# Market Design for Kidney Exchange*

Tayfun Sönmez† & M. Utku Ünver‡

*Boston College, Department of Economics, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA*

*April 25, 2011*

## 1 Introduction

The National Organ Transplant Act of 1984 makes it illegal to buy or sell a kidney in the US, thus making donation the only viable option for kidney transplantation. A transplanted kidney from a live donor survives significantly longer than one from a deceased donor (see e.g. Mandal et al. 2003). Hence, live donation is always the first choice for a patient. Moreover, there is a significant shortage of deceased donor kidneys.[1] There are two kidneys in the human body. One healthy kidney is more than enough for its healthy functioning. Since the risks associated with donation surgery and follow-up have decreased with the advancement of new medical and surgical techniques, live donation has been the increasing source of overall donations. Usually, a live donor is a relative or friend of the recipient and is willing to donate only if that particular recipient is going to receive a transplant. That is, she is a **directed live donor**. However, a recipient is often unable to receive a willing live-donor's kidney because of blood-type incompatibility or antibodies to one of the donor's proteins (aka a positive crossmatch). Medical doctor F. T. Rapaport (1986) proposed **live-donor paired kidney exchanges** between two such incompatible recipient-donor pairs: the donor in each pair gives a kidney to the other pair's compatible recipient.[2]

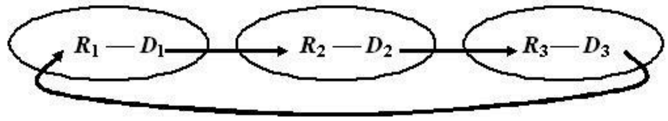
*This survey is prepared for the *Oxford Handbook of Market Design* edited by Z. Neeman, M. Niederle, and N. Vulkan.

†E-mail: sonmezt@bc.edu; URL: www2.bc.edu/~sonmezt

‡E-mail: unver@bc.edu; URL: www2.bc.edu/~unver

[1] About 79,000 patients were waiting for a deceased donor kidney transplant in the United States as of March 2009. In 2008, about 16,500 transplants were conducted, about 10,500 from deceased donors and 6,000 from living donors, while about 32,500 new patients joined the deceased donor waiting list and 4,200 patients died while waiting for a kidney (according to SRTR/OPTN national data retrieved at `http://www.optn.org` on 3/17/2009).

[2] Recently medical literature started to use the term *kidney paired donation* instead of *kidney exchange*.




A three-way kidney exchange. $R_i$ denotes the recipient and $D_i$ denotes the donor in each pair of the exchange.

In the 1990s, Korea and the Netherlands started to build databases to organize such swaps. Both programs recently reported that live-donor kidney exchanges make up more than 10% of the live-donor transplants in both countries (Park et al. 2004, de Klerk et al. 2005). Once the medical community in the US deemed the practice ethical (Abecassis et al. 2000), New England,[3] Ohio,[4] and Johns Hopkins transplant programs started conducting live-donor kidney exchange operations. The potential amount of such exchanges has been estimated to be 2,000 additional transplants per year in the US; however, it has yet to live up to expectations. The initial hurdle in organizing kidney exchanges was the lack of mechanisms to clear the market in an efficient and incentive-compatible manner. Roth, Sönmez, and Ünver (2004) proposed the first such mechanism. It was based on the core mechanism for the housing markets of Shapley and Scarf (1974), namely Gale's top trading cycles algorithm,[5] and a mechanism designed for the house allocation problem with existing tenants of Abdulkadiroğlu and Sönmez (1999), namely "you-request-my-house-I-get-your-turn" algorithm.[6] This new mechanism, called the **top-trading cycles and chains** (TTCC), is strategy-proof, i.e., it makes it the dominant strategy for recipients to reveal their preferences over compatible kidneys and all of their paired donors to the system. Moreover, it is Pareto-efficient. As the two of the coauthors of this study (Roth, Sönmez, and Ünver 2004) we showed through simulations that the potential benefits of switching to such a system would be huge.

However, one important aspect of kidney exchanges is that regardless of the number of pairs participating in an exchange, all transplants in the exchange must be conducted simultaneously. Otherwise, one or more of the live donors whose recipients receive a kidney in the previously conducted part of an exchange may back out from future donations of the same exchange. Since kidney donations are gifts, the donor can change her mind at any moment prior to the actual transplant, and it is not legal to contractually bind a donor to make future donations. This may put some recipient, whose paired-donor previously donated a kidney in the exchange, at harm.

This practice naturally places an upper limit on the number of kidney transplants that can be conducted simultaneously. The simulations showed that the TTCC mechanism may lead to large

[3] New England Program for Kidney Exchange, www.nepke.org

[4] Ohio Solid Organ Consortium, http://www.paireddonationnetwork.org

[5] See also Roth and Postlewaite (1977), Roth (1982), and Ma (1994).

[6] See also Pápai (2000), Sönmez and Ünver (2005, 2010a), Pycia and Ünver (2009), and a recent literature survey of discrete resource allocation by Sönmez and Ünver (2011).

exchanges with many recipient-donor pairs.

Another controversial issue in the market design for kidney exchange concerns the preferences of recipients over kidneys. A respected assumption in the field is that all compatible live-donor kidneys have the same likelihood of survival, following Gjertson and Cecka (2000), who statistically show this in their data set (see also Delmonico 2004).

Medical doctors also point out that if the paired-donor is compatible with the recipient, the latter will directly receive a kidney from her paired-donor and will not participate in the exchange.[7]

These institutional restrictions limit the applicability of the TTCC mechanism, which uses strict preferences information, opts in compatible pairs to the system, and results in possibly arbitrary lengths of exchange cycles. Thus, it is not immediately practical to implement this mechanism in the field.

Based on these restrictions, Roth, Sönmez, and Ünver (2005a) focused on exchanges consisting of two pairs, assuming recipients are indifferent among all compatible donors. They proposed two mechanisms, a priority mechanism and an egalitarian mechanism for strategy-proof and Pareto-efficient exchanges when recipients are indifferent among compatible donors.

The New England Program for Kidney Exchange (NEPKE) is the first US kidney exchange program that started to implement mechanisms for kidney exchange, and was established in 2004 the the collaboration of surgeon Francis Delmonico, tissue-typing expert Susan Saidman, Alvin Roth, and the authors. NEPKE started to implement a version of the priority mechanism proposed by Roth, Sönmez, and Ünver (2005a) in 2004 (see also Roth, Sönmez, and Ünver 2005b). It was followed by the Johns Hopkins Kidney Exchange Program (Segev et al. 2005), which adopted a similar algorithm due to Edmonds (1965) as proposed by Roth, Sönmez, and Ünver (2005a).

However, there was a significant gap between theory and implementation. Two-way exchanges were clearly the cornerstone of the kidney exchange paradigm. However, it was not clear what society at large was losing by restricting exchanges to two-way. Roth, Sönmez, and Ünver (2007) showed that in a large population, all the gains from exchange can be obtained by using 2&3&4-way exchanges. Especially, 2&3-way exchanges capture almost all the gains from exchange, and the marginal contribution of 3-way exchanges is significantly large. Thus, going from 2-way to 2&3-way exchanges nearly captures all the gains from exchange.

[7] This is a controversial point. Using European data, Opelz (1997) shows that indeed tissue-type matching matters even in living donations. Thus, there is no consensus in the medical community regarding tissue-type matching matters for long-term survival of live donor kidneys (other than immediate rejection). Of course, there are certain properties of donors that all authors agree as important, such as the age and health of the donor. Following the field practice of live donation established, the models and field applications surveyed here do not directly take these points into consideration, other than the ability of a recipient to report her willingness to receive or not to receive a compatible kidney.

Based on these observations, NEPKE started to implement a priority mechanism that could induce up to four-way exchanges.

In 2005, the Ohio-based Alliance for Paired Donation (APD)[8] was established through the collaboration of surgeon Michael Rees, computer programmer Jon Kopke, Alvin Roth, and the authors. This program immediately started to implement a mechanism based on maximizing the number of patients to be matched through up to four-way exchanges. It uses a priority-based solution in case there is more than one maximal matching.

The establishment of a national program for kidney exchange is in progress. The United Network for Organ Sharing (UNOS), the contractor for the national organization that maintains the deceased-donor waiting list, the Organ Procurement and Transplantation Network (OPTN), is developing this program with the consultation of economists, computer scientists, medical doctors, and administrators who have worked on the development and in the executive body of the exchange programs mentioned here and some other independent organizations. In late 2010, they launched a pilot program and two match runs have already been concluded.

In this survey, we will summarize the works of Roth, Sönmez, and Ünver (2005a, 2007), which we mentioned above, and Ünver (2010). The latter of these three studies extends the agenda of the first two papers and analyzes the kidney exchange problem as a dynamic problem in which patients arrive over time under a stochastic distribution. Then it proposes efficient mechanisms that maximize the total discounted number of patients matched under different institutional restrictions.

We will also discuss computational issues involved in solving the optimization problems with the mechanism design approach. Finally, we will talk about other paradigms in kidney exchange that are in implementation such as *list exchange*, *altruistic donor exchange*, and *altruistic donor chains*, and how these are incorporated in the market design paradigm.

## 2 Mechanics of Donation

In this section, we summarize the mechanics governing kidney donations. There are two sources of donation: *deceased donors* and *living donors.*

In the US and Europe a centralized priority mechanism is used for the allocation of deceased donor kidneys, which are considered social endowments. There have been studies regarding the effect of the choice of priority mechanism on efficiency, equity, and incentives, starting with Zenios (1995) (see also Zenios, Chertow, and Wein 2000, Votruba 2002, Su and Zenios 2006). In the US, a *soft opt-in* system is used to recruit such donors. On their drivers' licences, candidates can opt in to be deceased donors, that is they give consent to have their organs be used for transplantation

[8] www.paireddonation.org

upon their death. However, upon their death their relatives can override this decision. There are also other regimes in practice around the world, such as hard opt-in, hard opt-out, and soft opt-out.

As mentioned, live donations have been an increasing source of donations in the last decade. Live donors are generally significant others, family members, or friends of recipients. There are also some altruistic live donors who are kind enough to donate a kidney to a stranger. There is no single regulation governing live donations in the US. The only rule of thumb used is that live donors should not be coerced in donation through economic, psychological, or social pressure. In some countries, live donors are required to be blood-related or emotionally related (i.e., romantically related) to the recipient.

In this survey, we will deal with directed living donations, more specifically, the cases in which a living donor is willing to donate a kidney to a specific recipient but is incompatible with her intended recipient. We will also briefly comment on non-directed, i.e., altruistic, donations.

