteq A^k$ of $\nu$: Place all the hospitals that belong to a blocking coalition to the hospital stack $\hat{H}$ and return to Substep B.2.
- If there is no blocking coalition $K \subseteq A^k$ of $\nu$: Set $\mu^k = \nu$. If $(S \cup C) \setminus A^k = \emptyset$, then terminate the heuristic, and $\mu^k$ is the outcome; otherwise, continue with Step k+1.

■

The following is a direct extension of the main result in Roth and Vande Vate (1990) that inspired the sequential NRMP heuristic.

**Proposition 7** *When there are no couples, the sequential NRMP heuristic always terminates and finds the doctor-optimal stable matching.*

Roth and Peranson (1999) ran simulations on historical data and numerical experiments to inspect the consequences of the order in which we include applicants to the set $A^k$ in each step k and other sequencing decisions. Finally, they used the order in which all single doctors are first processed, and then couples are processed, and a certain stopping rule is added to Substep A.3. in case the heuristic does not terminate for designing the *new NRMP mechanism*. In this stopping rule, if an applicant is rejected by the same hospital in Subsetp A.3 (or hospitals if they are a couple) twice in running of the algorithm, then the algorithm is stopped after Step A.4. returning $\nu$ as the outcome matching. They reported that they did not observe too many differences in changing these orders. They also verified on three years of historical data from NRMP markets that the algorithms constructed using this heuristic always converged to a stable matching.

**Limitations of the Design.** The Sequential NRMP heuristic might cycle for general couple preferences when there are couples. It requires the aforementioned stopping condition as instituted by Roth and Peranson (1999) to prevent it from going into an infinite loop. That is why we refer to this procedure as a *heuristic* rather than an algorithm.

Klaus, Klijn, and Massó (2007) demonstrated that the Roth and Peranson (1999) design may fail to find a stable matching even when couples have responsive preferences. In this case, the applicant-proposing deferred acceptance algorithm finds a stable matching. Moreover, they showed that it can be manipulated by couples acting single. Another issue is that the Roth and Peranson design is no longer strategy-proof, even for single doctors. However, truth-telling becomes an approximate Bayesian Nash equilibrium under certain regularity conditions in large markets (Kojima, Pathak, and Roth, 2013).

Further investigations by Kojima, Pathak, and Roth (2013) offer conditions under which a stable matching exists in large couples problems, provided couples do not constitute a significant portion of the market. They showed that a truncated version of the sequential NRMP heuristic in which single doctors are processed first almost surely converges to a stable matching as the problem size grows. Additional insights were provided by Ashlagi, Braverman, and Hassidim (2014), who introduced a new sequential heuristic for the couples problem. This heuristic almost surely converges to a stable matching asymptotically even if the number of couples grew at a closer rate to the singles but not faster.[40]

[40] Although the new NRMP mechanism is used in several entry-level labor markets in North America, the Israeli Medical Intern Market was also recently redesigned by design economists and computer scientists to address the couples problem and a *parents* problem (Bronfman et al., 2018). This market mostly resembles

**Another Solution to the Couples Problem.** An interesting solution to the couples problem was proposed by Nguyen and Vohra (2018). However, this requires possibly changing the capacities of hospitals, albeit relatively modestly, and can only be viable when such adjustments are possible, e.g., in places where hospitals are jointly managed or can coordinate.

**Theorem 19 (Nguyen and Vohra, 2018)** *Consider a many-to-one matching market with couples and responsive hospital preferences. It is possible to construct a stable matching for an alternative market obtained by changing the capacity of each hospital at most by 2 positions and the overall capacity of hospitals by at most 4.*

## 6.5 Other Developments in the NRMP Matching Market

The NRMP matching market has also gone through legal problems involving an anti-trust case filed against the NRMP, which accused it of keeping wages artificially low for new physicians through the centralized mechanism. Although the case was eventually dismissed, Bulow and Levin (2006) showed in a model how such wage compression could happen in equilibrium in an environment that does not involve wage offers in its centralized clearinghouse as in the NRMP mechanism.

# 7 Course Allocation at Universities

In many universities, the registration procedure for courses operates as a dynamic queue mechanism based on priority tiers: students in the highest priority tier are allocated time windows to select their schedules sequentially, followed by subsequent priority tiers (for example, our institution, Boston College, employs a similar system). The exact priority order among students in the same priority tier is typically determined by lottery. Once a section of an offered course reaches its capacity, it becomes unavailable for further selection. This environment mirrors an economy with a mixture of common-ownership rights and priority-based entitlements. It involves multi-unit demand, as each student typically registers for multiple courses. The unit demand versions of such environments were explored in Chapter 1.

Some graduate business schools implement more sophisticated procedures for course registration, with three notable examples coming to mind. Two of them, designed by system operators, have been studied in the market design literature: the University of Michigan Ross Business School (UMBS) employs a *pseudo auction* mechanism (Sönmez and Ünver, 2010; Krishna and Ünver, 2008), and the Harvard Business School (HBS) utilizes a *draft mechanism* (Budish and Cantillon, 2012).[41] The third ex-

a common-ownership economy, and traditionally, a random priority mechanism was used. A new design similar to random priority, accommodating parents and couples, was designed by this team of authors.

[41] See Caspari (2020) for another investigation of mechanisms in the class of draft mechanisms.

ample, a recently developed economist-designed system (Budish, 2011; Budish et al., 2017; Budish and Kessler, 2022), has been adopted at the University of Pennsylvania Wharton School of Business. This design extends the competitive equilibrium from equal income (CEEI) concept proposed by Hylland and Zeckhauser (1979), explored in Chapter 1, Section 3, to multi-unit demand environments.

In this section, we will provide a brief overview of the first two mechanisms and their properties, followed by an exploration of the approximate-CEEI mechanism of Budish (2011), which forms the foundation of the new Wharton mechanism.

## 7.1 The Model and Earlier Mechanisms Adopted in the Field

Each *student* $i \in I$ can register for up to $m$ courses, and each *course* $c \in C$ can accommodate a maximum of $q_c$ students. Let $q = (q_c)_{c \in C}$ denote the *capacity* vector. Students have strict preferences, denoted by $\succ_i$, over *schedules* of courses—where a schedule consists of a group of up to $m$ courses. For each student $i \in I$, let $\mathcal{P}_i$ denote the set of strict preferences.

The list $[I, m, C, q, \succ]$ represents a *course allocation problem*. Let $[I, m, C, q, \mathcal{P}]$ denote a *course allocation environment*, where $\mathcal{P} = \bigtimes_{i \in I} \mathcal{P}_i$ is the set of all preference profiles. A *matching* is a function $\mu : I \to 2^C$ that assigns a schedule to each student such that no course $c \in C$ is assigned to more than $q_c$ students, and no student is assigned more than $m$ courses.

We can directly extend the definitions of properties of matchings, mechanisms, and lottery mechanisms, which we introduced in the house allocation model covered in Section 3 of Chapter 1, to this model with multi-unit demand. We skip them for brevity. We refer the reader to that section for their formal definitions in the unit-demand setting.

### 7.1.1 Random Priority Mechanism

We define a *priority* mechanism extending its definition in the unit-demand setting as follows. Let

$$\pi = i_1 - i_2 - \ldots - i_{|I|}$$

be a priority order over students.

**Priority mechanism induced by $\pi$ for course allocation.**

**Step 0.** Let $C^0 = C$.

**Step k.** (k $\geq$ 1) Suppose agent with the $k$'th highest priority under $\pi$, $i_k$, is assigned their best schedule among the courses in $C^{k-1}$. Let $C^k \subseteq C^{k-1}$ be the set of courses that have not filled their capacity.

We terminate after Step $|I|$. ■

This mechanism is Pareto efficient, individually rational, and strategy-proof.

Thus, a fair and strategy-proof way of assigning courses is a randomized version of a priority mechanism, which is the direct version of the random queue mechanism already used for course allocation in many schools when there is a single priority tier. Formally, the *random priority* (*RP*) mechanism is the lottery mechanism that is obtained by randomly determining a priority order over agents and implementing the induced priority mechanism.

We have the following result, generalizing a similar result for RP in the unit-demand environment.

**Proposition 8** *In any course allocation environment, the RP mechanism satisfies ex-post efficiency, individual rationality, anonymity, and strategy-proofness.*

We next consider two other mechanisms used at business schools in the US to allocate schedules to students.

#### 7.1.2 UMBS Pseudo Auction

Sönmez and Ünver (2010) reports that UMBS used a course allocation mechanism based on students bidding over courses using token money. This is an indirect mechanism. Each student is endowed with a budget of $B$ tokens. UMBS asks students to bid for their favorite courses, meaning students are asked to distribute $B$ tokens among courses. Let $b_i = (b_{i,c})_{c \in C}$ denote the *bid vector* of student $i$, where $\sum_{c \in C} b_{i,c} \leq B$ and $b_{i,c} \geq 0$ for each $c \in C$.

UMBS course allocation mechanism, referred as a *pseudo auction*, operates as follows:

**The UMBS Pseudo Auction.**

Each student $i$ submits a bid vector $b_i$.

Suppose $b_{i_1,c_1} > b_{i_2,c_2} > \ldots > b_{i_n,c_n} > 0$ is the indexed sequence of ordered bids after a random tie-breaker is applied to break ties among them if there are any.

Starting with the highest bid, courses are assigned with the following iterative process:

**Step 0.** Initially, no student has received any courses, and no course is assigned to any students.

**Step k.** ($n \geq k \geq 1$) If student $i_k$ has less than $m$ courses assigned so far and course $c_k$ is assigned to less than $q_{c_k}$ students in previous steps, then assign student $i_k$ course $c_k$.

■

An issue with the UMBS pseudo auction is that the mechanism implicitly interprets that a student $i$ prefers a course $c$ to a course $d$ if they bid higher for $c$ than for $d$, since it assigns course $c$ before course $d$ to the student even though their bids can clear both courses' cutoffs. However, student $i$ may prefer course $d$ to course $c$ and yet still choose to strategically bid higher for $c$ than for $d$, simply because $c$ is a more popular course. As a result, students may not get into a course $d$ they prefer to one of their assigned courses, even though their bid is higher than the lowest bid honored for $d$. This, in turn, may lead to an efficiency loss, one that can be avoided by the following mechanism proposed and analyzed by Sönmez and Ünver (2010):

**Pareto-dominant Cutoff Equilibrium Mechanism.**

Students bid for courses as in the UMBS pseudo auction, but at the same time they also submit their preferences over schedules. Students are prioritized for each course based on their submitted bids, with higher bids awarding higher priority. Then, the student-proposing deferred acceptance algorithm (DA) is used to find a cutoff equilibrium outcome for the submitted bid profile, extending the same concept defined in the unit-demand environment for student placement with school priorities induced by the bid profile (see Chapter 1, Section 4).

**Theorem 20 (Sönmez and Ünver, 2010)** *Consider a course allocation problem with responsive student preferences over schedules. Suppose* $b = (b_i)_{i \in I}$ *is a strict bid profile of students.*

1. *The student-proposing DA algorithm finds the student-optimal cutoff equilibrium matching at* $b$*. Let* $\mu$ *be this outcome.*
2. *If the same bid profile* $b$ *were used, the UMBS pseudo auction outcome might not be a cutoff equilibrium matching at* $b$*. Moreover, it cannot Pareto dominate* $\mu$*.*

On the other hand, for the same bid profile, the outcome of DA can Pareto dominate the outcome of the UMBS mechanism.

The Pareto-dominant cutoff equilibrium mechanism was evaluated against the UMBS mechanism through a combination of laboratory experiments, empirical analysis of field data, and a supplementary survey of UMBS students conducted by Krishna and Ünver (2008). The study demonstrated that the Pareto-dominant cutoff equilibrium mechanism significantly outperformed the UMBS pseudo auction in terms of efficiency.

### 7.1.3 HBS Draft Mechanism

Although the RP mechanism is anonymous and thus satisfies equal treatment of equals, it can be highly unfair ex post. A student who is drawn last in the priority order may be left with fewer choices, possibly missing out on many of their favorite courses, especially the most popular ones. This can lead to a significant level of unfairness.

A practical compromise is documented by Budish and Cantillon (2012). HBS employs a mechanism inspired by RP called the *draft mechanism*, akin to drafts used for selecting junior players for sports teams employed in professional sports leagues in the US.

Unlike the UMBS pseudo auction, this mechanism is direct, in which students only submit preferences over courses. It implicitly assumes that the schedule preferences of students are responsive to individual preferences over courses.

**The HBS Draft Mechanism.**

Randomly draw a priority order $\pi$ over students with the uniform distribution. Define

$$\pi = i_1 - i_2 - \ldots - i_{|I|}.$$

Let $\pi^r$ represent the reverse order of $\pi$, i.e.,

$$\pi^r = i_{|I|} - i_{|I|-1} - \ldots - i_1.$$

The allocation proceeds in $m$ rounds such that

- in each odd round ($k = 1, 3, 5, \ldots$), $\pi$ is activated, and
- in each even round ($k = 2, 4, 5, \ldots$), the reverse priority order $\pi^r$ is activated.

**Round k.** (k $\leq$ m) Implement the *unit-demand* version of the priority algorithm induced by the active priority order:

**Step k.j** (j $\leq$ $|I|$) The $j$'th highest priority student of the active priority order is assigned their favorite course that has an available seat and was not assigned to them in a previous round.

■

This mechanism is no longer strategy-proof, unlike the random priority mechanism.

Empirical evidence by Budish and Cantillon (2012) based on surveys and field data reveals extensive manipulation by students using the HBS draft mechanism, resulting in efficiency loss. Additionally, suboptimal manipulation strategies lead to further efficiency losses. Interestingly, they show that the RP mechanism, in which students behave more truthfully, wouldn't have fared better. The HBS draft mechanism, on average, generates higher welfare for students even under sub-optimally manipulated preferences. To illustrate this point, they provide an example, similar to the welfare comparison between the probabilistic serial mechanism and RP mechanism for unit demand by Bogomolnaia and Moulin (2001) explained in Section 3 of Chapter 1.

**Example 9** *Let $I = \{1, 2, 3, 4\}$ be the set of students. Consider four courses, $C = \{a, b, c, d\}$, each with a capacity of 2 seats. Students require $m = 2$ courses each, with preferences as follows:*

*For each $i \in \{1,2\}$, $\succ_i$: $a\ b\ c\ d$,*
*For each $j \in \{3,4\}$, $\succ_j$: $b\ a\ d\ c$.*

*The RP mechanism yields the following random assignment for both student types:*

| *Students* | $\{a,b\}$ | $\{c,d\}$ |
|---|---|---|
| 1,2 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| 3,4 | $\frac{1}{2}$ | $\frac{1}{2}$ |

*Under the HBS draft mechanism, the assignment is as follows:*

| *Students* | $\{a,c\}$ | $\{b,d\}$ |
|---|---|---|
| 1,2 | 1 | 0 |
| 3,4 | 0 | 1 |

*The HBS draft outcome stochastically dominates the RP outcome for each student.*

## 7.2 Approximate CEEI as a Course Allocation Mechanism

Inspired by the concept of competitive equilibrium from equal incomes (CEEI) in house allocation by Hylland and Zeckhauser (1979) (cf. Foley, 1967; Varian, 1974), Budish (2011) uses a deterministic approximation of the same concept for course allocation.

Suppose each student is provided a budget of $B$ units of *token money*. Instead of using an indirect mechanism that involves explicit bidding, Budish leverages the utility profile information of students over schedules to calculate an approximate competitive equilibrium from equal incomes of a *course allocation problem with utility representation*, $[I, m, C, q, u]$, where $u = (u_i)_{i \in I}$ represents a profile of *utility functions* of students over schedules.[42]

A *price vector* is a non-negative vector $p = (p_c)_{c \in C} \in \mathbb{R}_+^{|C|}$. The *budget set* of a student $i \in I$ is given as,

$$\mathcal{B}_i(p|B) = \{S \subseteq C : |S| \leq m, \sum_{c \in S} p_c \leq B\}.$$

A price vector-matching pair $(p, \mu)$ is a *competitive equilibrium from equal incomes CEEI* if:

1. For each $i \in I$,
$$\mu(i) \in \max_{u_i} \mathcal{B}_i(p|B)$$
such that $\mu(i)$ is the cheapest schedule with this property.
2. For each $c \in C$,
$$\left|\mu^{-1}(c)\right| < q_c \implies p_c = 0.$$

[42]Budish (2011) considers more complex constraints on course assignments through feasibility constraints. For expositional reasons, we only consider a simple version of these constraints here.

Note that, a CEEI may not exist for a given course allocation problem with a utility representation.

In pursuit of an approximate positive existence result, Budish (2011) allows some course capacities to be exceeded in allocation and token budgets of some students to be increased by small error margins. He defines an $(\alpha, \beta)$*-approximate CEEI* (*A-CEEI*) for given $\alpha, \beta \in \mathbb{R}_+$ as a price vector-matching-budget vector triple $(p, \mu; \overline{B})$ with $\overline{B} = (B_i)_{i \in I} \in \mathbb{R}_+^{|I|}$ such that:

1. For each $i \in I$,
$$\mu(i) \in \max_{u_i} \mathcal{B}_i(p|B_i).$$
2. The market clearing error is within $\alpha$: if the market clearing error for each course $c \in C$ is
$$e_c = \begin{cases} \max\left\{\left|\mu^{-1}(c)\right| - q_c, 0\right\} & \text{if } p_c = 0 \\ \left|\mu^{-1}(c)\right| - q_c & \text{if } p_c > 0 \end{cases},$$
then we have
$$\sqrt{\sum_{c \in C} e_c^2} \leq \alpha.$$
3. The ratio of maximum to minimum induced budgets is within $\beta$:
$$\frac{\max_{i \in I} B_i}{\min_{i \in I} B_i} \leq 1 + \beta.$$

**Theorem 21 (Budish, 2011)** *Consider a course allocation problem $[I, m, C, q, u]$ with utility representation. If we can adjust the capacities of courses, for any $\beta > 0$, there exists a $\left(\left(\sqrt{|C| \min\{2m, |C|\}}\right) \Big/ 2, \beta\right)$-A-CEEI.*

We can make the budgets of students as equal as possible, without making them exactly the same, and with an additional seat assignment of less than, on average, one seat per course, an A-CEEI exists.

An A-CEEI, as given in the above theorem, may not be Pareto efficient, as positively priced courses are not necessarily fully filled, unlike in exact CEEI; however, the inefficiency is small as the market clearing error is small.

In addition, the outcome matching is not necessarily envy-free, unlike a matching supporting a CEEI, as the token budgets of agents can be different. But it satisfies the following approximate fairness concept.

A matching $\mu$ is *envy-free by up to one course* if for any pair of students $i, j \in I$,

$$u_i(\mu(i)) < u_i(\mu(j)) \quad \Longrightarrow \quad u_i(\mu(i)) \geq u_i(\mu(j) \setminus \{c\}) \text{ for some } c \in \mu(j).$$

**Proposition 9 (Budish, 2011)** *Consider a course allocation problem $[I, m, C, q, u]$ with util-*

*ity representation. If we can adjust the capacities and $\beta < \frac{1}{m-1}$, then any matching supporting a $\left(\left(\sqrt{|C| \min\{2m, |C|\}}\right) \big/ 2, \beta\right)$-A-CEEI satisfies envy-freeness by up to one course.*

An *A-CEEI mechanism induced by $\beta > 0$* is a direct mechanism that assigns each student $i$ a uniformly random budget in $B_i \in [B, (1+\beta)B]$, and computes its $\left(\left(\sqrt{|C| \min\{2m, |C|\}}\right) \big/ 2, \beta\right)$-A-CEEI.[43]

Though not strategy-proof, Budish (2011) demonstrates the mechanism's favorable incentive properties in larger market sizes under certain regularity assumptions (also see Chapter 2 of this handbook about large market results).

Implementing A-CEEI as a real-life course allocation mechanism faces challenges due to the complexity of utility functions over schedules that each student must report. Budish and Kessler (2022) reported implementation of this mechanism in the University of Pennsylvania's Wharton School of Business using a fairly simple utility reporting language. It has been successfully used for course allocation to MBA students. The A-CEEI implementation for course registration represents another successful application of matching theory that has changed real-life practices.

## 7.3 Other Designs for Course Allocation

Bichler and Merting (2021) report a successful design of a course allocation mechanism using student-proposing deferred acceptance algorithm in the Computer Science Depratment of Technical University of Munich in Germany (also see Diebold et al., 2014 for the underlying study).

This design does not allow schedule-based allocation involving complementarities in preferences, although courses have priority rankings over students. Bichler and Merting conducted an experimental study involving two other alternatives. One is a design by Nguyen, Peivandi, and Vohra (2016) based on a generalization of the probabilistic serial mechanism of Bogomolnaia and Moulin (2001) (covered in Section 3 of Chapter 1) when individuals have preferences allowing complementarities over schedules. Instead of each individual *eating* from one object at a time, they consume from a whole schedule until feasibility constraints kick in. Nguyen, Peivandi, and Vohra (2016) show that the resulting probabilistic assignments can be implemented as lotteries over bundles by some slight error regarding the correct capacity of the courses, extending the Birkhoff (1946)-von Neumann (1953) Theorem to this combinatorial setting. In a similar vein, this mechanism is asymptotically strategy-proof in

[43]This random mechanism is not uniquely defined because the random assignment outcome for each student depends on the method one uses to solve the computational A-CEEI problem given the error bounds, as there is not necessarily a unique A-CEEI for a given error bound vector and budget vector for students. The computation of an A-CEEI is quite cumbersome, and Budish et al. (2017) proposes a feasible computational method to implement it.

a large market. Bichler and Merting (2021) find favorable evidence to adopt the *bundled probabilistic serial* mechanism in their empirical study based on experiments and empirical studies.