There are two tests that a donor must pass before she is deemed compatible with the recipient, *blood compatibility* and *tissue compatibility*, aka *crossmatch*, tests:

- *Blood compatibility test:* There are four human blood types, O, A, B, and AB. Blood type is determined by the existence or absence of one or two of the blood-type proteins called A and B. As a rule of thumb, a donor can donate a kidney to a recipient who has all the blood-type proteins that the donor possesses.[9] Thus:
  - O blood-type kidneys are blood-type compatible with all recipients;
  - A blood-type kidneys are blood-type compatible with A and AB blood-type recipients;
  - B blood-type kidneys are blood-type compatible with B and AB blood-type recipients;
  - AB blood-type kidneys are blood-type compatible with AB blood-type recipients.
- *Tissue compatibility (or crossmatch) test:* Six human leukocyte antigen (HLA) proteins on DNA determine tissue type. There does not need to be a 100% match of the HLA proteins between the donor and the recipient for tissue compatibility. If antibodies form in the blood of the recipient against the donor's tissue types, then there is **tissue rejection** (or **positive crossmatch**), and the donor is tissue-type incompatible with the recipient. The reported chance of positive crossmatch in the literature is around 11% between a random blood-type compatible donor and a random recipient (Zenios, Woodle, and Ross 2001).

If either test fails, the donation cannot go forward. We refer to such a pair as **incompatible**. This pair then becomes available for paired kidney exchange, which is the topic of the rest of the survey.

[9] O type is referred to as 0 (zero) in many languages, and it refers to non-existence of any blood-type proteins.

# 3 A Model of Kidney Exchanges

Let $N$ be the set of groups of incompatible donors and their recipients, that is: each $i = (R_i, \{D_i^1, ..., D_i^{n_i}\}) \in N$ is a *group* (if $n_i = 1$, a **pair**) and is represented by a **recipient** $R_i$ and her paired incompatible **donors** $D_i^1, ..., D_i^{n_i}$. We permit each recipient to have more than one incompatible donor. However, only one of these donors will donate a kidney if, and only if, the recipient receives one. We will sometimes refer to $i$ simply as a **recipient**, since we treat the donors through their kidneys, which are objects, and consider the recipients as the decision makers, i.e. agents.

For each $i \in N$, let $\succsim_i$ be a preference relation on $N$ with three indifference classes. Option $j \in N \setminus \{i\}$ refers to the recipient $i$ receiving a kidney from the best donor of $j$ for $i$. Option $i$ refers to remaining unmatched. Let $\succ_i$ be the acyclic (i.e., strict preference) portion of $\succsim_i$ and $\sim_i$ be the cyclic (i.e., indifference) portion of $\succsim_i$. For any $j, k \in N \setminus \{i\}$, we have

- $j \succ_i i$ if at least one donor of $j$ is compatible with $i$;
- $j \sim_i k$ if at least one donor of each of $j$ and $k$ is compatible with $i$;
- $i \succ_i j$ if all donors of $j$ are incompatible with $i$; and
- $j \sim_i k$ if all donors of $j$ and $k$ are incompatible with $i$.

That is, a recipient with a compatible donor is preferred by $i$ to remaining unmatched, which is, in turn, preferred to a recipient with incompatible donors. All recipients with only incompatible donors are indifferent for $i$. Similarly, all recipients each with at least one compatible donor are indifferent for $i$.

A **problem** is denoted by the recipients, their donors and preferences. An outcome of a problem is a *matching*. A **matching** $\mu : N \to N$ is a one-to-one and onto mapping. For each $i \in N$, recipient $i_1$ receives a kidney from some donor of recipient $\mu(i)$. We do not specify which donor in our notation, since at most one donor of a recipient is going make a donation in any matching. Thus, for our purposes $i$ can be matched with any compatible donor of $\mu(i)$. A matching $\mu$ is **individually rational** if for all recipients $i \in N$, $\mu(i) \succsim_i i$. We will focus on only individually rational matchings. Thus, when we say a matching it will be individually rational from now on. Let $\mathcal{M}$ be the set of matchings. A $k-$**way exchange** for some $k \geq 1$ is a list $(i_1, i_2, ..., i_k)$ such that $i_1$ receives a kidney from a compatible donor of $i_k$, $i_2$ receives a kidney from a compatible donor of $i_1$, ..., and $i_k$ receives a kidney from a compatible donor of $i_{k-1}$. Similarly, all exchanges we will consider will be individually rational. A degenerate exchange $(i)$ denotes the case in which recipient $i$ is unmatched. Alternatively, we represent a matching $\mu$ as a set of exchanges such that each recipient participates in one and only one exchange.

Besides deterministic outcomes, we will also define stochastic outcomes. A stochastic outcome is a **lottery** $\lambda = (\lambda_\mu)_{\mu\in\mathcal{M}}$ that is a probability distribution on all matchings. Although in many matching problems, there is no natural definition of von Neumann - Morgenstern utility functions, there is one for this problem: It takes value 1 if the recipient is matched and 0 otherwise. We can define the (**expected**) **utility** of the recipient of a pair $i$ under a lottery $\lambda$ as the probability of the recipient getting a transplant and we denote it by $u_i(\lambda)$. The **utility profile** of lottery $\lambda$ is denoted by $u(\lambda) = (u_i(\lambda))_{i\in N}$.

A matching is **Pareto-efficient** if there is no other matching that makes every recipient weakly better off and some recipient strictly better off. A lottery is **ex-post efficient** if it gives positive weight to only Pareto-efficient matchings. A lottery is **ex-ante efficient** if there is no other lottery that makes every recipient weakly better off and some recipient strictly better off.

A **mechanism** is a systematic procedure that assigns a lottery for each problem.

A mechanism is **strategy-proof** if for each problem $(N, \succsim)$, it is a dominant strategy for each pair $i$

- to report its true preference $\succsim_i$in a preference profile set $\mathcal{P}(\succsim_i)$ where for all $\succsim_i' \in \mathcal{P}(\succsim_i)$, $j \succ_i' i \Longrightarrow j \succ_i i$, i.e. one pair can never report a group with only incompatible donors as compatible; and

- to report full set of incompatible donors to the problem.

The first bullet point above underlines the fact that it is possible to detect incompatible donors through blood tests, thus, we will assume that no recipient can reveal an incompatible donor to be compatible. On the other hand, some idiosyncratic factors can lead a recipient to reveal compatible donors to be incompatible.

We will survey different Pareto-efficient and strategy-proof mechanisms for different institutional constraints.

# 4 Two-way Kidney Exchanges

First, we restrict our attention in this section to individually rational two-way exchanges. This section follows Roth, Sönmez, and Ünver (2005a). Formally, for given any problem $(N, \succsim)$, we are interested in matchings $\mu \in \mathcal{M}$ such that for all $i \in N$, $\mu(\mu(i)) = i$. To make our notation simpler, we define the following concept: Recipients $i, j$ are **mutually compatible** if $j$ has a compatible donor for $i$ and $i$ has a compatible donor for $j$. We can focus on a *mutual compatibility matrix* that summarizes the feasible exchanges and preferences. A **mutual compatibility matrix** $C = [c_{i,j}]_{i\in N, j\in N}$ is defined as for any $i, j \in N$,

$$c_{i,j} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are mutually compatible} \\ 0 & \text{otherwise} \end{cases}.$$

The induced **two-way kidney exchange problem** from problem $(N, \succsim)$ is denoted by $(N, C)$. A **subproblem** of $(N, C)$ is denoted as $(I, C_I)$ where $I \subseteq N$ and $C_I$ is the restriction of $C$ to the pairs in $I$. Thus, all relevant information regarding preferences is summarized by the mutual compatibility matrix $C$.

Observe that a problem $(N, C)$ can be represented by an undirected *graph* in which each recipient is a *node*, and there is an *edge* between two nodes if and only if these two recipients are mutually compatible. Hence, we define the following graph-theoretic concepts for two-way kidney exchange problems:

A problem is **connected** if the corresponding graph of the problem is connected, i.e., one can traverse between any two nodes of the graph using the edges of the graph. A **component** is a largest connected subproblem. We refer to a component as **odd** if it has an odd number of recipients, and as **even** if it has an even number of recipients.

Although in many matching domains ex-ante and ex-post efficiency are not equivalent (see e.g. Bogomolnaia and Moulin 2001), they are equivalent for two-way kidney exchanges with 0-1 preferences because of the following lemma:

**Lemma 1** ***(Roth, Sönmez, and Ünver 2005a)*** *The same number of recipients is matched at each Pareto-efficient matching, which is the maximum number of recipients that can be matched.*

Thus, finding a Pareto-efficient matching is equivalent to finding a matching that matches the maximum number of recipients. In graph theory, such a problem is known as a *cardinality matching problem* (see e.g. Lóvasz and Plummer 1986 for an excellent survey of this and other matching problems regarding graphs). Various intuitive polynomial time algorithms are known to find one Pareto-efficient matching (starting with Edmonds' 1965 algorithm).

The above lemma would not hold if exchange were possible among three or more recipients. Moreover, we can state the following lemma regarding efficient lotteries:

**Lemma 2** ***(Roth, Sönmez, and Ünver 2005a)*** *A lottery is ex-ante efficient if and only it is ex-post efficient.*

There are many Pareto-efficient matchings, and finding all of them is not computationally feasible (i.e., it is NP-complete). Therefore, we will focus on two selections of Pareto-efficient matchings and lotteries that have nice fairness features.

## 4.1 Priority Mechanism

In many situations, recipients may be ordered by a natural priority ordering. For example, the sensitivity of a recipient to the tissue types of others, known as panel reactive antibody (PRA), is a criterion also accepted by medical doctors. Some recipients may be sensitive to almost all tissue types other than their own and have a PRA=99%, meaning that they will reject based solely on tissue incompatibility 99% of donors from a random sample. So, one can order the recipients from high to low with respect to their PRA's and use the following *priority mechanism*:

Given a priority ordering of recipients, a **priority mechanism**
matches **Priority 1** recipient if she is mutually compatible with a recipient, and skips her otherwise.

$$\vdots$$

matches **Priority k** recipient in addition to all the previously matched recipients if possible, and skips her otherwise.

Thus, the mechanism determines which recipients are to be matched first, and then one can select a Pareto-efficient matching that matches those recipients. Thus, the mechanism is only unique-valued for the utility profile induced. Any matching inducing this utility profile can be the final outcome. The following result makes a priority mechanism very appealing:

**Theorem 1** *A two-way priority mechanism is Pareto-efficient and strategy-proof.*

## 4.2 The Structure of Pareto-Efficient Matchings

We can determine additional properties Pareto-efficient matchings (even though finding all such matchings is exhaustive and hence, NP-complete) thanks to the results of Gallai (1963, 1964) and Edmonds (1965) in graph theory. We can partition the recipients into three sets as $N^U, N^O, N^P$. The members of these sets are defined as follows:

An **underdemanded recipient** is one for whom there exists a Pareto-efficient matching that leaves her unmatched. Set $N^U$ is formed by underdemanded recipients, and we will refer to this set as the set of underdemanded recipients. An **overdemanded recipient** is one who is not underdemanded, yet is mutually compatible with an underdemanded recipient. Set $N^O$ is formed by overdemanded recipients. A **perfectly matched recipient** is one who is neither underdemanded nor mutually compatible with any underdemanded recipient. Set $N^P$ is formed by perfectly matched recipients.