### 7.4 Extension: Allocation of Food to Food Banks

Another interesting application of a variant of the multi-unit demand model with common ownership we discussed here pertains to the allocation of food donations to food banks (Prendergast, 2017). Feeding America is the largest non-profit in America, operating a large national network of food banks that provide food to meal programs and food pantries. These, in return, distribute meals to the needy. Feeding America needs to manage the allocation of food donations from large grocery store chains as well as other local retailers to the local food banks on a regular basis, depending on the needs of the banks.

Prendergast (2017) reports a successful design and implementation of a dynamic bidding system, similar in vein to the UMBS course bidding mechanism. The system endows food banks with token money, which they use to bid to a centrally managed online bidding system for various food items depending on their needs.

The desire to adopt such a mechanism stemmed from the fact that the headquarters of Feeding America did not have a good estimate of the heterogeneous needs of local food banks, and thus, a preference revelation system was needed. For efficient allocation and incentivizing truthful preference revelation, the headquarters considered adopting an auction using real money. They eventually decided that this would be against the philosophy of the charity, and as a result, they settled on a design involving bidding through token money, similar to allocating courses to students.

## 8 Other Notable Applications

We conclude the chapter with brief discussions of several other notable applications of matching theory.

### 8.1 Centralized School Admissions Through Exams as an Application of Matching under Priority-based Entitlements

Many countries employ centralized college and high school admission through exams, first observed by Balinski and Sönmez (1999) within the context of Turkey. Some of the other countries utilizing centralized methods through exam scores include China (**chen/jiang/kesten:20-empirical**; Chen and Kesten, 2017, 2019), Greece, Hungary (Biró, 2008), Taiwan (Dur et al., 2022), and others.

Centralized school admissions via exams are a direct application of models discussed in Chapter 1, Section 4 on matching under priority-based entitlements.

Two of these countries, Turkey (Balinski and Sönmez, 1999) and Hungary (Biró, 2008), have historically adopted mechanisms that satisfy no justified envy. These mechanisms are designed by system operators without the intervention of design economists.

The Turkish mechanism, on the other hand, utilized a Pareto-inferior mechanism prone to certain other unfairness and the possibility of another type of manipulation besides preference manipulation by the students. A student may intentionally underperform in an exam and get a lower score to improve their assigned college (Balinski and Sönmez, 1999). However, the Hungarian system employs the Gale-Shapley student-optimal stable mechanism, which is devoid of such issues.

It is also worthwhile to discuss a "hybrid" mechanism, which combines design elements of both the two-sided college admissions model by Gale and Shapley (1962) and the priority-based student placement model proposed by Balinski and Sönmez (1999). This hybrid approach is relevant in German college admissions, as examined by Westkamp (2013). Initially, seats at universities are allocated to students with strong claims, followed by consideration of university preferences for the remaining seats based on a specific formula. However, as shown by Westkamp (2013), this mechanism creates strong incentives for student manipulation.

## 8.2 Design for School Choice Mechanisms as an Application of Matching under Priority-based Entitlements

Many school districts worldwide use a matching mechanism to assign children to public schools instead of automatically enrolling them in their neighborhood schools. However, one of the historically most popular mechanisms, referred as the "Boston" mechanism, was shown to have many unappealing properties by Abdulkadiroğlu and Sönmez (2003). As a result of their efforts, the Boston public school district was persuaded to adopt a strategy-proof mechanism without justified envy, the Gale-Shapley student-optimal stable mechanism (Abdulkadiroğlu et al., 2005, 2006; Pathak and Sönmez, 2008; Sönmez, 2023). This constitutes a direct application of matching markets with priority-based entitlements, as covered in Chapter 1.

There was also an implementation in the New York City school district that featured both a two-sided structure with private property rights and matching under priority-based entitlements. Some schools were free to submit their own preference lists over applicants and capacities as in two-sided matching markets. However, most of the schools were objects whose priority orders were determined by the school district (Abdulkadiroğlu, Pathak, and Roth, 2005, 2009). This district also implemented the Gale-Shapley student-optimal stable mechanism.

These applications and many other applications in school choice are partially based on Section 4 of Chapter 1, matching under priority-based entitlements. As one of the most notable applications of matching theory, Chapter 4 of this handbook is devoted to school choice.

## 8.3 Design of Reserve Systems for the Allocation of Scarce Critical Care Resources in Times of Public Health Crisis

The COVID-19 pandemic has reignited discussions about the allocation guidelines for limited medical resources such as ventilators, ICU beds, vaccines, and antivirals during public health crises. Traditionally, prevailing guidelines in emergency healthcare medicine predominantly relied on priority systems as the primary means of resource allocation. A priority system relies on a priority ranking of individuals or categories of individuals, and using this priority order assigns scarce resources. However, there are usually multiple ethical criteria that policymakers would like to respect when allocating scarce medical resources, not just one. For example, in the case of ventilators and ICU beds, besides a criterion that ranks individuals based on the expected success of the ventilator treatment in saving them, doctors would also like to give some priority to essential personnel who get sick. However, a priority system is not flexible enough to use two different criteria to allocate ventilators.

Starting with Pathak et al. (2023), a series of recent studies in economics, bioethics, and emergency healthcare medicine have challenged the limitations of priority systems, documenting instances where decision-makers faced challenges in adequately integrating and balancing ethical values in their guidelines (Pathak, Sönmez, and Ünver, 2020, Schmidt, 2020, Schmidt et al., 2020, Persad, Peek, and Emanuel, 2020, Galiatsatos et al., 2020, Sönmez et al., 2021, Pathak, Sönmez, and Ünver, 2021, Makhoul and Drolet, 2021, Persad et al., 2022). They advocated adopting a reserve system.

The theoretical basis of such classes of "generalized" reserve systems was discussed in Chapter 1, Section 4, using the theoretical framework and main results of Pathak et al. (2023). For example, in the case of ventilator and ICU bed allocation, a reserve system can be used to create a reserve category for essential personnel besides the general category of individuals that are prioritized based on the metric of saving most lives.

As the pandemic unfolded, evolving in real-time, these studies influenced policy decisions. Notably, Massachusetts embraced the reserve system for allocation of monoclonal antibodies (Rubin et al., 2021), while Pennsylvania implemented dynamic allocation methods for therapeutics using lotteries to implement equivalent reserve systems (White et al., 2022; McCreary et al., 2023). Furthermore, informed by these studies and the advocacy of bioethicists, the US National Academies of Sciences, Engineering, and Medicine (NASEM) recommended an *over-and-above reserve system* based on geographical social vulnerability for the allocation of COVID-19 vaccines (NASEM, 2020). Consequently, over 15 US states adopted various versions of reserve systems in their vaccine allocation protocols.

Tennessee was the first state to adopt a reserve system for vaccine allocation (TDH, 2020). Several states followed suit, emphasizing equity and social justice in COVID-19 vaccine allocation, including Massachusetts, California, New Hampshire, North Carolina, Connecticut, Florida, Minnesota, Colorado, Mississippi, Maryland, Nebraska, New Mexico, Georgia, Illinois, Richmond and Henrico Counties in Virginia, and Washington, DC (see Pathak et al., 2023 for details of these applications).

## 8.4 Design of Israeli Psychology Master's Match as an Application of Two-Sided Matching with Contracts

As a market design effort utilizing the matching with contracts model, Hassidim, Romm, and Shorrer (2017) report the successful adoption of a mechanism that they designed for matching students to Psychology Master's programs in Israel (see Chapter 9 of this Handbook for more on the matching with contracts model by Hatfield and Milgrom, 2005). Before 2014, this market was decentralized and suffered from a chaotic admission process akin to the period before the adoption of the NRMP's centralized matching procedure in 1950's (see Section 6). The authors proposed to the authorities to design a centralized matching mechanism akin to the NRMP matching mechanism with a few notable differences. Although Master's programs and students constitute the two sides of the market, the choices of students are more complicated than those of doctors due to different scholarship options, and the programs have non-substitutable choice functions, unlike hospitals. Moreover, these program-specific choice functions are highly heterogeneous across different programs. Despite the availability of different tracks, scholarship options, and other complicated constraints imposed by programs, in general, these choice functions satisfy a weaker condition called *substitutable completability*, originally introduced by Hatfield and Kominers (2015), guaranteeing the existence of a stable matching. Moreover, these choice functions also satisfy size monotonicity (or the law of aggregate demand; see Alkan and Gale, 2003 and Hatfield and Milgrom, 2005), and hence, the student-proposing deferred acceptance procedure turns out to be strategy-proof for students for submitting their preferences (Hatfield and Kominers, 2015). A variation of this procedure, which permits the couples to submit joint preferences and utilizes a version of the sequential-offer heuristic (cf. Section 6) was adopted. Hassidim, Romm, and Shorrer (2017) report that the main challenge in the design was coming up with the correct preference expression language and choice function representation for the programs. This mechanism was successfully adopted in 2014 and has been in use as of the 2024 run. Additional insights into the running of this market and evidence regarding the suboptimal behavior of the applicants in submitting their preferences are provided in Hassidim, Romm, and Shorrer (2021).

## 8.5 Applications on Matching with Distributional Constraints

Kamada and Kojima (2015) observed that new physician markets in Japan differ from the NRMP matching program in the US in one substantial way. In 2008, the Japanese government instated a "regional cap," restricting the total number of residents matched within each of the country's prefectures. This move aimed to balance the geographical distribution of doctors, counteracting the concentration in urban centers at the detriment of rural areas. Following the adoption of these regional caps, a modified mechanism known as the Japan Residency Matching Program (JRMP) was introduced. This adjusted system addresses the caps by equalizing hospital capacities if the total exceeds the regional limit. Essentially, if the combined capacity of hospitals surpasses the cap, each hospital's capacity is scaled down proportionally to align with the regional restriction. Subsequently, the doctor-proposing deferred acceptance algorithm, previously used before 2008, is employed under these adjusted capacities.[44]

Kamada and Kojima (2015) also drew attention to other examples with similar aggregate capacity constraints. China's graduate school admissions strategy mirrors Japan's approach in its new physician market. Annually placing over 400,000 students since 2009 in academic and professional tracks, China aimed to increase its professional master's degree holders. Implementing limitations on admissions to academic master's programs in 2010, the government reduced available seats in each program by around 25% by 2015.

Similarly, Ukrainian college admissions impose hard caps on the number of state-funded positions, establishing an aggregate capacity constraint for these slots.

Kamada and Kojima (2015) have shown that, tweaking the notion of "no justified envy" slightly and endogenously adjusting hospital capacities could pave the way for a more efficient and strategy-proof mechanism than JRMP mechanism. Moreover, more intricate distributional constraints beyond aggregate hierarchical capacity constraints, such as matroidal constraints in priority-based entitlements (Hafalir et al., 2022), can also be integrated. Expanding on this, Kamada and Kojima (2023b) generalizes the theory of constraints for matching under priority-based entitlements and two-sided matching, characterizing constraints that allow for a student-optimal matching without justified envy (also see Kamada and Kojima, 2018).