The following result, due to Gallai and Edmonds, is the key to understanding the structure of Pareto-efficient matchings:

**Lemma 3** ***Gallai (1963,1964)-Edmonds (1965) Decomposition (GED)****: Let $\mu$ be* any *Pareto-efficient matching for the original problem $(N, C)$ and $(I, C_I)$ be the* ***subproblem*** *for $I = N \setminus N^O$. Then we have:*

1. *Each overdemanded recipient is matched with an underdemanded recipient under $\mu$.*
2. *$J \subseteq N^P$ for any* even component *$J$ of the subproblem $(I, C_I)$ and all recipients in $J$ are matched with each other under $\mu$.*
3. *$J \subseteq N^U$ for any* odd component *$J$ of the subproblem $(I, C_I)$ and for any recipient $i \in J$, it is possible to match all remaining recipients with each other under $\mu$. Moreover, under $\mu$*
   - *either one recipient in $J$ is matched with an overdemanded recipient and all others are matched with each other,*

     *or*
   - *one recipient in $J$ remains unmatched while the others are matched with each other.*

We can interpret this lemma as follows: There exists a competition among odd components of the subproblem $(I, C_I)$ for overdemanded recipients. Let $\mathcal{O} = \{O_1, \ldots, O_p\}$ be the set of odd components remaining in the problem when overdemanded recipients are removed. By the GED Lemma, all recipients in each odd component are matched but at most one, and all of the other recipients are matched under each Pareto-efficient matching. Thus, such a matching leaves $|\mathcal{O}| - |N^O|$ unmatched recipients each of whom is in a distinct odd component.

First, suppose that we determine the set of overdemanded recipients, $N^O$. After removing those from the problem, we mark the recipients in odd components as *underdemanded*, and recipients in even components as *perfectly matched*. Moreover, we can think of each odd component as a single entity, which is competing to get one overdemanded recipient for its recipients under a Pareto-efficient matching.

It turns out that the sets $N^U, N^O, N^P$ and the GED decomposition can also be found in polynomial time thanks to Edmonds' algorithm and related results in the literature.

## 4.3 Egalitarian Mechanism

Recall that the utility for a recipient under a lottery is the probability of receiving a transplant. Equalizing utilities as much as possible may be considered very desirable from an equity perspective, which is also in line with the Rawlsian notion of fairness (Rawls 1971). We define a central notion in Rawlsian egalitarianism:

A feasible utility profile is **Lorenz-dominant** if

- the least fortunate recipient receives the highest utility among all feasible utility profiles, and

  $\vdots$

- the sum of utilities of the $k$ least fortunate recipients is the highest among all feasible utility profiles.[10]

Is there a feasible Lorenz-dominant utility profile? Roth, Sönmez, and Ünver (2005a) answer this question affirmatively. This utility profile is constructed with the help of the GED of the problem. Let

- $\mathcal{J} \subseteq \mathcal{O}$ be an arbitrary set of odd components of the subproblem obtained by removing the overdemanded recipients,
- $I \subseteq N^O$ be an arbitrary set of overdemanded recipients, and
- $N(\mathcal{J}, I) \subseteq I$ denote the **neighbors** of $\mathcal{J}$ among $I$, that is, each overdemanded recipient in $N(\mathcal{J}, I)$ is in $I$ and is mutually compatible with a recipient in an odd component of the collection $\mathcal{J}$.

Suppose only overdemanded recipients in $I$ are available to be matched with underdemanded recipients in $\bigcup_{J \in \mathcal{J}} J$. Then, what is the upper bound of the utility that can be received by the *least fortunate* recipient in $\bigcup_{J \in \mathcal{J}} J$? The answer is

$$f(\mathcal{J}, I) = \frac{\left|\bigcup_{J \in \mathcal{J}} J\right| - \left(|\mathcal{J}| - |N(\mathcal{J}, I)|\right)}{\left|\bigcup_{J \in \mathcal{J}} J\right|}$$

and it can be received only if

1. all underdemanded recipients in $\bigcup_{J \in \mathcal{J}} J$ receive the same utility, and
2. all overdemanded recipients in $N(\mathcal{J}, I)$ are committed for recipients in $\bigcup_{J \in \mathcal{J}} J$.

The function $f$ is the key in constructing an **egalitarian utility profile** $u^E$. The following procedure can be used to construct it:

Partition $\mathcal{O}$ as $\mathcal{O}_1, \mathcal{O}_2, \ldots$ and $N^O$ as $N_1^O, N_2^O, \ldots$ as follows:

[10] By *k least fortunate recipients* under a utility profile, we refer to the $k$ recipients whose utilities are lowest in this utility profile.

**Step 1**.

$$\mathcal{O}_1 = \arg\min_{\mathcal{J}\subseteq\mathcal{O}} f\left(\mathcal{J}, N^O\right) \text{ and}$$
$$N_1^O = N\left(\mathcal{O}_1, N^O\right)$$

$$\vdots$$

**Step k.**

$$\mathcal{O}_k = \arg\min_{\mathcal{J}\subseteq\mathcal{O}\setminus\bigcup_{\ell=1}^{k-1}\mathcal{O}_\ell} f\left(\mathcal{J}, N^O \Big\backslash \bigcup_{\ell=1}^{k-1} N_\ell^O\right) \text{ and}$$
$$N_k^O = N\left(\mathcal{O}_k, N^O \Big\backslash \bigcup_{\ell=1}^{k-1} N_\ell^O\right)$$

Construct the vector $u^E = (u_i^E)_{i\in N}$ as follows:

1. For any overdemanded recipient and perfectly matched recipient $i \in N \setminus N^U$,
$$u_i^E = 1.$$

2. For any underdemanded recipient $i$ whose odd component left the above procedure at Step $k(i)$,
$$u_i^E = f(\mathcal{O}_{k(i)}, N_{k(i)}^O).$$

We provide an example explaining this construction:

**Example 1** Let $N = \{1, \ldots, 16\}$ be the set of recipients and let the reduced problem be given by the graph in Figure 1. $N^U = \{3, \ldots, 16\}$ is the set of underdemanded recipients. Since both recipients 1 and 2 have edges with recipients in $N^U$, $N^O = \{1, 2\}$ is the set of overdemanded recipients.

$$\mathcal{O} = \{O_1, \ldots, O_6\}$$

where

$$O_1 = \{3\}, O_2 = \{4\}, O_3 = \{5\}, O_4 = \{6, 7, 8\}$$
$$O_5 = \{9, 10, 11\}, O_6 = \{12, 13, 14, 15, 16\}$$

Consider $J_1 = \{O_1, O_2\} = \{\{3\}, \{4\}\}$. Note that by the GED Lemma, an odd component that has $k$ recipients guarantees $\frac{k-1}{k}$ utility for each of its recipients. Since $f(J_1, N^O) = \frac{1}{2} < \frac{2}{3} < \frac{4}{5}$, none of the

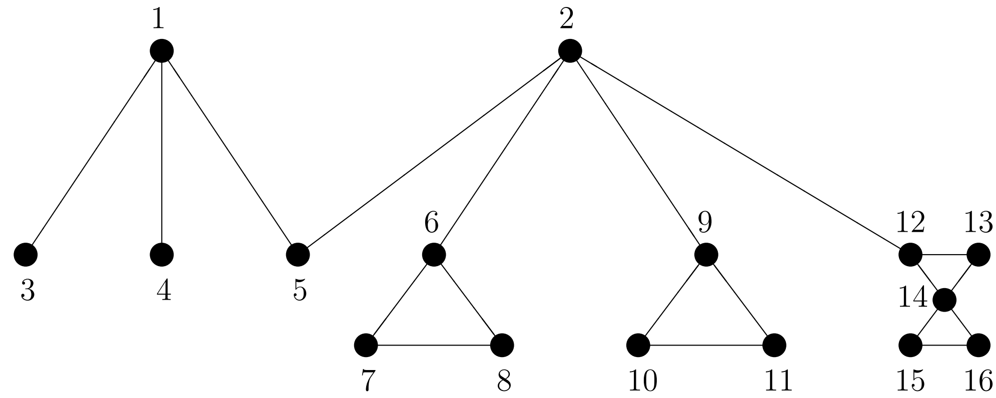



Figure 1: Graphical Representation for Example 1.

multi-recipient odd components is an element of $O_1$. Moreover, recipient 5 has two overdemanded neighbors and $f(J, N^O) > f(J_1, N^O)$ for any $J \subseteq \{\{3\},\{4\},\{5\}\}$ with $\{5\} \in J$. Therefore

$$\mathcal{O}_1 = \mathcal{J}_1 = \{\{3\},\{4\}\}, \quad N_1^O = \{1\},$$
$$u_3^E = u_4^E = \frac{1}{2}.$$

Next consider $J_2 = \{O_3, O_4, O_5\} = \{\{5\},\{6,7,8\},\{9,10,11\}\}$. Note that $f(J_2, N^O \setminus N_1^O) = \frac{7-(3-1)}{7} = \frac{5}{7}$. Since $f(J_2, N^O \setminus N_1^O) = \frac{5}{7} < \frac{4}{5}$, the 5-recipient odd component $O_6$ is not an element of $O_2$. Moreover,

$$\begin{aligned} f(\{O_3\}, N^O \setminus N_1^O) &= f(\{O_4\}, N^O \setminus N_1^O) \\ &= f(\{O_5\}, N^O \setminus N_1^O) = 1, \\ f(\{O_3, O_4\}, N^O \setminus N_1^O) &= f(\{O_3, O_5\}, N^O \setminus N_1^O) = \frac{3}{4}, \\ f(\{O_4, O_5\}, N^O \setminus N_1^O) &= \frac{5}{6}. \end{aligned}$$

Therefore,

$$\begin{aligned} &\mathcal{O}_2 = \mathcal{J}_2 = \{\{5\},\{6,7,8\},\{9,10,11\}\}, \\ &N_2^O = \{2\}, \\ \text{and} \quad &u_5^E = \cdots = u_{11}^E = \frac{5}{7}. \end{aligned}$$

Finally since $N^O \setminus (N_1^O \cup N_2^O) = \emptyset$,

$$\begin{aligned} &\mathcal{O}_3 = \{\{12,13,14,15,16\}\}, \\ &N_3^O = \emptyset, \\ \text{and} \quad &u_{12}^E = \cdots = u_{16}^E = \frac{4}{5}. \end{aligned}$$

Hence the egalitarian utility profile is

$$u^E = (1, 1, \frac{1}{2}, \frac{1}{2}, \frac{5}{7}, \frac{5}{7}, \frac{5}{7}, \frac{5}{7}, \frac{5}{7}, \frac{5}{7}, \frac{5}{7}, \frac{4}{5}, \frac{4}{5}, \frac{4}{5}, \frac{4}{5}, \frac{4}{5}).$$

$\triangledown$

Roth, Sönmez, and Ünver (2005a) proved the following results:

**Theorem 2** ***(Roth, Sönmez, and Ünver 2005a)*** *The vector $u^E$ is a feasible utility profile.*

In particular, the proof of Theorem 2 shows how a lottery that implements $u^E$ can be constructed.

**Theorem 3** ***(Roth, Sönmez, and Ünver 2005a)*** *The utility profile $u^E$ Lorenz-dominates any other feasible utility profile (efficient or not).*

The egalitarian mechanism is a lottery mechanism that selects a lottery whose utility profile is $u^E$. It is only unique-valued for the utility profile induced. As a mechanism, the egalitarian approach also has appealing properties:

**Theorem 4** ***(Roth, Sönmez, and Ünver 2005a)*** *The egalitarian mechanism is ex–ante efficient and strategy-proof.*

The egalitarian mechanism can be used for cases in which there is no exogenous way to distinguish among recipients.