[44] They note that similar restrictions exist in the UK residency matching market for new doctors as of 2015.

## 8.6 Applications of Matching with Reassignment

**Applications of Matching under Mixed Priority-based Entitlements and Ownership.** Inspired by dormitory allocation in US college campuses, a model of mixed-ownership economies by Abdulkadiroğlu and Sönmez (1999) was discussed in Chapter 1, Section 3. This model accounts for upperclassmen retaining their college dorm rooms for the upcoming year if they wanted to. Dormitory allocation is, therefore, an application of matching models with reassignment of individuals, who have inherited certain property rights due to their past assignments.

Notably, Abdulkadiroğlu and Sönmez, 1999 also discusses an intriguing dormitory assignment mechanism employed in an MIT dorm called NH4. As shown by Guillen and Kesten, 2012, this mechanism is equivalent to implementing the student-proposing deferred acceptance (DA) algorithm by using a constructed priority profile that assigns the highest priority to the tenants of rooms. Under this mechanism, dorm rooms hold priorities over students, as in the priority-based entitlement model of Chapter 1, Section 4. However, a student already occupying a room is automatically moved to the top priority in that room's priority order, regardless of their priority ranking in other rooms. Subsequently, the student-proposing DA algorithm is employed using this priority structure. The priority order doesn't align with the intended initial order, so the outcome does not necessarily satisfy no justified envy. Nevertheless, the NH4 mechanism is individually rational, unlike the random priority mechanism with squatting rights used in many colleges and highlighted in Abdulkadiroğlu and Sönmez (1999). Further analysis of the NH4 mechanism is carried out by Guillen and Kesten (2012) both theoretically and experimentally.

Interestingly, Compte and Jehiel, 2008 observed that the NH4 mechanism is precisely the mechanism used in the centralized assignment and reassignment of tens of thousands of secondary education teachers annually in France. Teaching positions are public service jobs in France, and a centralized mechanism is used to assign and reassign teachers to schools. Priority orders of schools over teachers are determined through various factors. To encourage voluntary participation, teachers who seek reassignment are given the highest priority at their own schools, regardless of how they would be ranked by the French Ministry of Education priority system. Then, the teacher-proposing DA algorithm with the updated priority rankings is used. This mechanism satisfies both individual rationality and strategy-proofness. Moreover, Combe, 2023 demonstrates other desirable properties of this mechanism, such as minimizing justified envy among all individually rational and strategy-proof mechanisms. However, Combe, Tercieux, and Terrier, 2022 observe that this mechanism lacks important efficiency features, resulting in too many blocks impeding mobility and causing excessive efficiency loss. Consequently, they propose a new strategy-proof and efficient mechanism akin to the top-trading-cycles-like mechanisms in the many-to-one setting, such as the "You-Request-My-House I Get-Your-Turn" (YRMH-IGYT) mechanism of Abdulkadiroğlu and Sönmez (1999), discussed extensively in Section 3 of Chapter 1.

A more recent paper Combe et al. (2022) also embeds the distributional objective of making teacher quality as equal as possible across schools – a stated goal of the Ministry of Education in France. They develop an inequality-decreasing, efficient, individually rational, and strategy-proof mechanism.

**Balancedness Axiom and Its Applications.** Related class applications require an explicit *balancedness* condition to be satisfied. In these problems, there are no new applicants being assigned to vacant positions, and every matching is a reassignment of individuals with an initial match.

Dur and Ünver (2019) introduced the balancedness condition and examined an application known as *tuition exchange* among US colleges. Given the substantial cost of college education in the US, many colleges offer free education for the children of their faculty members, provided the student gains admission to the school. However, smaller colleges might not offer all necessary programs, limiting the choices available to faculty members for their children's education. Tuition exchange programs facilitate opportunities among faculty members of different member institutions to swap tuition benefits for their children.

Maintaining a balance between outgoing and incoming students with tuition benefits is crucial for colleges in these programs. However, achieving stability and balancedness might be conflicting objectives. Dur and Ünver (2019) documented that decentralized methods used in these markets led to the shutdown of some programs. Currently, tens of thousands of students use these programs annually. Dur and Ünver (2019) proposed a two-sided matching mechanism that satisfies balancedness. Their mechanism is strategy-proof for students, individually rational for both colleges and students, and under reasonable assumptions about their preferences, colleges would truthfully disclose their capacities.

Additionally, certain longstanding exchange programs worldwide, such as those for doctors and teachers, operate based on similar principles of balancedness.

Similarly, diversity considerations substantially enrich this problem domain when the balancedness condition is embedded in reassignment. An application of reassignment under priority-based entitlements and diversity constraints was proposed in Hafalir, Kojima, and Yenmez (2022) for inter-district school choice. Embedding diversity goals in the Erasmus student assignment program across European countries also presents an intriguing application when efficiency is the goal (Dur, Kesten, and

Ünver, 2015).

Kamada and Kojima (2023a) considered an aggregate balancedness axiom for student assignment to kindergartens in Tokyo. A kindergarten district can accommodate a student living in a different region as long as one of its resident students is assigned to an out-of-district kindergarten. Tokyo has many kindergarten districts that they refer to as regions. Each region's kindergarten system is autonomous and locally financed, making the balancedness axiom relevant. They consider matchings and mechanisms that satisfy individual rationality, balancedness, and no justified envy. Unlike the previously discussed problems, non-wastefulness can be violated in this domain, as non-wastefulness and balancedness may not be compatible with each other. Kamada and Kojima (2023a) study constrained efficient mechanisms among the ones that satisfy the three properties: individual rationality, balancedness, and no justified envy when priorities are weak or strict for districts. They show that there is no strategy-proof mechanism satisfying these three properties along with constrained efficiency. Nevertheless, they propose a mechanism, naturally not strategy-proof, for achieving these goals after characterizing the graph-theoretic properties implied by a matching that satisfies individual rationality, balancedness, no justified envy, and constrained efficiency.

There are also applications in two-sided matching markets with status-quo matchings where schools or firms are involved as transacting parties with preferences and stakes in allocation. In these applications, the reassignment of workers is a crucial aspect of the market.

## 8.7 Refugee Assignment as a Market Design Problem

Delacrétaz, Kominers, and Teytelboym (2023) report that over 79.5 million people were displaced due to conflicts in 2019 in the world, among whom 20 million are categorized as *refugees* by UNHCR, the United Nations Refugee Agency. While around 1.44 million refugees are unable to safely return home, they are eligible for resettlement in countries offering permanent residence. The process of determining where refugees are resettled has recently been studied within the matching theory framework. Jones and Teytelboym (2016, 2017) and Andersson (2019) are some of the first papers that attracted the attention of market designers to the refugee resettlement problem. With the efforts of design economists, operations researchers, and computer scientists, a machine-learning-based assignment system, Annie MOORE, is now being used in the US by The Hebrew Immigrant Aid Society (HIAS), a major refugee resettlement agency, to find suitable places for refugees (Ahani et al., 2021).

Although crafted by market design experts, this mechanism still carries limitations. One notable drawback is its failure to account for the preferences of refugees.

This oversight is particularly significant considering that the initial resettlement location substantially impacts refugees' future prospects, as many do not move again for years.

To address this issue, Delacrétaz, Kominers, and Teytelboym (2023) propose some new mechanisms utilizing refugee preferences and locality priorities into account. The problem differs from the ones we covered in Chapter 1 and in this chapter in an important aspect. The refugees are usually allocated as families, and each family has multi-dimensional constraints to be taken into account in matching, for example, how many of the family members need jobs, how many of them require schooling, how big of a house is needed for the family, etc. These multi-dimensional constraints span an interesting, difficult, and much-studied operations research problem motivated by *knapsack packing*: how do we fit the maximum number of *items* with different dimensional sizes in a *knapsack* with a given capacity in each dimension? Taking into account the family preferences over localities, locality priorities, and these knapsack constraints of families, Delacrétaz, Kominers, and Teytelboym (2023) propose different mechanisms based on different interpretations of property rights over the locality slots. To accommodate priority-based entitlements, they define a new and interesting concept of no justified envy that they refer to as *weak envy-freeness* and propose a mechanism fulfilling this objective along with strategy-proofness for families (the mechanism leads to some waste as fulfilling all these objectives is not possible). This mechanism uses a variant of family-proposing deferred acceptance algorithm and finds a family-optimal weakly envy-free outcome.

Additionally, they propose an efficient allocation mechanism that is strategy-proof, Pareto efficient, and individually rational, building on the idea of top-trading cycles. Also interpreting the Annie MOORE system's assignment as initial endowments for refugee families, they propose an efficient mechanism based on trading these localities among families, akin to a version of top-trading cycles.

There has been increased interest in different aspects of this problem from market designers, especially due to the immense refugee flow to Europe and other Middle Eastern countries during the Syrian civil war of the 2010's and to other European countries during the Russia-Ukraine war in the 2020's. For example, Caspari (2019) studies how to design a centralized asylum seeker processing and assignment mechanism that fits into the European Union's by-laws, utilizing the preferences of asylum seekers. Andersson and Ehlers (2020) consider housing for refugee settlement in Sweden as a matching mechanism design problem.[45]

[45] See Ahani et al. (2023) and Caria et al. (2023) for further work on refugee resettlement studies from a market design point of view.

# References

Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth (2005). "The New York City High School Match." *American Economic Review Papers and Proceedings*, 95, 364–367.

Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth (2009). "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match." *American Economic Review*, 99(5), 1954–1978.

Atila Abdulkadiroğlu, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez (2005). "The Boston Public School Match." *American Economic Review Papers and Proceedings*, 95, 368–372.

Atila Abdulkadiroğlu, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez (2006). "Changing the Boston School Choice Mechanism." NBER working paper 11965.

Atila Abdulkadiroğlu and Tayfun Sönmez (1999). "House Allocation with Existing Tenants." *Journal of Economic Theory*, 88, 233–260.

Atila Abdulkadiroğlu and Tayfun Sönmez (2003). "School Choice: A Mechanism Design Approach." *American Economic Review*, 93, 729–747.

David Abraham, Avrim Blum, and Tuomas Sandholm (2007). "Clearing Algorithms for Barter Exchange Markets: Enabling Nationwide Kidney Exchanges." *Proceedings of ACM-EC 2007: the Eighth ACM Conference on Electronic Commerce*.

Nikhil Agarwal, Itai Ashlagi, Eduardo M Azevedo, Clayton R Featherstone, and Ömer Karaduman (2019). "Market failure in kidney exchange." *American Economic Review*, 109 (11), 4026–4070.