The related literature for this section includes four other papers, two of which are by Bogomolnaia and Moulin (2004), who inspect a two-sided matching problem with the same setup as the model above, and by Dutta and Ray (1989), who introduce the egalitarian approach for convex TU-cooperative games. Morrill (2008) inspects a model similar to the one surveyed here for two-way exchanges, with the exception that preferences are strict. He considers Pareto-efficient matchings and proposes a polynomial time algorithm for finding one starting from a status-quo matching (see Section 7). Yilmaz (2011) considers an egalitarian kidney exchange mechanism when multi-way list exchanges (Section 7) are possible. He considers a hybrid model between Roth, Sönmez, and Ünver (2004) and (2005a).

# 5 Multi-Way Kidney Exchanges

Roth, Sönmez, and Ünver (2007) explored what is lost when the central authority conducts only two-way kidney exchanges rather than multi-way exchanges. More specifically, they examined the upper bound of marginal gains from conducting 2&3-way exchanges instead of only two-way exchanges, 2&3&4-way exchanges instead of only 2&3-way exchanges, and unrestricted multi-way exchanges instead of only 2&3&4-way exchanges. The setup is very similar to that given in the previous section with only one difference: a matching does not necessarily consist of two-way exchanges. All results in this section are due to Roth, Sönmez, and Ünver (2007) unless otherwise noted.

In this section, a recipient will be assumed to have a single incompatible donor, and thus, the recipient and her incompatible donor will be referred to as a pair. The blood types of the recipient $R_i$ and donor $D_i$ are denoted as X-Y for pair $i$, where the recipient is of blood type X and donor is of blood type Y.

An example helps illustrate why the possibility of a three-way exchange is important:

**Example 2** Consider a sample of 14 incompatible recipient-donor pairs. There are nine pairs, who are blood-type incompatible, of types A-AB, B-AB, O-A, O-A, O-B, A-B, A-B, A-B, and B-A; and five pairs, who are incompatible because of tissue rejection, of types A-A, A-A, A-A, B-O, and AB-O. For simplicity in this example there is no tissue rejection between recipients and other recipients' donors.

- If only two-way exchanges are possible:

  (A-B,B-A); (A-A,A-A); (B-O,O-B); (AB-O,A-AB) is a possible Pareto-efficient matching.

- If three-way exchanges are also feasible:

  (A-B,B-A); (A-A,A-A,A-A); (B-O,O-A,A-B); (AB-O, O-A, A-AB) is a possible maximal Pareto-efficient matching.

The 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), and

2. a pair with a donor who has a blood type more desirable than her recipient's to facilitate three transplants rather than only two. Here, the AB-O type pair helps two pairs with recipients having less desirable blood type than their donors (O-A and A-AB), while the B-O type pair helps one pair with a recipient having a less desirable blood type than her donor (O-A) and a pair of type A-B. Here, note that another A-B type pair is already matched with a B-A type, and this second A-B type pair is in excess. $\triangledown$

First, we introduce two upper-bound assumptions and find the size of Pareto-efficient exchanges with only two-way exchanges:

**Assumption 1** ***(Upper Bound Assumption)*** *No recipient is tissue-type incompatible with another recipient's donor.*

**Assumption 2** ***(Large Population of Incompatible Recipient-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 feasible set of exchanges. We simply assume there is an arbitrarily many number of O-A, O-B, O-AB, A-AB, and B-AB type pairs.*

The following observations concern the feasibility of exchanges:

**Observation 1** *A pair of type* $X$*-*$Y \in \{$*O-A, O-B, O-AB, A-AB, B-AB*$\}$ *can participate in a two-way exchange only with a pair of its reciprocal type* $Y$*-*$X$ *or type AB-O.*

**Observation 2** *A pair of O-O, A-A, B-B, AB-AB, A-B or B-A can participate in a two-way exchange only with its reciprocal type pair or a pair belonging to some of the types among A-O, B-O, AB-O, AB-A, AB-B.*

**Observation 3** *A pair of type* $X$*-*$Y \in \{$*A-O, B-O, AB-O, AB-A, AB-B*$\}$ *can participate in a two-way exchange with a pair of not only its own type (and possibly some other types in the same set), but also some types among O-A, O-B, O-AB, A-AB, B-AB, O-O, A-A, B-B, AB-AB, A-B, B-A, as well.*

Based on the above observations and the intuition given in Example 2, we formally classify the types of pairs into 4 (Ünver 2010):

**Overdemanded** types: $\mathcal{T}^O = \{$A-O, B-O, AB-O, AB-A, AB-B$\}$

**Underdemanded** types: $\mathcal{T}^U = \{$O-A, O-B, O-AB, A-AB, B-AB$\}$

**Self-demanded** types: $\mathcal{T}^S = \{$O-O, A-A, B-B, AB-AB$\}$

**Reciprocally demanded** types: $\mathcal{T}^R = \{$A-B, B-A$\}$

Observe that the definitions of overdemanded and underdemanded types in this chapter are different from their definitions used in Section 4 for the GED Lemma. We will use these definitions in the next two sections as well. Both definitions are in the same flavor, yet they are not equivalent.

The first result is about the greatest lower bound of the size of two-way Pareto-efficient matchings:

**Proposition 1** ***(Roth, Sönmez, and Ünver 2007) The Maximal Size of Two-Way Matchings:*** *For any recipient population obeying Assumptions 1 and 2, the maximum number of recipients who can be matched with only two-way exchanges is:*

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

*where* $\lfloor a \rfloor$ *refers to the largest integer smaller than or equal to* $a$ *and* $\#(x\text{-}y)$ *refers to the number of x-y type pairs.*

We can generalize Example 2 in a proposition for three-way exchanges. We introduce an additional assumption for ease of notation. The symmetric case implies replacing types "A" with "B" and "B" with "A" in all of the following results.

**Assumption 3** $\#(A\text{-}B) > \#(B\text{-}A)$.

The following is a simplifying assumption.

**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.*

When three-way exchanges are also feasible, as we noted earlier, Lemma 1 no longer holds. Thus, we consider the largest of the Pareto-efficient matchings under 2&3-way matching technology.

On the other hand, an overdemanded AB-O type pair can potentially save two underdemanded type pairs of types O-A and A-AB or O-B and B-AB under a three-way exchange (see Figure 2).

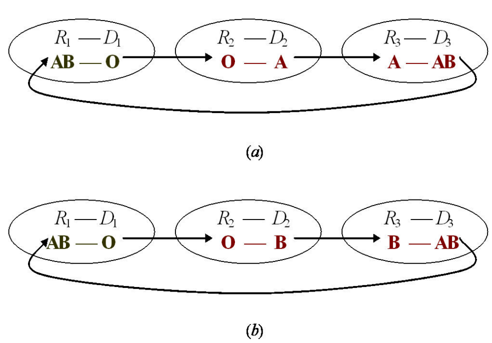



Figure 2: AB-O type pair saving 2 underdemanded pairs in a 3-way exchange.

When the number of A-B type pairs is larger than the number of B-A type pairs in a static pool (Assumption 3):

- All B-A type pairs can be matched with A-B type pairs in two-way exchanges.
- Each B-O type pair can potentially save one O-A type pair and one *excess* A-B type pair in a three-way exchange.
- Each AB-A type pair can potentially save one *excess* A-B type and one B-AB type pair in a three-way exchange (see Figure 3).

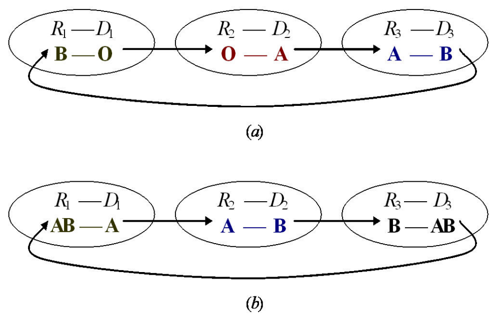



Figure 3: Overdemanded pairs B-O / AB-A each saving one underdemanded pair and an A-B type pair in a 3-way exchange.

The above intuition can be stated as a formal result:

**Proposition 2** ***(Roth, Sönmez, and Ünver 2007) The Maximal Size of 2&3-Way Matchings:*** *For any recipient population for which Assumptions 1-4 hold, the maximum number of recipients who can be matched with two-way and three-way exchanges is:*

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

*And to summarize, the marginal effect of availability of 2&3-way kidney exchanges over two-way exchanges is:*

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

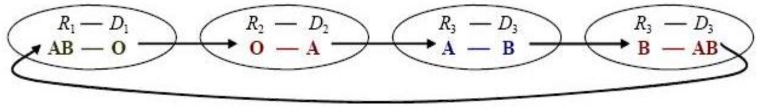



Figure 4: An overdemanded AB-O type pair can save 3 underdemanded pairs in a four-way kidney exchange.

What about the marginal effect of 2&3&4-way exchanges over 2&3-way exchanges? It turns out that there is only a slight improvement in the maximal matching size with the possibility of four-way exchanges.

We illustrate this using the above example:

**Example 3** (Example 2 Continued) If four-way exchanges are also feasible, instead of the exchange (AB-O, O-A, A-AB) we can now conduct a four-way exchange (AB-O, O-A, A-B, B-AB). Here, the valuable AB-O type pair helps an additional A-B type pair in excess in addition to two pairs with less desirable blood-type donors than their recipients. ▽

Thus, each AB-O type pair can potentially save one O-A type pair, one *excess* A-B type pair, and one B-AB type pair in a four-way exchange (See Figure 4).

We formalize this intuition as the following result:

**Proposition 3** ***(Roth, Sönmez, and Ünver 2007) The Maximal Size of 2&3&4-Way Matchings:*** *For any recipient population in which Assumptions 1-4 hold, the maximum number of recipients who can be matched with two-way, three-way, and four-way exchanges is:*

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

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

It turns out that under the above assumptions, larger exchanges do not help to match more recipients. This is stated as follows:

**Theorem 5** ***(Roth, Sönmez, and Ünver 2007) Availability of Four-Way Exchange Suffices:*** *Consider a recipient population for which Assumptions 1, 2, 4 hold and let $\mu$ be any maximal matching (when there is no restriction on the size of the exchanges). Then there exists a maximal matching $\nu$ that consists only of two-way, three-way, and four-way exchanges, under which the same set of recipients benefits from exchange as in matching $\mu$.*

What about incentives, when these maximal solution concepts are adopted in a kidney exchange mechanism? The strategic properties of multi-way kidney exchange mechanisms are inspected by Hatfield (2005) in the 0-1 preference domain. This result is a generalization of Theorem 1.