Dhiraj Agrawal, Subhash Gupta, and Sanjiv Saigal (2023). "Paired exchange living donor liver transplantation: Indications, stumbling blocks, and future considerations." *Journal of Hepatology*, 78 (3), 643–651.

Dhiraj Agrawal, Sanjiv Saigal, Shekhar Singh Jadaun, Shweta A Singh, Shaleen Agrawal, and Subhash Gupta (2022). "Paired exchange living donor liver transplantation: a nine-year experience from North India." *Transplantation*, 106 (11), 2193–2199.

Narges Ahani, Tommy Andersson, Alessandro Martinello, Alexander Teytelboym, and Andrew C Trapp (2021). "Placement optimization in refugee resettlement." *Operations Research*, 69 (5), 1468–1486.

Narges Ahani, Paul Gölz, Ariel D Procaccia, Alexander Teytelboym, and Andrew C Trapp (2023). "Dynamic placement in refugee resettlement." *Operations Research*.

Mohammad Akbarpour, Julien Combe, Yinghua He, Victor Hiller, Robert J Shimer, and Olivier Tercieux (2020). "Unpaired Kidney Exchange: Overcoming Double Coincidence of Wants without Money." *Review of Economic Studies*, forthcoming.

Ahmet Alkan and David Gale (2003). "Stable schedule matching under revealed preference." *Journal of Economic Theory*, 112, 289–306.

Allahabad High Court (2016). "Ashish Kumar Pandey And 24 Others vs. State Of U.P. And 29 Others on 16 March 2016." https://indiankanoon.org/doc/74817661/.

Frederike Ambagtsheer, Bernadette Haase-Kromwijk, Frank J M F Dor, Greg Moorlock, Franco Citterio, Thierry Berney, and Emma K Massey (2020). "Global Kidney Exchange: opportunity or exploitation? An ELPAT/ESOT appraisal." *Transplant International*, 33 (9), 989–998.

Ross Anderson, Itai Ashlagi, David Gamarnik, Michael Rees, Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2015). "Kidney Exchange and the Alliance for Paired Donation: Operations Research Changes the Way Kidneys Are Transplanted." *Interfaces*, 45 (1), 26–42.

Tommy Andersson (2019). "Refugee Matching as a Market Design Application." *The Future of Economic Design: The Continuing Development of a Field as Envisioned by Its Researchers*. Ed. by Jean-Francois Laslier, Herve Moulin, M. Remzi Sanver, and William S. Zwicker. Springer International Publishing, pp. 445–450.

Tommy Andersson and Lars Ehlers (2020). "Assigning Refugees to Landlords in Sweden: Efficient, Stable, and Maximum Matchings*." *The Scandinavian Journal of Economics*, 122 (3), 937–965.

Tommy Andersson and Jörgen Kratz (May 2019). "Pairwise Kidney Exchange over the Blood Group Barrier." *The Review of Economic Studies*, 87 (3), 1091–1133.

Itai Ashlagi, Mark Braverman, and Avinatan Hassidim (2014). "Stability in Large Matching Markets with Complementarities." *Operations Research*, 62 (4), 713–732.

Itai Ashlagi, Felix Fischer, Ian A Kash, and Ariel D Procaccia (2015). "Mix and match: A strategyproof mechanism for multi-hospital kidney exchange." *Games and Economic Behavior*, 91, 284 –296.

Itai Ashlagi and Alvin E Roth (2014). "Free riding and participation in large scale, multi-hospital kidney exchange." *Theoretical Economics*, 9, 817–865.

Itai Ashlagi and Alvin E Roth (2021). "Kidney exchange: an operations perspective." *Management Science*, 67 (9), 5455–5478.

Lawrence M Ausubel and Thayer Morrill (2014). "Sequential Kidney Exchange." *American Economic Journal: Microeconomics*, 6 (3), 265–85.

Christopher Avery, Christine Jolls, Richard A Posner, and Alvin E Roth (2001). "The market for federal judicial law clerks." *University of Chicago Law Review*, 68, 793.

Orhan Aygün and Inácio Bó (2021). "College admission with multidimensional privileges: The Brazilian affirmative action case." *American Economic Journal: Microeconomics*, 13 (3), 1–28.

Orhan Aygün and Bertan Turhan (2020). "Dynamic Reserves in Matching Markets." *Journal of Economic Theory*, 188, 1050–1069.

Michel Balinski and Tayfun Sönmez (1999). "A Tale of Two Mechanisms: Student Placement." *Journal of Economic Theory*, 84, 73–94.

Surender Baswana, Partha Pratim Chakrabarti, Sharat Chandran, Yashodhan Kanoria, and Utkarsh Patange (2019). "Centralized Admissions for Engineering Colleges in India." *INFORMS Journal on Applied Analytics*, 49 (5), 338–354.

Martin Bichler and Soeren Merting (2021). "Randomized scheduling mechanisms: Assigning course seats in a fair and efficient way." *Production and Operations Management*, 30 (10), 3540–3559.

AW Bingaman, FH Wright Jr, M Kapturczak, L Shen, S Vick, and CL Murphey (2012). "Single-center kidney paired donation: the Methodist San Antonio experience." *American Journal of Transplantation*, 12 (8), 2125–2132.

Garrett Birkhoff (1946). "Three Observations on Linear Algebra." *Revi. Univ. Nac. Tucuman, ser A*, 5, 147–151.

Péter Biró (2008). "Student admissions in Hungary as Gale and Shapley envisaged." *University of Glasgow Technical Report TR-2008-291*.

Péter Biró, Bernadette Haase-Kromwijk, Tommy Andersson, Eyjólfur Ingi Ásgeirsson, Tatiana Baltesová, Ioannis Boletis, Catarina Bolotinha, Gregor Bond, Georg Böhmig, Lisa Burnapp, Katarina Cechlárová, Paola Di Ciaccio, Jiri Fronek, Karine Hadaya, Aline Hemke, Christian Jacquelinet, Rachel Johnson, Rafal Kieszek, Dirk R Kuypers, Ruthanne Leishman, et al. (2019). "Building kidney exchange programmes in Europe—an overview of exchange practice and activities." *Transplantation*, 103 (7), 1514.

Péter Biró, Flip Klijn, Xenia Klimentova, and Ana Viana (2023). "Shapley–Scarf Housing Markets: Respecting Improvement, Integer Programming, and Kidney Exchange." *Mathematics of Operations Research*.

Péter Biró, Joris Van de Klundert, David Manlove, William Pettersson, Tommy Andersson, Lisa Burnapp, Pavel Chromy, Pablo Delgado, Piotr Dworczak, Bernadette Haase, Aline Hemke, Rachel Johnson, Xenia Klimentova, Dirk Kuypers, Alessandro Nanni Costa, Bart Smeulders, Frits Spieksma, Maria O Valentin, and Ana Viana (2021). "Modelling and optimisation in European kidney exchange programmes." *European Journal of Operational Research*, 291 (2), 447–456.

Péter Biró, David F Manlove, and Romeo Rizzi (2009). "Maximum weight cycle packing in directed graphs, with application to kidney exchange programs." *Discrete Mathematics, Algorithms and Applications*, 1 (04), 499–517.

Anna Bogomolnaia and Hervé Moulin (2001). "A New Solution to the Random Assignment Problem." *Journal of Economic Theory*, 100, 295–328.

Anna Bogomolnaia and Hervé Moulin (2004). "Random Matching Under Dichotomous Preferences." *Econometrica*, 72, 257–279.

Bombay High Court (2016). "Asha Ramnath Gholap vs. President, District Selection on 30 March 2016." https://indiankanoon.org/doc/178693513/.

Bombay High Court (2019a). "Shilpa Sahebrao Kadam And Another vs. The State Of Maharashtra on 8 August 2019." https://indiankanoon.org/doc/89017459/.

Bombay High Court (2019b). "Smt. Tejaswini Raghunath Galande vs. The Chairman, Maharashtra Public ... on 22 January 2019." https://indiankanoon.org/doc/128600601/.

Slava Bronfman, Noga Alon, Avinatan Hassidim, and Assaf Romm (2018). "Redesigning the Israeli medical internship match." *ACM Transactions on Economics and Computation (TEAC)*, 6 (3-4), 1–18.

Eric Budish (2011). "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes." *Journal of Political Economy*, 119(6), 1061–1103.

Eric Budish, Gérard P Cachon, Judd B Kessler, and Abraham Othman (2017). "Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation." *Operations Research*, 65 (2), 314–336.

Eric Budish and Estelle Cantillon (Aug. 2012). "The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard." *American Economic Review*, 102 (5), 2237–2271.

Eric Budish and Judd B Kessler (2022). "Can market participants report their preferences accurately (enough)?" *Management Science*, 68 (2), 1107–1130.

Eric Budish, Robin Lee, and John Shim (2019). "A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?" NBER Working Paper 25855.

Jeremy Bulow and Jonathan Levin (2006). "Matching and Price Competition." *American Economic Review*, 96, 652–668.

A Stefano Caria, Grant Gordon, Maximilian Kasy, Simon Quinn, Soha Osman Shami, and Alexander Teytelboym (2023). "An adaptive targeted field experiment: Job search assistance for refugees in Jordan." *Journal of the European Economic Association*, jvad067.

Gian Caspari (2019). "An alternative approach to asylum assignment." Working paper.

Gian Caspari (2020). "Booster draft mechanism for multi-object assignment." *ZEW-Centre for European Economic Research Discussion Paper*, (20-074).

Central Government Act (1949). "Article 16 in The Constitution Of India 1949." https://indiankanoon.org/doc/211089/.

Yan Chen and Onur Kesten (2017). "Chinese college admissions and school choice reforms: A theoretical analysis." *Journal of Political Economy*, 125 (1), 99–139.

Yan Chen and Onur Kesten (2019). "Chinese college admissions and school choice reforms: An experimental study." *Games and Economic Behavior*, 115, 83–100.

Yao Cheng and Zaifu Yang (2021). "Efficient kidney exchange with dichotomous preferences." *Journal of Health Economics*, 80, 102536.

Youngsub Chun, Eun Jeong Heo, and Sunghoon Hong (2021). "Kidney Exchange with Immunosuppressants." *Economic Theory*, 72, 1–19.

Julien Combe (2023). "Reallocation with priorities and minimal envy mechanisms." *Economic Theory*, 76 (2), 551–584.

Julien Combe, Umut Dur, Olivier Tercieux, Camille Terrier, and M Utku Ünver (2022). "Market Design for Distributional Objectives in (Re) assignment: An Application to Improve the Distribution of Teachers in Schools." Working paper.

Julien Combe, Olivier Tercieux, and Camille Terrier (2022). "The design of teacher assignment: Theory and evidence." *The Review of Economic Studies*, 89 (6), 3154–3222.

Olivier Compte and Philippe Jehiel (2008). "Voluntary participation and reassignment in two-sided matching." Unpublished. http://www.enpc.fr/ceras/compte/matchingparticipation.pdf.

N Cowan, H A Gritsch, N Nassiri, J Sinacore, and J Veale (2017). "Broken Chains and Reneging: A Review of 1748 Kidney Paired Donation Transplants." *American Journal of Transplantation*, 17 (9), 2451–2457.

Ettore Damiano, Hao Li, and Wing Suen (2005). "Unravelling of Dynamic Sorting." *The Review of Economic Studies*, 72 (4), 1057–1076.

Marry De Klerk, Karin M Keizer, Frans H J Claas, Marian Witvliet, Bernadette J J M Haase-Kromwijk, and Willem Weimar (2005). "The Dutch National Living Donor Kidney Exchange Program." *American Journal of Transplantation*, 5 (9), 2302–2305.

David Delacrétaz, Scott Duke Kominers, and Alexander Teytelboym (2023). "Matching Mechanisms for Refugee Resettlement." *American Economic Review*, 113 (10), 2689–2717.

Francis L Delmonico (2004). "Exchanging Kidneys: Advances in Living-Donor Transplantation." *New England Journal of Medicine*, 350, 1812–1814.

Francis L Delmonico and Nancy L Ascher (2017). "Opposition to irresponsible global kidney exchange." *American Journal of Transplantation*, 17 (10), 2745–2746.

John P Dickerson and Tuomas Sandholm (2016). "Organ Exchanges: A Success Story of AI in Healthcare." Thirtieth Conference on Artificial Intelligence Tutorial Forum.

Franz Diebold, Haris Aziz, Martin Bichler, Florian Matthes, and Alexander Schneider (2014). "Course allocation via stable matching." *Business & Information Systems Engineering*, 6, 97–110.

Umut Dur, Onur Kesten, and M Utku Ünver (2015). "Maintaining Diversity in Student Exchange." Available at SSRN 4064921.

Umut Dur, Scott Duke Kominers, Parag A Pathak, and Tayfun Sönmez (2018). "Reserve Design: Unintended Consequences and The Demise of Boston's Walk Zones." *Journal of Political Economy*, 126, 2457–2479.

Umut Dur, Parag A Pathak, Fei Song, and Tayfun Sönmez (2022). "Deduction dilemmas: The Taiwan assignment mechanism." *American Economic Journal: Microeconomics*, 14 (1), 164–185.

Umut Dur, Parag A Pathak, and Tayfun Sönmez (2020). "Explicit vs. Statistical Preferential Treatment in Affirmative Action: Theory and Evidence from Chicago's Exam Schools." *Journal of Economic Theory*, 187, 104996.

Umut Dur and M Utku Ünver (2019). "Two-Sided Matching via Balanced Exchange." *Journal of Political Economy*, 127, 1156–1177.

Federico Echenique and Juan Sebastián Pereyra (2016). "Strategic complementarities and unraveling in matching markets." *Theoretical Economics*, 11 (1), 1–39.

Federico Echenique and M Bumin Yenmez (2015). "How to Control Controlled School Choice." *American Economic Review*, 105 (8), 2679–2694.

Jack Edmonds (1965). "Paths, Trees, and Flowers." *Canadian Journal of Mathematics*, 17, 449–467.

Jack Edmonds and Richard M Karp (1972). "Theoretical improvements in algorithmic efficiency for network flow problems." *Journal of the ACM*, 19, 248ñ264.

Haluk Ergin, Tayfun Sönmez, and M Utku Ünver (2018). "Efficient and incentive-compatible liver exchange." Boston College Working Papers in Economics 951, https://ideas.repec.org/p/boc/bocoec/951.html.

Haluk Ergin, Tayfun Sönmez, and M Utku Ünver (2020). "Efficient and Incentive-Compatible Liver Exchange." *Econometrica*, 88 (3), 965–1005.

Haydar Evren and Manshu Khanna (2024). "Affirmative Action's Cumulative Fractional Assignments." arXiv preprint arXiv:2111.11963.

Paolo Ferrari, Linda Cantwell, Joseph Ta, Claudia Woodroffe, Lloyd D'Orsogna, and Rhonda Holdsworth (2017). "Providing Better-Matched Donors for HLA Mismatched Compatible Pairs Through Kidney Paired Donation." *Transplantation*, 101 (3).

David Foley (1967). "Resource Allocation and the Public Sector." *Yale Economic Essays*, 7, 45–98.

Guillaume Frechette, Alvin E Roth, and M Utku Ünver (2007). "Unraveling Results from Inefficient Matching: Evidence from Post-Season College Football Bowls." *RAND Journal of Economics*, 38, 967–982.

Lucrezia Furian, Antonio Nicolò, Caterina Di Bella, Massimo Cardillo, Emanuele Cozzi, and Paolo Rigotti (2020). "Kidney exchange strategies: new aspects and applications with a focus on deceased donor-initiated chains." *Transplant International*, 33 (10), 1177–1184.

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

Panagis Galiatsatos, Allen Kachalia, Harolyn ME Belcher, Mark T Hughes, Jeffrey Kahn, Cynda H Rushton, Jose I Suarez, Lee Daugherty Biddison, and Sherita H Golden (2020). "Health equity and distributive justice considerations in critical care resource allocation." *The Lancet Respiratory Medicine*, 8 (8), 758–760.

S E Gentry, D L Segev, M Simmerling, and R A Montgomery (2007). "Expanding Kidney Paired Donation Through Participation by Compatible Pairs." *American Journal of Transplantation*, 7 (10), 2361–2370.

John S Gill, Kathryn Tinckam, Marie Chantal Fortin, Caren Rose, Kara Shick-Makaroff, Kimberly Young, Julie Lesage, Edward H Cole, Maeghan Toews, David N Landsberg, et al. (2017). "Reciprocity to increase participation of compatible living donor and recipient pairs in kidney paired donation." *American Journal of Transplantation*, 17 (7), 1723–1728.

David W Gjertson and J Michael Cecka (2000). "Living Unrelated Donor Kidney Transplantation." *Kidney International*, 58, 491–499.

Kyle Greenberg, Parag A Pathak, and Tayfun Sönmez (2023). "Redesigning the US Army's Branching Process: A Case Study in Minimalist Market Design." Forthcoming, *American Economic Review*.

Pablo Guillen and Onur Kesten (2012). "Matching Markets With Mixed Ownership: The Case For A Real-Life Assignment Mechanism." *International Economic Review*, 53 (3), 1027–1046.

Gujarat High Court (2020). "Tamannaben Ashokbhai Desai vs Shital Amrutlal Nishar on 5 August 2020." https://indiankanoon.org/doc/101656671/.

Vikraman Gunabushanam, Swaytha Ganesh, Kyle Soltys, George Mazariegos, Armando Ganoza, Michele Molinari, Amit Tevar, Christopher Hughes, and Abhinav Humar (2022). "Increasing living donor liver transplantation using liver paired exchange." *Journal of the American College of Surgeons*, 234 (2), 115–120.

Isa E Hafalir, Fuhito Kojima, and M Bumin Yenmez (2022). "Interdistrict school choice: A theory of student assignment." *Journal of Economic Theory*, 201, 105441.

Isa E Hafalir, Fuhito Kojima, M Bumin Yenmez, and Koji Yokote (2022). "Design on Matroids: Diversity vs. Meritocracy." ArXiv preprint https://arxiv.org/abs/2301.00237.

Isa E Hafalir, M Bumin Yenmez, and Muhammed A Yildirim (May 2013). "Effective affirmative action in school choice." *Theoretical Economics*, 8 (2), 325–363.

Xiang Han, Onur Kesten, and M Utku Ünver (2021). "Blood Allocation with Replacement Donors: A Theory of Multi-unit Exchange with Compatibility-based Preferences." WP.

Ernan Haruvy, Alvin E Roth, and M Utku Ünver (2006). "The Dynamics of Law Clerk Matching: An Experimental and Computational Investigation of Proposals for Reform of the Market." *Journal of Economic Dynamics and Control*, 30, 457–486.

Avinatan Hassidim, Assaf Romm, and Ran I Shorrer (2017). "Redesigning the Israeli Psychology Master's Match." *American Economic Review, Papers and Proceedings*, 107 (5), 205–09.

Avinatan Hassidim, Assaf Romm, and Ran I Shorrer (2021). "The limits of incentives in economic matching procedures." *Management Science*, 67 (2), 951–963.

John William Hatfield (2005). "Pairwise Kidney Exchange: Comment." *Journal of Economic Theory*, 125, 189–193.

John William Hatfield and Scott Duke Kominers (2015). "Hidden Substitutes." Working Paper.

John William Hatfield and Paul Milgrom (2005). "Matching with Contracts." *American Economic Review*, 95, 913–935.

Kyu Ha Huh, Myoung Soo Kim, Man Ki Ju, Hye Kyung Chang, Hyung Joon Ahn, Su Hyung Lee, Jong Hoon Lee, Soon Il Kim, Yu Seun Kim, and Kiil Park (2008). "Exchange Living-Donor Kidney Transplantation: Merits and Limitations." *Transplantation*, 86 (3).

Shin Hwang, Sung-Gyu Lee, Deok-Bog Moon, Gi-Won Song, Chul-Soo Ahn, Ki-Hun Kim, Tae-Yong Ha, Dong-Hwan Jung, Kwan-Woo Kim, Nam-Kyu Choi, Gil-Chun Park, Young-Dong Yu, Young-Il Choi, Pyoung-Jae Park, and Hea-Seon Ha (2010). "Exchange living donor liver transplantation to overcome ABO incompatibility in adult patients." *Liver Transplantation*, 16 (4), 482–490.

Aanund Hylland and Richard J Zeckhauser (1979). "The Efficient Allocation of Individuals to Positions." *Journal of Political Economy*, 87(2), 293–314.

Will Jones and Alexander Teytelboym (2016). "Choices, preferences and priorities in a matching system for refugees." *Forced Migration Review*, 51, 80–82.

Will Jones and Alexander Teytelboym (2017). "The Local Refugee Match: Aligning Refugees' Preferences with the Capacities and Priorities of Localities." *Journal of Refugee Studies*, 31(2), 152–178.

Dong-Hwan Jung, Shin Hwang, Chul-Soo Ahn, Ki-Hun Kim, Deok-Bog Moon, Tae-Yong Ha, Gi-Won Song, Gil-Chun Park, and Sung-Gyu Lee (2014). "Section 16. Update on Experience in Paired-Exchange Donors in Living Donor Liver Transplantation For Adult Patients at ASAN Medical Center." *Transplantation*, 97.

John Kagel and Alvin E Roth (2000). "Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment." *Quarterly Journal of Economics*, 115(1), 201–235.

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

Yuichiro Kamada and Fuhito Kojima (2018). "Stability and strategy-proofness for matching with constraints: A necessary and sufficient condition." *Theoretical Economics*, 13 (2), 761–793.

Yuichiro Kamada and Fuhito Kojima (2023a). "Ekkyo Matching: How to Integrate Fragmented Matching Markets for Welfare Improvement." Working Paper.