A deterministic kidney exchange mechanism is **consistent** if whenever it only selects a multi-way matching in set $\mathcal{X} \subseteq \mathcal{M}$ as its outcome, where all matchings in $\mathcal{X}$ generate the same utility profile when the set of feasible individually rational matchings is $\mathcal{M}$, then for any other problem for the same set of pairs such that the set of feasible individually rational matchings is $\mathcal{N} \subset \mathcal{M}$ with $\mathcal{X} \cap \mathcal{N} \neq \varnothing$, it selects a multi-way matching in set $\mathcal{X} \cap \mathcal{N}$.[11]

A deterministic mechanism is **non-bossy** if whenever one recipient manipulates her preferences/number of donors and cannot change her outcome, defined as either being matched to a compatible donor or remaining unmatched, then she cannot change other recipients' outcome under this mechanism with the same manipulation.

The last result of this section is as follows:

**Theorem 6** ***(Hatfield 2005):*** *If a deterministic mechanism is non-bossy and strategy-proof then it is consistent. Moreover, a consistent mechanism is strategy-proof.*

Thus, it is straightforward to create strategy-proof mechanisms using *maximal-priority* or *priority* multi-way exchange rules. By maximal-priority mechanisms, we mean mechanisms that maximize the number of patients matched (under an exchange restriction such as 2, 3, 4, etc., or no exchange size restriction) and then use a priority criterion to select among such matchings.

# 6 Simulations Using National Recipient Characteristics

In this section we dispense with the simplifying assumptions made so far, and turn to simulated data reflecting national recipient characteristics. Specifically, we now look at populations in which a recipient may have tissue type incompatibilities with many donors. This will allow us to assess the accuracy of the approximations derived under the above assumption that exchange is limited only by blood-type incompatibilities.

[11] Recall that a kidney exchange mechanism may select many matchings that are utility-wise equivalent in the 0-1 preference domain. A two-way priority mechanism is an example.

The simulations reported here follow those of Saidman et. al. (2006) and Roth, Sönmez, and Ünver (2007). We will see that the formulas predict the actual number of exchanges surprisingly well. That is, the upper bounds on the maximal number of exchanges when exchange is limited only by blood-type incompatibility are not far above the numbers of exchanges that can actually be realized. In addition, only a small number of exchanges involving more than four pairs are needed to achieve efficiency in the simulated data.

## 6.1 Recipient-Donor Population Construction

We consider samples of non-blood-related recipient-donor pairs to avoid complications due to the impact of genetics on immunological incompatibilities. The characteristics such as the blood-types of recipients and donors, the PRA distribution of the recipients, donor relation of recipients, and the gender of the recipients are generated using the empirical distributions of the data from an OPTN subsidiary in the US, the Scientific Registry of Transplant Recipients (SRTR) (see Table 1). We consider all ethnicity in the data.

| **A. Patient ABO Blood Type** | **Frequency (percent)** |
|---|---|
| O | 48.14 |
| A | 33.73 |
| B | 14.28 |
| AB | 3.85 |
| **B. Patient Gender** | **Frequency (percent)** |
| Female | 40.90 |
| Male | 59.10 |
| **C. Unrelated Living Donors** | **Frequency (percent)** |
| Spouse | 48.97 |
| Other | 51.03 |
| **E. PRA Distribution** | **Frequency (percent)** |
| Low PRA | 70.19 |
| Medium PRA | 20.00 |
| High PRA | 9.81 |

Table 1: Patient and living donor distributions used in simulations based on OPTN/SRTR Annual Report in 2003, for the period 1993-2002, retrieved from `http://www.optn.org` on 11/22/2004. Patient characteristics are obtained using the new waiting list registrations data, and living donor relational type distribution is obtained from living donor transplants data.

In our simulations, we randomly simulate a series of recipient-donor pairs using the population characteristics explained above. Whenever a pair is compatible (both blood-type compatible and tissue-type compatible), the donor can directly donate to the intended recipient and therefore we do not include them in our sample. Only when they are either blood-type or tissue-type incompatible do we keep them, until we reach a sample size of $n$ incompatible pairs. We use a Monte-Carlo simulation size of 500 random population constructions for three population sizes of 25, 50, and 100.

## 6.2 Tissue-Type Incompatibility

Tissue-type incompatibility (a **positive crossmatch**) is independent of blood-type incompatibility and arises when a recipient has preformed antibodies against a donor tissue-type.

Recipients in the OPTN/SRTR database are divided into the following three groups based on the odds that they have a crossmatch with a random donor:

1. Low PRA (Percent Reactive Antibody) recipients: Recipients who have a positive crossmatch with less than 10 percent of the population.

2. Medium PRA recipients: Recipients who have a positive crossmatch with 10-80 percent of the population.

3. High PRA recipients: Recipients who have a positive crossmatch with more than 80 percent of the population.

Frequencies of low, medium, and high PRA recipients reported in the OPTN/SRTR database are given in Table 1. Since a more detailed PRA distribution is unavailable in the medical literature, we will simply assume that:

- each low PRA recipient has a positive crossmatch probability of 5 percent with a random donor,

- each medium PRA recipient has a positive crossmatch probability of 45 percent with a random donor, and

- each high PRA recipient has a positive crossmatch probability of 90 percent with a random donor.

We have already indicated that when the recipient is female and the potential donor is her husband, it is more likely that they have a positive crossmatch due to pregnancies. Zenios, Woodle, and Ross (2001) indicate that while positive crossmatch probability is 11.1 percent between random pairs, it is 33.3 percent between female recipients and their donor husbands. Equivalently, female recipients' **negative crossmatch** probability (i.e. the odds that there is no tissue-type incompatibility) with their husbands is approximately 75 percent of the negative crossmatch probability with a random donor. Therefore, we accordingly adjust the positive crossmatch probability between a female recipient and her donor husband using the formula

$$\text{PRA}^* = 100 - 0.75(100 - \text{PRA})$$

and assume that

- each low PRA female recipient has a positive crossmatch probability of 28.75 percent with her husband,
- each medium PRA female recipient has a positive crossmatch probability of 58.75 percent with her husband, and
- each high PRA female recipient has a positive crossmatch probability of 92.25 percent with her husband.

## 6.3 Outline of the Simulations

For each sample of $n$ incompatible recipient-donor pairs, we find the maximum number of recipients who can benefit from an exchange when both blood-type and tissue-type incompatibilities are considered, and

a. only two-way exchanges are allowed,

b. two-way and three-way exchanges are allowed,

c. two-way, three-way, and four-way exchanges are allowed, and

d. any size exchange is allowed.

In our simulations, to find the maximal number of recipients who can benefit from an exchange when only two-way exchanges are allowed, we use a version of Edmonds' (1965) algorithm (see Roth, Sönmez, and Ünver 2005a), and to find the maximal number of recipients who can benefit from an exchange when larger exchanges are allowed, we use various integer programming techniques.

We compare these numbers with those implied by the analytical expressions in the above propositions, to see whether those formulas are close approximations or merely crude upper-bounds. Since many high PRA recipients cannot be part of any exchange due to tissue-type incompatibilities, we report two sets of upper-bounds induced by the formulas we developed:

1. For each sample we use the formulas with the raw data, and
2. for each sample we restrict our attention to recipients each of whom can participate in at least one feasible exchange.

That is, in Table 2, "Upper bound 1" for each maximal allowable size exchange is the formula developed above for that size exchange (i.e. Propositions 1, 2, and 3 for maximal exchange sizes 2, 3, or 4 pairs) with the population size of $n = 25$, 50, or 100. However, in a given sample of size $n = 25$, for example, there may be some recipients who have no compatible donor because of tissue-type incompatibilities, and hence cannot possibly participate in an exchange. In this population there is therefore a smaller number $n' < n$ of pairs actually available for exchange, and "Upper bound 2" in Table 2 reports the average over all populations for the formulas using this smaller population of incompatible recipient-donor pairs. Clearly Upper bound 2 provides a more precise (i.e. lower) upper bound to the number of exchanges that can be found. The fact that the difference between the two upper bounds diminishes as the population size increases reflects that, in larger populations, even highly sensitized recipients are likely to find a compatible donor.

## 6.4 Discussion of the Simulation Results

The static simulation results (which include tissue-type incompatibilities) are very similar to the theoretical upper-bounds we develop for the case with only blood-type incompatibilities. While two-way exchanges account for most of the potential gains from exchange, the number of recipients who benefit from exchange significantly increases when three or more pair exchanges are allowed, and, consistent with the theory, three-way exchanges account for a large share of the remaining potential gains. For example, for a population size of 25 pairs, an average of:

- 11.99 pairs can be matched when any size exchange is feasible,
- 11.27 pairs can be matched when only two-way and three-way exchange are feasible, and
- 8.86 pairs can be matched when only two-way exchange is feasible.

Hence for $n = 25$, two-way exchanges account for 74 percent (i.e. $\frac{8.86}{11.99}$) of the potential gains from exchange whereas three-way exchanges account for 77 percent (i.e. $\frac{11.27-8.86}{11.99-8.86}$) of the remaining potential gains. These rates are 78 percent and 87 percent for a population size of 50 pairs, and 82 percent and 94 percent for a population size of 100 pairs. The theory developed in the absence of crossmatches is still predictive when there are crossmatches: virtually all possible gains from trade

| Pop. Size | Method | Type of Exchange | | | |
|---|---|---|---|---|---|
| | | **Two-way** | **Two-way, Three-way** | **Two-way, Three-way, Four-way** | **No Constraint** |
| **n=25** | **Simulation** | 8.86 (3.4866) | 11.272 (4.0003) | 11.824 (3.9886) | 11.992 (3.9536) |
| | **Upper bound 1** | 12.5 (3.6847) | 14.634 (3.9552) | 14.702 (3.9896) | |
| | **Upper bound 2** | 9.812 (3.8599) | 12.66 (4.3144) | 12.892 (4.3417) | |
| **n=50** | **Simulation** | 21.792 (5.0063) | 27.266 (5.5133) | 27.986 (5.4296) | 28.09 (5.3658) |
| | **Upper bound 1** | 27.1 (5.205) | 30.47 (5.424) | 30.574 (5.4073) | |
| | **Upper bound 2** | 23.932 (5.5093) | 29.136 (5.734) | 29.458 (5.6724) | |
| **n=100** | **Simulation** | 49.708 (7.3353) | 59.714 (7.432) | 60.354 (7.3078) | 60.39 (7.29) |
| | **Upper bound 1** | 56.816 (7.2972) | 62.048 (7.3508) | 62.194 (7.3127) | |
| | **Upper bound 2** | 53.496 (7.6214) | 61.418 (7.5523) | 61.648 (7.4897) | |

Table 2: Simulation results about average number of patients actually matched and predicted by the formulas to be matched. The standard errors of the population are reported in parentheses. The standard errors of the averages are obtained by dividing population standard errors by square root of the simulation number, 22.36.

are achieved with two-way, three-way, and four-way exchanges, especially when the population size is large (see Table 2).[12]

# 7 Dynamic Kidney Exchange

The above two models consider a static situation when there is a pool of recipients with their directed incompatible donors. These models answer how we can organize kidney exchanges in an efficient and incentive-compatible way.

However, in real life, the recipient pool is not static but evolves over time. Ünver (2010) considered a model in which the exchange pool evolves over time by pairs of a recipient and her directed donor arriving with a Poisson distribution in continuous time with an expected arrival rate of $\lambda$. The question answered by this paper is that if there is a constant unit cost of waiting in the pool for each recipient, what is the mechanism that should be run to conduct the exchanges so that the expected discounted exchange surplus is maximized? (It turns out that this is equivalent to maximizing the expected discounted number of recipients to be matched.)