Yuichiro Kamada and Fuhito Kojima (2023b). "Fair matching under constraints: Theory and applications." *Review of Economic Studies*, rdad046.

Süleyman Kerimov, Itai Ashlagi, and Itai Gurvich (2023). "On the optimality of greedy policies in dynamic matching." *Operations Research*.

Jong Man Kim (2022). "Increasing Living Liver Donor Pools: Liver Paired Exchange Versus ABO-incompatible Living Donor Liver Transplantation." *Transplantation*, 106 (11), 2118–2119.

Bettina Klaus and Flip Klijn (2007). "Paths to stability for matching markets with couples." *Games and Economic Behavior*, 58 (1), 154–171.

Bettina Klaus, Flip Klijn, and Jordi Massó (2007). "Some things couples always wanted to know about stable matchings (but were afraid to ask)." *Review of Economic Design*, 11, 175–184.

Fuhito Kojima, Parag A Pathak, and Alvin E Roth (Aug. 2013). "Matching with Couples: Stability and Incentives in Large Markets*." *The Quarterly Journal of Economics*, 128 (4), 1585–1632.

Fuhito Kojima and M Utku Ünver (2008). "Random paths to pairwise stability in many-to-many matching problems: a study on market equilibration." *International Journal of Game Theory*, 36, 473–488.

Scott Duke Kominers and Tayfun Sönmez (2016). "Matching with slot-specific priorities: Theory." *Theoretical Economics*, 11 (2), 683–710.

Leonieke W Kranenburg, Willij Zuidema, Willem Weimar, Jan Passchier, Medard Hilhorst, Marry De Klerk, Jan N M IJzermans, and Jan J V Busschbach (2006). "One donor, two transplants: willingness to participate in altruistically unbalanced exchange donation." *Transplant International*, 19 (12), 995–999.

Jörgen Kratz (2021). “Triage in Kidney Exchange.” WP.

Aradhna Krishna and M Utku Ünver (2008). “Improving the Efficiency of Course Bidding at Business Schools: Field and Laboratory Studies.” *Marketing Science*, 27, 262–282.

Aradhna Krishna and Yu Wang (2007). “The Relationship Between Top Trading Cycles Mechanism and Top Trading Cycles and Chains Mechanism.” *Journal of Economic Theory*, 132, 539–547.

Hao Li and Sherwin Rosen (1998). “Unraveling in Matching Markets.” *American Economic Review*, 88(3), 371–387.

Shengwu Li (2017). “Ethics and market design.” *Oxford Review of Economic Policy*, 33 (4), 705–720.

Jinpeng Ma (1994). “Strategy-proofness and the strict core in a market with indivisibilities.” *International Journal of Game Theory*, 23, 75–83.

Alan T Makhoul and Brian C Drolet (2021). “A reserve system for the equitable allocation of a severe acute respiratory syndrome Coronavirus 2 vaccine.” *Chest*, 159 (3), 1292–1293.

David F Manlove and Gregg O'Malley (2012). “Paired and Altruistic Kidney Donation in the UK: Algorithms and Experimentation.” *Experimental Algorithms: 11th International Symposium, SEA 2012, Bordeaux, France, June 7-9, 2012. Proceedings*. Ed. by Ralf Klasing. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 271–282.

Erin K McCreary, Utibe R Essien, Chung-Chou H Chang, Rachel A Butler, Parag Pathak, Tayfun Sönmez, M Utku Ünver, Ashley Steiner, Maddie Chrisman, Derek C Angus, and Douglas B White (2023). “Weighted Lottery to Equitably Allocate Scarce Supply of COVID-19 Monoclonal Antibody.” *JAMA Health Forum*, 4 (9), e232774–e232774.

D G McVitie and L B Wilson (1971). “The stable marriage problem.” *Communications of the ACM*, 14, 486–490.

Susan Mongell and Alvin E Roth (1991). “Sorority rush as a two-sided matching mechanism.” *The American Economic Review*, 441–464.

Robert A Montgomery, Sommer E Gentry, William H Marks, Daniel S Warren, Janet Hiller, Julie Houp, Andrea A Zachary, J Keith Melancon, Warren R Maley, Hamid Rabb, et al. (2006). “Domino paired kidney donation: a strategy to make best use of live non-directed donation.” *The Lancet*, 368 (9533), 419–421.

NASEM (2020). “A Framework for Equitable Allocation of Vaccine for the Novel Coronavirus.” https://www.nap.edu/resource/25917/25914.pdf.

Thành Nguyen, Ahmad Peivandi, and Rakesh Vohra (2016). “Assignment problems with complementarities.” *Journal of Economic Theory*, 165, 209–241.

Thanh Nguyen and Rakesh Vohra (2018). "Near-feasible stable matchings with couples." *American Economic Review*, 108 (11), 3154–3169.

Antonio Nicolò and Carmelo Rodríguez-Álvarez (2012). "Transplant quality and patients' preferences in paired kidney exchange." *Games and Economic Behavior*, 74 (1), 299–310.

Antonio Nicolò and Carmelo Rodríguez-Álvarez (2017). "Age-based preferences in paired kidney exchange." *Games and Economic Behavior*, 102, 508–524.

Muriel Niederle and Alvin E Roth (2003). "Unraveling reduces mobility in a labor market: Gastroenterology with and without a centralized match." *Journal of Political Economy*, 111, 1342–1352.

Muriel Niederle and Alvin E Roth (2004). "The Gastroenterology Fellowship Match: How It Failed and Why It Could Succeed Once Again." *Gastroenterology*, 127, 658–666.

Muriel Niederle, Alvin E Roth, and M Utku Ünver (2013). "Unraveling Results from Comparable Demand and Supply: An Experimental Investigation." *Games*, 4, 243–282.

Gerhard Opelz (1997). "Impact of HLA Compatibility on Survival of Kidney Transplants from Unrelated Live Donors." *Transplantation*, 64, 1473–1475.

Sonia T Orcutt, Katsuhiro Kobayashi, Mark Sultenfuss, Brian S Hailey, Anthony Sparks, Bighnesh Satpathy, and Daniel A Anaya (2016). "Portal vein embolization as an oncosurgical strategy prior to major hepatic resection: anatomic, surgical, and technical considerations." *Frontiers in surgery*, 3, 14.

Parag A Pathak, Alex Rees-Jones, and Tayfun Sönmez (2020). "Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms." forthcoming, *Review of Economics and Statistics*.

Parag A Pathak and Tayfun Sönmez (2008). "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review*, 98(4), 1636–1652.

Parag A Pathak, Tayfun Sönmez, and M Utku Ünver (2020). "Improving Ventilator Rationing through Collaboration with Experts on Resource Allocation." *JAMA Network Open*.

Parag A Pathak, Tayfun Sönmez, and M Utku Ünver (2021). "Reserve Systems for Allocation of Scarce Medical Resources During the COVID-19 Pandemic: The Path From April 2020 to April 2021." *CHEST*, 160 (4), 1572–1575.

Parag A Pathak, Tayfun Sönmez, M Utku Ünver, and M Bumin Yenmez (2023). "Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing." *Management Science*, forthcoming.

Govind Persad, Parag A Pathak, Tayfun Sönmez, and M Utku Ünver (2022). "Fair access to scarce medical capacity for non-covid-19 patients: a role for reserves." *bmj*, 376.

Govind Persad, Monica E Peek, and Ezekiel J Emanuel (2020). "Fairly prioritizing groups for access to COVID-19 vaccines." *Jama*, 324 (16), 1601–1602.

Richard A Posner, Christopher Avery, Christine Jolls, and Alvin E Roth (2007). "The New Market for Federal Judicial Law Clerks." *University of Chicago Law Review*, 74, 447–486.

Canice Prendergast (2017). "How food banks use markets to feed the poor." *Journal of Economic Perspectives*, 31 (4), 145–162.

Rajasthan High Court - Jodhpur (2013). "Smt. Megha Shetty vs. State Of Raj. & Anr on 26 July 2013." https://indiankanoon.org/doc/78343251/.

Rajasthan High Court (2013). "Rajeshwari vs. State (Panchayati Raj Dep) and Others on 15 March 2013." https://indiankanoon.org/doc/128221069/.

F T Rapaport (1986). "The case for a living emotionally related international kidney donor exchange registry." *Transplantation Proceedings*, 18 (3), 5–9.

M A Rees, T B Dunn, C S Kuhr, C L Marsh, J Rogers, S E Rees, A Cicero, L J Reece, A E Roth, O Ekwenna, D E Fumo, K D Krawiec, J E Kopke, S Jain, M Tan, and S R Paloyo (2017). "Kidney Exchange to Overcome Financial Barriers to Kidney Transplantation." *American Journal of Transplantation*, 17 (3), 782–790.

Michael Rees, Alvin E Roth, Ignazio Marino, Kimberly Krawiak, Susan Rees, Krista Sweeeney, Ty Dunn, Jeff Punch, Michael Zimmerman, Feroz Aziz, Ali Abdul Kareem Al Obaidli, Itai Ashlagi, Vaughn Whittaker, Aparna Rege, Rachel Forbes, Christian Kuhr, Ricardo Correa-Rotter, Citterio Franco, Obi Okwenna, Siegfredo Paloyo, et al. (2022). "The First 52 Global Kidney Exchange Transplants: Overcoming Multiple Barriers to Transplantation." *Transplantation*, 106 (9S), S469.

L F Ross and E S Woodle (2000). "Ethical issues in increasing living kidney donations by expanding kidney paired exchange programs." *Transplantation*, 69, 1539–1543.

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

Alvin E Roth (1990). "New Physicians: A Natural Experiment in Market Organization." *Science*, 250, 1524–1528.

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

Alvin E Roth (2008). "Deferred acceptance algorithms: History, theory, practice, and open questions." *international Journal of game Theory*, 36, 537–569.

Alvin E Roth and Elliot Peranson (1999). "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." *American Economic Review*, 89 (4), 748–780.

Alvin E Roth and Uriel Rothblum (1999). "Truncation Strategies in Matching Markets: In Search of Advice for Participants." *Econometrica*, 67, 21–43.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2003). "Kidney Exchange." NBER Working paper 10002.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2004). "Kidney Exchange." *Quarterly Journal of Economics*, 119 (2), 457–488.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2005a). "A Kidney Exchange Clearinghouse in New England." *American Economic Review Papers and Proceedings*, 95, 376–380.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2005b). "Pairwise Kidney Exchange." *Journal of Economic Theory*, 125 (2), 151–188.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2005c). "Transplant Center Incentives in Kidney Exchange." Unpublished mimeo.

Alvin E Roth, Tayfun Sönmez, and M Utku Ünver (2007). "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences." *American Economic Review*, 97 (3), 828–851.

Alvin E Roth, Tayfun Sönmez, M Utku Ünver, Francis L Delmonico, and Susan L Saidman (2006). "Utilizing List Exchange and Non-directed Donation through 'Chain' Paired Kidney Donations." *American Journal of Transplantation*, 6 (11), 2694–2705.

Alvin E Roth and John Vande Vate (1990). "Random Path to Stability in Two-sided Matching." *Econometrica*, 58(6), 1475–1480.

Alvin E Roth and Xiaolin Xing (1994). "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions." *American Economic Review*, 84, 992–1044.

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

Emily Rubin, Scott L Dryden-Peterson, Sarah P Hammond, Inga Lennes, Alyssa R Letourneau, Parag Pathak, Tayfun Sönmez, and M Utku Ünver (2021). "A novel approach to equitable distribution of scarce therapeutics: institutional experience implementing a reserve system for allocation of COVID-19 monoclonal antibodies." *Chest*, 160 (6), 2324–2331.

Susan L Saidman, Alvin E Roth, Tayfun Sönmez, M Utku Ünver, and Francis L Delmonico (2006). "Increasing the Opportunity of Live Kidney Donation by Matching for Two- and Three-Way Exchanges." *Transplantation*, 81 (5), 773–782.

Saad Salman, Muhammad Arsalan, and Faisal Saud Dar (2023). “Launching Liver Exchange and the First 3-Way Liver Paired Donation.” *JAMA surgery*, 158 (2), 210–211.

Harald Schmidt (2020). “Vaccine rationing and the urgency of social justice in the Covid-19 response.” *Hastings Center Report*, 50 (3), 46–49.

Harald Schmidt, Parag Pathak, Tayfun Sönmez, and M Utku Ünver (2020). “Covid-19: how to prioritize worse-off populations in allocating safe and effective vaccines.” *bmj*, 371.

Lloyd S Shapley and Herbert Scarf (1974). “On Cores and Indivisibility.” *Journal of Mathematical Economics*, 1, 23–37.

AS Soin, Prashant Bhangui, Amit Rastogi, Tarun Piplani, Narendra Choudhary, Swapnil Dhampalwar, Fysal Kollantavalappil, Kamal Yadav, Ankur Gupta, Nikunj Gupta, et al. (2023). “Simultaneous 3-way Paired Exchange Liver Transplantation Without Nondirected Donation: Novel Strategy to Expand the Donor Pool.” *Transplantation*, 107 (6), e175–e177.

Tayfun Sönmez (1999a). “Can Pre-arranged Matches be Avoided in Two-Sided Matching Markets?” *Journal of Economic Theory*, 86, 148–156.

Tayfun Sönmez (1999b). “Strategy-Proofness and Essentially Single-Valued Cores.” *Econometrica*, 67, 677–690.

Tayfun Sönmez (2013). “Bidding for Army Career Specialties: Improving the ROTC Branching Mechanism.” *Journal of Political Economy*, 121(1), 186–219.

Tayfun Sönmez (2023). “Minimalist market design: A framework for economists with policy aspirations.” ArXiv preprint https://arxiv.org/abs/2401.00307.

Tayfun Sönmez, Parag A Pathak, M Utku Ünver, Govind Persad, Robert D Truog, and Douglas B White (2021). “Categorized priority systems: a new tool for fairly allocating scarce medical resources in the face of profound social inequities.” *Chest*, 159 (3), 1294–1299.

Tayfun Sönmez and Tobias Switzer (2013). “Matching with (Branch-of-Choice) Contracts at the United States Military Academy.” *Econometrica*, 81(2), 451–488.

Tayfun Sönmez and M Utku Ünver (2010). “Course Bidding at Business Schools.” *International Economic Review*, 51 (1), 99–123.

Tayfun Sönmez and M Utku Ünver (2013). “Market Design for Kidney Exchange.” *The Handbook of Market Design*. Ed. by Zvika Neeman Alvin E. Roth and Nir Vulkan. OUP, pp. 93–137.

Tayfun Sönmez and M Utku Ünver (2014). “Altruistically unbalanced kidney exchange.” *Journal of Economic Theory*, 152, 105–129.

Tayfun Sönmez and M Utku Ünver (2015). "Enhancing the Efficiency of and Equity in Transplant Organ Allocation via Incentivized Exchange." Available as Boston College Working Paper 868.

Tayfun Sönmez and M Utku Ünver (2017). "Market design for living-donor organ exchanges: An economic policy perspective." *Oxford Review of Economic Policy*, 33 (4), 676–704.

Tayfun Sönmez and M Utku Ünver (2023). "Influencing Policy and Transforming Institutions: Lessons from Kidney/Liver Exchange." *New Directions in Market Design*. Ed. by Irene Lo, Michael Ostrovsky, and Parag A. Pathak. Chicago IL: University of Chicago Press.

Tayfun Sönmez, M Utku Ünver, and M Bumin Yenmez (2020). "Incentivized Kidney Exchange." *American Economic Review*, 110 (7), 2198–2224.

Tayfun Sönmez, M Utku Ünver, and Özgür Yilmaz (2018). "How (Not) to Integrate Blood Subtyping Technology to Kidney Exchange." *Journal of Economic Theory*, 176, 193–231.

Tayfun Sönmez and M Bumin Yenmez (2019). *Affirmative Action in India via Vertical and Horizontal Reservations*. Boston College Working Papers in Economics 977. Boston College Department of Economics.

Tayfun Sönmez and M Bumin Yenmez (2022a). "Affirmative action in India via vertical, horizontal, and overlapping reservations." *Econometrica*, 90 (3), 1143–1176.

Tayfun Sönmez and M Bumin Yenmez (2022b). "Constitutional implementation of affirmative action policies in India." ArXiv preprint https://arxiv.org/abs/2203.01483.

David Steinberg (2011). "Compatible-Incompatible Live Donor Kidney Exchanges." *Transplantation*, 91 (3).

Supreme Court of India (1992). "Indra Sawhney, etc. vs. Union Of India and Others, Etc. on 16 November 1992." https://indiankanoon.org/doc/1363234/.

Supreme Court of India (1995). "Anil Kumar Gupta, etc vs. State of Uttar Pradesh and Others on 28 July 1995." https://indiankanoon.org/doc/1055016/.

Supreme Court of India (2020). "Saurav Yadav vs. The State Of Uttar Pradesh on 18 December 2020." https://indiankanoon.org/doc/27820739/.

TDH (2020). "Covid-19 Vaccination Plan." Version 1.0, October 16, https://www.cdc.gov/vaccines/covid-19/downloads/tennessee-jurisdiction-executive-summary.pdf.

Ashutosh Thakur (2021). "Matching in the civil service: A market design approach to public administration and development." ECONtribute Discussion Paper.

The Declaration of Istanbul Custodian Group (2020). "Statement of The Declaration of Istanbul Custodian Group Concerning Ethical Objections to the Proposed Global

Exchange Program." https://www.declarationofistanbul.org/images/stories/resources/policy_documents/DICGStatementonGKEP_Nov28_final.pdf reached on 11/15/2023.

Panos Toulis and David C Parkes (2015). "Design and analysis of multi-hospital kidney exchange mechanisms using random graphs." *Games and Economic Behavior*, 91, 360 –382.

M Utku Ünver (2001). "Backward unraveling over time: The evolution of strategic behavior in the entry level British medical labor markets." *Journal of Economic Dynamics and Control*, 25 (6), 1039 –1080.

M Utku Ünver (2005). "On the survival of some unstable two-sided matching mechanisms." *International Journal of Game Theory*, 33 (2), 239–254.

M Utku Ünver (2010). "Dynamic Kidney Exchange." *Review of Economic Studies*, 77 (1), 372–414.

Hal R Varian (1974). "Equity, Envy, and Efficiency." *Journal of Economic Theory*, 9, 63–91.

Jeffrey L Veale, Alexander M Capron, Nima Nassiri, Gabriel Danovitch, H Albin Gritsch, Amy Waterman, Joseph Del Pizzo, Jim C Hu, Marek Pycia, Suzanne McGuire, Marian Charlton, and Sandip Kapur (2017). "Vouchers for Future Kidney Transplants to Overcome 'Chronological Incompatibility' Between Living Donors and Recipients." *Transplantation*, Online First.

R M Veatch (2006). "Organ Exchanges: Fairness to the O-Blood Group." *American Journal of Transplantation*, 6 (1), 1–2.

John von Neumann (1953). "A certain zero-sum two-person game equivalent to the optimal assignment problem." *Contributions to the Theory of Games, Vol. II*. Ed. by Harold W Kuhn and Albert W Tucker. Annals of Mathematics Studies 28. Princeton, New Jersey: Princeton University Press.

Ilse Duus Weinreich, Tommy Andersson, Margrét Birna Andrésdóttir, Mats Bengtsson, Alireza Biglarnia, Claus Bistrup, Line Boulland, Helle Bruunsgaard, Ilkka Helanterä, Kulli Kölvald, et al. (2023). "Scandiatransplant Exchange Program (STEP): Development and Results From an International Kidney Exchange Program." *Transplantation direct*, 9 (11), e1549.

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

Douglas B White, Erin K McCreary, Chung-Chou H Chang, Mark Schmidhofer, J Ryan Bariola, Naudia N Jonassaint, Govind Persad, Robert D Truog, Parag Pathak, Tayfun Sönmez, et al. (2022). "A multicenter weighted lottery to equitably allocate scarce COVID-19 therapeutics." *American Journal of Respiratory and Critical Care Medicine*, 206 (4), 503–506.

Chris Wiebe, David N Rush, Thomas E Nevins, Patricia E Birk, Tom Blydt-Hansen, Ian W Gibson, Aviva Goldberg, Julie Ho, Martin Karpinski, Denise Pochinco, et al. (2017). "Class II eplet mismatch modulates tacrolimus trough levels required to prevent donor-specific antibody development." *Journal of the American Society of Nephrology: JASN*, 28 (11), 3353.

Özgür Yilmaz (2011). "Kidney exchange: An egalitarian mechanism." *Journal of Economic Theory*, 146 (2), 592 –618.

Sezai Yilmaz, Ahmet Kizilay, Nuru Bayramov, Ahmet Tekin, and Sukru Emre (2023a). "Multiple Swaps Tested: Rehearsal for Triple and Five-Liver Paired Exchanges." *Transplantation Proceedings*, 11, S0041–1345(23)00026–X.

Sezai Yilmaz, Tayfun Sönmez, M Utku Ünver, Volkan Ince, Sami Akbulut, Burak Isik, and Sukru Emre (2023b). "The first 4-way liver paired exchange from an interdisciplinary collaboration between health care professionals and design economists." *American Journal of Transplantation*, 23 (10), 1612–1621.

Stefanos A Zenios, E Steve Woodle, and Lainie Friedman Ross (2001). "Primum Non Nocere: Avoiding Increased Waiting Times for Individual Racial and Blood-type Subsets of Kidney Wait List Candidates in a Living Donor/cadaveric Donor Exchange Program." *Transplantation*, 72, 648–654.

Congyi Zhou and Tong Wang (2021). "Purchasing Seats for High School Admission and Inequality." Available at SSRN 3579819.