There are also operation research and computer science articles answering different aspects of the dynamic problem:

Zenios (2002) considers a continuous arrival model with pairs of recipients and their directed donors. The model is stylistic in the sense that all blood types are not modeled and all exchanges are two-way. However, the preferences are not 0-1 and the outside option is list exchange.

Awasthi and Sandholm (2009) consider an online mechanism design approach to find optimal dynamic mechanisms for kidney exchange when there are no waiting costs but pairs can exit the pool randomly. They look at mechanisms that are obtained heuristically by sampling future possibilities depending on the current and past matches. Their model has a very large state space; thus, online sampling is used to simplify the optimization problem.

## 7.1 Exchange Pool Evolution

We continue with Ünver's (2010) model. For any pair type X-Y $\in \mathcal{T}$, let $q_{\text{X-Y}}$ be the probability of a random pair being of type X-Y. We refer to $q_{\text{X-Y}}$ as the **arrival probability** of pair type X-Y $\in \mathcal{T}$. We have $\sum_{\text{X-Y} \in \mathcal{T}} q_{\text{X-Y}} = 1$.

Once a pair arrives, if it is not compatible, it becomes available for exchange. If it is compatible, the donor immediately donates a kidney to the recipient of the pair and the pair does not participate

[12] When the population size is 100 incompatible pairs, in 485 of the 500 simulated populations the maximum possible gains from trade are achieved when no more than four pairs are allowed to participate in an exchange.

in exchanges. The **exchange pool** is the set of the pairs which have arrived over time and whose recipient has not yet received a transplant.

Let $p_c$ be the *positive crossmatch* probability that determines the probability that a donor and a recipient will be *tissue-type incompatible*. Let $p_{\text{X-Y}}$ denote the **pool entry probability** of any arriving pair type X-Y. Since blood-type incompatible pairs always join the exchange pool, we have $p_{\text{X-Y}} = 1$ for any blood-type incompatible X-Y. Since blood-type-compatible pairs join the pool if and only if they are not tissue-type compatible, we have $p_{\text{X-Y}} = p_c$ for any blood-type-compatible X-Y. Let $\lambda^p = \lambda \sum_{\text{X-Y} \in \mathcal{T}} p_{\text{X-Y}} q_{\text{X-Y}}$ be the expected number of pairs that enter the pool for exchange per unit time interval.

## 7.2 Time- and Compatibility-Based Preferences

Each recipient has preferences over donors and time of waiting in the pool. For any incompatible pair $i$, recipient $R_i$'s preferences are denoted by $\succsim_i$ and defined over donor-time interval pairs. Recipient $R_i$'s preferences over donors fall into three indifference classes (as in Sections 4 and 5): compatible donors are preferred to being unmatched - an option denoted by being matched with her paired *incompatible* donor $D_i$ - and, in turn, being unmatched is preferred to being matched with incompatible donors. Moreover, time spent in the exchange pool is another dimension in the preferences of recipients: waiting is costly. Formally, preferences of $R_i$ over donors and time spent in the pool are defined as follows:[13]

1. for any two compatible donors $D$ and $D'$ with $R_i$, and time period $t$, $(D, t) \sim_i (D', t)$ (indifference over compatible donors if both transplants occur at the same time),
2. for any compatible donor $D$ with $R_i$ and time periods $t$ and $t'$ such that $t < t'$, $(D, t) \succ_i (D, t')$ (waiting for a compatible donor is costly),
3. for any compatible donor $D$ with $R_i$ and time periods $t$ and $t'$, $(D, t) \succ_i (D_i, t')$ (compatible donors are preferred to remaining unmatched),
4. for any incompatible donor $D \neq D_i$ and time periods $t$ and $t'$, $(D_i, t) \succ_i (D, t')$ (remaining unmatched is preferred to being matched with incompatible donors).

For each pair, we associate waiting in the pool with a monetary cost and we assume that the **unit time cost** of waiting for a transplant by undergoing continuous dialysis is equal to $c > 0$ for each recipient. The alternative to a transplant is dialysis. A recipient can undergo dialysis continuously.

[13] Let $\sim_i$ denote the indifference relation and $\succ_i$ denote the strict preference relation associated with the preference relation $\succsim_i$.

It is well known that receiving a transplant causes the recipient to resume a better life (Overbeck et al. 2005). Also, health care costs for dialysis are higher than those for transplantation in the long term (Schweitzer et al. 1998). We model all the costs associated with undergoing continuous dialysis by the unit time cost $c$.

## 7.3 Dynamically Efficient Mechanisms

A (**dynamic**) **matching mechanism** is a *dynamic* procedure such that at each time $t \geq 0$ it selects a (possibly empty) matching of the pairs available in the pool. Once a pair is matched at time $t$ by a matching mechanism, it leaves the pool and its recipient receives the assigned transplant.

Let $\#^A(t)$ be the total number of pairs that have arrived until time $t$. If mechanism $\phi$ is executed (starting time 0), $\#^{A,\phi}(t)$ is the total number of pairs matched by mechanism $\phi$. There are $\#^A(t) - \#^{A,\phi}(t)$ pairs available at the pool at time $t$.

There is a health authority that oversees the exchanges.

Suppose that the health authority implements a matching mechanism $\phi$. For any time $t$, the current value of expected cost at time $t$ under matching mechanism $\phi$ is given as[14]

$$E_t\left[\mathcal{C}^\phi(t)\right] = \int_t^\infty cE_t\left[\#^A(\tau) - \#^{A,\phi}(\tau)\right] e^{-\rho(\tau-t)} d\tau,$$

where $\rho$ is the **discount rate**.

For any time $\tau, t$ such that $\tau > t$, we have $E_t\left[\#^A(\tau)\right] = \lambda^p(\tau - t) + \#^A(t)$, where the first term is the expected number of recipients to arrive at the exchange pool in the interval $[t, \tau]$ and the second term is the number of recipients that arrived at the pool until time $t$. Therefore, we can rewrite $E_t\left[\mathcal{C}^\phi(t)\right]$ as

$$E_t\left[\mathcal{C}^\phi(t)\right] = \int_t^\infty c(\lambda^p(\tau - t) + \#^A(t) - E_t\left[\#^{A,\phi}(\tau)\right])e^{-\rho(\tau-t)} d\tau.$$

Since $\int_t^\infty e^{-\rho(\tau-t)} d\tau = \frac{1}{\rho}$ and $\int_t^\infty (\tau - t) e^{-\rho(\tau-t)} d\tau = \frac{1}{\rho^2}$, we can rewrite $E_t\left[\mathcal{C}^\phi(t)\right]$ as

$$E_t\left[\mathcal{C}^\phi(t)\right] = \frac{c\lambda^p}{\rho^2} + \frac{\#^A(t)}{\rho} - \int_t^\infty cE_t\left[\#^{A,\phi}(\tau)\right])e^{-\rho(\tau-t)} d\tau. \tag{1}$$

Only the last term in Equation 1 depends on the choice of mechanism $\phi$. The previous terms cannot be controlled by the health authority, since they are the costs associated with the number of recipients arriving at the pool. We refer to this last term as the **exchange surplus at time** $t$ **for** mechanism $\phi$ and denote it by

[14] $E_t$ refers to the expected value at time $t$.

$$\mathcal{ES}^{\phi}(t) = \int_t^{\infty} cE_t\left[\#^{\phi}(\tau)\right] e^{-\rho(\tau-t)} d\tau.$$

We can rewrite it as

$$\begin{aligned}
\mathcal{ES}^{\phi}(t) &= \int_t^{\infty} c\left(E_t\left[\#^{\phi}(\tau) - \#^{\phi}(t)\right] + \#^{A,\phi}(t)\right) e^{-\rho(\tau-t)} d\tau \\
&= \frac{c\#^{\phi}(t)}{\rho} + \int_t^{\infty} c\left(E_t\left[\#^{\phi}(\tau) - \#^{A,\phi}(t)\right]\right) e^{-\rho(\tau-t)} d\tau.
\end{aligned}$$

The first term above is the exchange surplus attributable to all exchanges that have been done until time $t$ and at time $t$ and the second term is the **future exchange surplus** attributable to the exchanges to be done in the future. The central health authority cannot control the number of past exchanges at time $t$ either. Let $n^{\phi}(\tau)$ be the **number of matched recipients at time $\tau$ by mechanism $\phi$**, and we have

$$\#^{\phi}(t) = \left(\sum_{\tau<t} n^{\phi}(\tau)\right) + n^{\phi}(t).$$

We focus on the **present and future exchange surplus** which is given as

$$\widetilde{\mathcal{ES}}^{\phi}(t) = \frac{cn^{\phi}(t)}{\rho} + \int_t^{\infty} c\left(E_t\left[\#^{A,\phi}(\tau) - \#^{A,\phi}(t)\right]\right) e^{-\rho(\tau-t)} d\tau. \tag{2}$$

A dynamic matching mechanism $\nu$ is (**dynamically**) **efficient** if for any $t$, it maximizes the *present and future exchange surplus* at time $t$ given in Equation 2. We look for solutions of the problem independent of initial conditions and time $t$. We will define a steady-state formally. If such solutions exist, they depend only on the "current state of the pool" (defined appropriately) but not on time $t$ or the initial conditions.

## 7.4 Dynamically Efficient *Two-way* Exchange

In this subsection, we derive the dynamically optimal *two-way* matching mechanism. Throughout this subsection we will maintain two assumptions, Assumptions 1 and 2, introduced in Section 5.

We are ready to state Theorem 7.

**Theorem 7** ***(Ünver 2010)*** *Let dynamic matching mechanism $\nu$ be defined as a mechanism that matches only X-Y type pairs with their reciprocal Y-X type pairs, immediately when such an exchange is feasible. Then, under Assumptions 1 and 2, mechanism $\nu$ is a dynamically optimal two-way matching mechanism.*

*Moreover, a dynamically optimal* two-way *matching mechanism conducts a two-way exchange whenever one becomes feasible.*

Next we show that Assumption 2 will hold in the long run under the most reasonable pair-type arrival distributions; thus, it is not a restrictive assumption.

**Proposition 4** *(**Ünver 2010**) Suppose that* $p_c\left(q_{AB\text{-}O}+q_{X\text{-}O}\right) < q_{O\text{-}X}$ *for all* $X \in \{A,B\}$, $p_c\left(q_{AB\text{-}O}+q_{AB\text{-}X}\right) < q_{X\text{-}AB}$ *for all* $X \in \{A,B\}$ *and* $p_c q_{AB\text{-}O} < q_{O\text{-}AB}$. *Then, Assumption 2 holds in the long run regardless of the two-way matching mechanism used.*

The hypothesis of the above proposition is very mild and will hold for sufficiently small crossmatch probability. Moreover, it holds for real-life blood frequencies. For example, assuming that the recipient and the paired-donor are blood-unrelated, the arrival rates reported in the simulations section of the paper satisfy these assumptions, when the crossmatch probability is $p_c = 0.11$, as reported by Zenios, Woodle, and Ross (2001).

## 7.5 Dynamically Efficient *Multi-way* Exchanges

In this section, we consider matching mechanisms that allow for not only two-way exchanges, but larger exchanges as well. Roth, Sönmez, and Ünver (2010) have studied the importance of three-way and larger exchanges in a static environment, and we summarized these results in Section 5. The results in this subsection follow this intuition, and are due to Ünver (2010). We can state the following observation motivated by the results in Section 5:

**Observation 4** *In an exchange that matches an underdemanded pair, there should be at least one overdemanded pair. In an exchange that matches a reciprocally demanded pair, there should at least be one reciprocal type pair or an overdemanded pair.*

Using the above illustration, under realistic blood-type distribution assumptions, we will prove that Assumption 2 still holds, when the applied matching mechanism is unrestricted. Recall that through Assumption 2, we assumed to have arbitrarily many underdemanded type pairs available in the long-run states of the exchange pool, regardless of the dynamic matching mechanism used to achieve the long run.

**Proposition 5** *(**Ünver 2010**) Suppose that* $p_c\left(q_{AB\text{-}O}+q_{X\text{-}O}\right) + \min\{p_c q_{Y\text{-}O}, q_{X\text{-}Y}\} < q_{O\text{-}X}$ *for all* $\{X,Y\} = \{A,B\}$, $p_c\left(q_{AB\text{-}O}+q_{AB\text{-}X}\right) + \min\{p_c q_{AB\text{-}Y}, q_{Y\text{-}X}\} < q_{X\text{-}AB}$ *for all* $\{X,Y\} = \{A,B\}$ *and* $p_c q_{AB\text{-}O} < q_{O\text{-}AB}$. *Then, Assumption 2 holds in the long run regardless of the unrestricted matching mechanism used.*

The hypothesis of the above proposition is also very mild and will hold for sufficiently small crossmatch probability $p_c$. Moreover, it holds for real-life blood frequencies and crossmatch probability.

For example, assuming that the recipient and the paired-donor are blood-unrelated, the arrival rates reported in the simulations section of the paper satisfy these assumptions. Thus, we can safely use Assumption 2 in this section, as well.

Next, we characterize the dynamically efficient mechanism.

In a dynamic setting, the structure of three-way and four-way exchanges discussed above may cause the second part of Theorem 7 not to hold when these larger exchanges are feasible. More specifically, we can benefit from not conducting all feasible exchanges currently available, and holding on to some of the pairs that can currently participate in an exchange in expectation of saving more pairs in the near future.

We maintain Assumption 1 as well as Assumption 2 in this subsection. We state one other assumption. First, we state that as long as the difference between A-B and B-A type arrival frequencies is not large, overdemanded type pairs will be matched immediately.

**Proposition 6** ***(Ünver 2010)*** *Suppose Assumptions 1 and 2 hold. If $q_{A\text{-}B}$ and $q_{B\text{-}A}$ are sufficiently close, then under the dynamically efficient multi-way matching mechanism, overdemanded type pairs are matched as soon as they arrive at the exchange pool.*

**Assumption 5** ***(Assumption on Generic Arrival Rates of Reciprocally Demanded Types):*** *A-B and B-A type pairs arrive at relatively close frequency to each other so that Proposition 6 holds.*

Under Assumptions 1, 2, and 5, we will only need to make decisions in situations in which multiple exchanges of different sizes are feasible: For example, consider a situation in which an A-O type pair arrives at the pool, while a B-A type pair is also available. Since by Assumption 2, there is an excess number of O-A and O-B type pairs in the long run, there are two sizes of feasible exchanges, a three-way exchange (for example, involving A-O, O-B, and B-A type pairs) or a two-way exchange (for example, involving A-O and O-A type pairs). Which exchange should the health authority choose?

To answer this question, we analyze the dynamic optimization problem. Since the pairs arrive according to a Poisson process, we can convert the problem to an embedded **Markov decision process**. We need to define a state space for our analysis. Since the pairs in each type are symmetric by Assumption 1, the natural candidate for a **state** is a 16-dimensional vector, which shows the number of pairs in each type available. In our exchange problem, there is additional structure to eliminate some of these state variables. We look at overdemanded, underdemanded, self-demanded, and reciprocally demanded types separately:

- **Overdemanded types**: If an overdemanded pair $i$ of type X-Y $\in \mathcal{T}^O$ arrives, by Proposition 6, pair $i$ will be matched immediately in some exchange. Hence, the number of overdemanded pairs remaining in the pool is always 0.

- **Underdemanded types**: By Assumption 2 as well as Assumption 1, there will be an arbitrarily large number of underdemanded pairs. Hence, the number of underdemanded pairs is always $\infty$.

- **Self-demanded types**: Whenever a self-demanded pair $i$ of type X-X $\in \mathcal{T}^S$ is available in the exchange pool, it can be matched through two ways under a multi-way matching mechanism:

  1. If another X-X type pair $j$ arrives, by Assumption 1, $i$ and $j$ will be mutually compatible, and a two-way exchange $(i, j)$ can be conducted.
  2. If an exchange $E = (i_1, i_2, ..., i_k)$, with Y blood-type donor $D_{i_k}$ and Z blood-type recipient $R_{i_1}$, becomes feasible, and blood-type Y donors are blood-type compatible with blood-type X recipients, while blood-type X donors are blood-type compatible with blood-type Z recipients, then pair $i$ can be inserted in exchange $E$ just after $i_k$, and by Assumption 1, the new exchange $E' = (i_1, i_2, ..., i_k, i)$ will be feasible.

  By Observation 4, a self-demanded type can never save an overdemanded or reciprocally demanded pair without the help of an overdemanded or another reciprocally demanded pair. Suppose that there are $n$ X-X type pairs. Then, they should be matched in two-way exchanges to save $2\left\lfloor \frac{n}{2} \right\rfloor$ of them (which is possible by Assumption 1). This and the above observations imply that under a dynamically efficient matching mechanism, for any X-X $\in \mathcal{T}^S$, at steady-state there will be either 0 or 1 X-X type pair.

  Therefore, in our analysis, existence of self-demanded types will be reflected by four additional state variables, each of which gets values either 0 or 1. We will derive the efficient dynamic matching mechanism by ignoring the self-demanded type pairs:

**Assumption 6** ***(No Self-Demanded Types Assumption):*** *There are no self-demanded types available for exchange and $q_{X\text{-}X} = 0$ for all X-X $\in \mathcal{T}$.*

- **Reciprocally demanded types**: By the above analysis, there are no overdemanded or self-demanded type pairs available and there are infinitely many underdemanded type pairs. Therefore, the state of the exchange pool can simply be denoted by the number of A-B type pairs and B-A type pairs. By Assumption 1, an A-B type pair and B-A type pair are mutually compatible with each other, and they can be matched in a two-way exchange. Moreover, by Observation 4, an A-B or B-A type pair cannot save an underdemanded pair in an exchange without the help of an overdemanded pair. Hence, the most optimal use of A-B and B-A type pairs is being matched with each other in a two-way exchange. Therefore, under the optimal matching mechanism, an A-B and B-A type pair will never remain in the pool together but will be matched via a two-way exchange. By this observation, we can simply denote the state of the exchange pool by an integer $s$, such that if $s > 0$, then $s$ refers to the number of A-B type pairs in the exchange pool, and if $s < 0$, then $|s|$ refers to the number of B-A type pairs in the exchange pool. Formally $s$ is the difference between the number of A-B type pairs and B-A type pairs in the pool, and *only one of these two numbers can be non-zero.* Let $S = \mathbb{Z}$ be the state space (i.e., the set of integers).

### 7.6 Markov Chain Representation

In this subsection, we characterize the transition from one state to another under a dynamically optimal matching mechanism by a Markov chain given Assumptions 1, 2, 5, and 6:

First suppose $s > 0$, i.e. there are some A-B type pairs and no B-A type pairs. Suppose a pair of type X-Y $\in \mathcal{T}$ becomes available. In this case, three subcases are possible for pair $i$:

1. X-Y $\in \mathcal{T}^U = \{$O-A, O-B, O-AB, A-AB, B-AB$\}$: By Observation 4, in any exchange involving an underdemanded pair, there should be an overdemanded pair. Since there are no overdemanded pairs available under the optimal mechanism, no new exchanges are feasible. Moreover, the state of the exchange pool remains as $s$.

2. X-Y $\in \mathcal{T}^O = \{$A-O, B-O, AB-O, AB-A, AB-B$\}$: If pair $i$ is compatible (which occurs with probability $1 - p_c$), donor $D_i$ donates a kidney to recipient $R_i$, and pair $i$ does not arrive at the exchange pool. If pair $i$ is incompatible (which occurs with probability $p_c$), pair $i$ becomes available for exchange. Three cases are possible:

   (a) X-Y $\in \{$A-O, AB-B$\}$: Since $s > 0$, there are no B-A type pairs available. In this case, there is one type of exchange feasible: a two-way exchange including pair $i$, and a mutually compatible pair $j$ of type Y-X. By Assumption 2, such a Y-X type pair exists. By Proposition 6, this exchange is conducted, resulting with 2 matched pairs, and the state of the pool remains as $s$. There is no decision problem in this state.

   (b) X-Y $\in \{$B-O, AB-A$\}$: Since $s > 0$, there are A-B type pairs available. There are two types of exchanges that can be conducted: a two-way exchange and a three-way exchange:

      - By Assumption 2, there is a mutually compatible pair $j$ of type Y-X, and $(i, j)$ is a feasible two-way exchange.
      - If X-Y = B-O: By Assumption 2, there is an arbitrary number of O-A type pairs. Let pair $j$ be an O-A type pair. Let $k$ be an A-B type pair in the pool. By Assumption 2, $(i, j, k)$ is a feasible three-way exchange (see Figure 3).

If X-Y = AB-A: By Assumption 2, there is an arbitrary number of B-AB type pairs. Let $k$ be a B-AB type pair. Let $j$ be an A-B type pair in the pool. By Assumption 1, $(i, j, k)$ is a feasible three-way exchange.

Let action $a_1$ refer to *conducting a smaller exchange* (i.e., two-way), and action $a_2$ refer to *conducting a larger exchange* (i.e., three-way). If action $a_1$ is chosen, 2 pairs are matched, and the state of the pool does not change. If action $a_2$ is chosen, then 3 pairs are matched, and the state of the pool decreases to $s - 1$.

(c) X-Y = AB-O: Since $s > 0$, there are three types of exchanges that can be conducted: a two-way exchange, a three-way exchange, or a four-way exchange:

- By Assumption 2 and Observation 1, for any W-Z $\in \mathcal{T}^U$, there is a mutually compatible pair $j$ of type W-Z for pair $i$. Hence, $(i, j)$ is a feasible two-way exchange.
- By Assumption 2, there are pair $j$ of type O-B and pair $k$ of type B-AB such that $(i, j, k)$ is a feasible three-way exchange. Also by Assumption 2, there are pair $g$ of type O-A and pair $h$ of type A-AB such that $(g, h, i)$ is a feasible three way-exchange (see Figure 2a,b). By Assumption 2, there is an arbitrarily large number of underdemanded pairs independent of the matching mechanism, therefore, conducting either of these two three-way exchanges has the same effect on the future states of the pool. Hence, we will not distinguish these two types of exchanges.
- By Assumptions 1 and 2, a pair $h$ of type B-AB, a pair $j$ of type O-A, and a pair $k$ of type A-B form the four-way exchange $(h, i, j, k)$ with pair $i$ (see Figure 4).

Two-way and three-way exchanges do not change the state of the pool. Therefore, conducting a three-way exchange dominates conducting a two-way exchange. Hence, under the optimal mechanism, we rule out conducting a two-way exchange, when an AB-O type pair arrives. Let action $a_1$ refer to *conducting a smaller* (i.e., three-way) *exchange*, and let action $a_2$ refer to *conducting a larger* (i.e., four-way) *exchange*. If action $a_1$ is chosen, 3 pairs are matched, and the state of the pool remains as $s$. If action $a_2$ is chosen, 4 pairs are matched, and the state of the pool decreases to $s - 1$.

3. X-Y $\in \mathcal{T}^R =$ {A-B, B-A}: Two cases are possible:

(a) X-Y = A-B: By Observation 4, an A-B type pair can only be matched using a B-A type pair or an overdemanded pair. Since there are no overdemanded and B-A type pairs, there is no possible exchange. The state of the pool increases to $s + 1$.

(b) X-Y = B-A: By Assumption 1, a feasible two-way exchange can be conducted using an A-B type pair $j$ in the pool and pair $i$. This is the only feasible type of exchange. Since matching a B-A type pair with an A-B type pair is the most optimal use of these types of pairs, we need to conduct such a two-way exchange and the state of the pool decreases to $s - 1$.

Note that we do not need to distinguish decisions regarding two-way versus three-way exchanges and three-way versus four-way exchanges. We denote all actions regarding smaller exchanges by $a_1$ and all actions regarding larger exchanges by $a_2$. Since the difference between a smaller exchange and a larger exchange is always one pair, i.e., an A-B type pair, whenever the state of the pool dictates that a three-way exchange is chosen instead of a two-way exchange when a B-O or AB-A type pair arrives, then it will also dictate that a four-way exchange will be chosen instead of a three-way exchange when an AB-O type pair arrives.

For $s < 0$, that is, when $|s|$ B-A type pairs are available in the exchange pool, we observe the symmetric version of the above evolution. For $s = 0$, that is, when there are no A-B or B-A type pairs available in the exchange pool, the evolution is somewhat simpler. At state 0, the only state transition occurs, when an A-B type pair arrives (to state 1), or when a B-A type pair arrives (to state -1). Actions involving largest exchanges for the case $s > 0$, referred to as *action* $a_2$, are infeasible at state 0, implying that there is no decision problem. Moreover, there are no exchanges involving A-B or B-A type pairs. In this state, a maximum size exchange is conducted when it becomes feasible.

## 7.7 The Dynamically Efficient Multi-way Matching Mechanism

A **(deterministic) Markov matching mechanism** $\phi$ is a matching mechanism that chooses the same action whenever the Markov chain is in the same state. In our reduced state and action problem, a Markov matching mechanism chooses either action $a_1$, conducting the smaller exchange, or action $a_2$, conducting the largest exchange, at each state, except state 0. The remaining choices of the Markov mechanism are straightforward: It chooses a maximal exchange when such an exchange becomes feasible (for negative states by interchanging the roles of A and B blood types as outlined in the previous subsection). Formally, $\phi : S \to \{a_1, a_2\}$ is a Markov matching mechanism.

Next we define a class of Markov matching mechanisms. A Markov matching mechanism $\phi^{\overline{s},\underline{s}} : S \to \{a_1, a_2\}$ is a **threshold matching mechanism with thresholds** $\overline{s} \geq 0$ and $\underline{s} \leq 0$, if

$$\phi^{\overline{s},\underline{s}}(s) = \begin{cases} a_1 & \text{if } \underline{s} \leq s \leq \overline{s} \\ a_2 & \text{if } s < \underline{s} \text{ or } s > \overline{s} \end{cases} .$$

A threshold matching mechanism conducts the largest exchanges that do not use existing A-B or B-A type pairs ("the smaller exchanges") as long as the numbers of A-B or B-A type pairs are not greater than the threshold numbers, $\overline{s}$ and $|\underline{s}|$ respectively; otherwise it conducts the largest possible exchanges including the existing A-B or B-A type pairs ("the larger exchanges").

Our next Theorem is as follows:

**Theorem 8** *(**Ünver 2010**) Suppose Assumptions 1, 2, 5, and 6 hold. There exist $\overline{s}^* = 0$ and $\underline{s}^* \leq 0$, or $\overline{s}^* \geq 0$ and $\underline{s}^* = 0$ such that $\phi^{\overline{s}^*, \underline{s}^*}$ is a dynamically efficient multi-way matching mechanism.*

The dynamically optimal matching mechanism uses a threshold mechanism. It stocks A-B or B-A type pairs, and does not use them in larger exchanges as long as the stock of the control group is less than or equal to $\overline{s}^*$ or $|\underline{s}^*|$, respectively. Under the optimal matching mechanism, either the number of A-B type pairs or B-A type pairs is the state variable, but not both. Under the first type of solution, the number of B-A type pairs is the state variable. As long as the number of B-A type pairs in the pool is zero, regardless of the number of A-B type pairs, when the next arrival occurs, the first type of optimal mechanism conducts the maximal size exchanges possible. If there are B-A type pairs and their number does not exceed the threshold number $|\underline{s}^*|$, then these pairs are exclusively used to match incoming A-B type pairs in two-way exchanges. On the other hand, if the number of B-A type pairs exceeds the threshold number $|\underline{s}^*|$, they should be used in maximal exchanges, which can be (1) a two-way exchange involving an A-B type pair if the incoming pair type is A-B, (2) a three-way exchange involving A-O and O-B type pairs or A-AB and AB-B type pairs if the incoming pair type is A-O or AB-B, respectively, and (3) a four-way exchange involving A-AB, AB-O, and O-B type pairs if the incoming pair type is AB-O. The other types of maximal exchanges are conducted by the optimal mechanism as soon as they become feasible. The second possible solution is the symmetric version of the above mechanism taking the number of A-B type pairs as a state variable.

Next, we specify the optimal mechanism more precisely.

**Theorem 9** *(**Ünver 2010**) Suppose Assumptions 1, 2, 5, and 6 hold. Then,*

- *If $q_{A\text{-}B} \geq q_{B\text{-}A}$, that is, A-B type arrives at least as frequently as B-A type, and $q_{B\text{-}O} + q_{AB\text{-}A} < q_{A\text{-}O} + q_{AB\text{-}B}$, that is, the types that can match A-B type pairs in larger exchanges arrive less frequently than those for the B-A type, then $\phi^{0,\underline{s}^*}$ is the dynamically efficient multi-way matching mechanism for some $\underline{s}^* \leq 0$.*

- *If $q_{A\text{-}B} = q_{B\text{-}A}$ and $q_{B\text{-}O} + q_{AB\text{-}A} = q_{A\text{-}O} + q_{AB\text{-}B}$, then $\phi^{0,0}$ is the dynamically efficient multi-way matching mechanism. That is, maximal size exchanges are conducted whenever they become feasible.*

- *If $q_{A\text{-}B} \leq q_{B\text{-}A}$ and $q_{B\text{-}O} + q_{AB\text{-}A} > q_{A\text{-}O} + q_{AB\text{-}B}$, then $\phi^{\overline{s}^*,0}$ is the dynamically efficient multi-way matching mechanism for some $\overline{s}^* \geq 0$.*

According to the arrival frequencies reported in Table 1, for pairs forming between random donors and recipients, we expect the mechanism reported in the first bullet-point to be the efficient mechanism.

# 8 Concluding Remarks

We conclude our survey by surveying other topics that have attracted the attention of researchers and practitioners alike.

## 8.1 Computational Issues

Following Roth, Sönmez, and Ünver (2007), one can write an integer program to solve the maximal kidney exchange problem.

We give the explicit formulation of finding the maximal number of patients who can benefit from two-way and up to $k$-way exchanges for any number $k$ such that $|N| \geq k \geq 2$.

Suppose $E = \left(R_{i_1} - D_{i_1},...,R_{i_k} - D_{i_k}\right)$ denotes a $k$-way exchange in which pairs $i_1, ..., i_k$ participate. Let $|E|$ be the number of transplants possible under $E$; hence we have $|E| = k$.

Let $\mathcal{E}^k$ be the set of feasible two-way through $k$-way exchanges possible among the pairs in $N$. For any pair $i$, let $\mathcal{E}^k(i)$ denote the set of exchanges in $\mathcal{E}^k$ such that pair $i$ can participate. Let $x = (x_E)_{E \in \mathcal{E}^k}$ be a vector of 0s and 1s such that $x_E = 1$ denotes that exchange $E$ is going to be conducted, and $x_E = 0$ denotes that exchange $E$ is not going to be conducted. Our problem of finding a maximal set of patients who will benefit from two-way,..., and $k$-way exchanges is given by the following integer program:

$$\begin{aligned} & \max_x \sum_{E \in \mathcal{E}^k} |E| \, x_E \\ & \text{subject to} \\ & x_E \in \{0,1\} \qquad \forall E \in \mathcal{E}^k \ , \\ & \sum_{E \in \mathcal{E}^k(i)} x_E \leq 1 \qquad \forall i \in N \ . \end{aligned}$$

This problem is solved using Edmonds' (1965) algorithm for $k = 2$ (i.e., only for two-way exchanges) in polynomial time. However, for $k \geq 3$ this problem is NP-complete.[15] (See also Abraham, Blum, and Sandholm 2007.)

[15]The observation that the mixed 2- and 3-way problem is NP complete was made by Kevin Cheung and Michel Goemans (personal communication.)

We also formulate the following version of the integer programming problem, which does not require ex-ante construction of the sets $\mathcal{E}^k$:

Let $C^* = \left[c^*_{i,j}\right]_{i\in N, j\in N}$ be a matrix of 0s and 1s such that if Recipient $R_i$ is compatible with Donor $D_j$ we have $c^*_{i,j} = 1$ and if $R_i$ is not compatible with donor $D_j$ we have $c^*_{i,j} = 0$. Let $X = [x_{i,j}]_{i\in N, j\in N}$ be the **assignment matrix** of 0s and 1s such that $x_{i,j} = 1$ denotes that recipient $R_i$ receives a kidney from donor $D_j$ and $x_{i,j} = 0$ denotes that recipient $R_i$ does not receive a kidney from donor $D_j$ under the proposed assignment $X$. We solve the following integer program to find a maximal set of two-way,...,and $k$-way exchanges:

$$\max_X \sum_{i\in N, j\in N} x_{i,j}$$

subject to

$$x_{i,j} \in \{0,1\} \qquad \forall i,j \in N\ , \tag{3}$$

$$x_{i,j} \leq c^*_{i,j} \qquad \forall i,j \in N\ , \tag{4}$$

$$\sum_{j\in N} x_{i,j} \leq 1 \qquad \forall i \in N\ , \tag{5}$$

$$\sum_{j\in N} x_{i,j} = \sum_{j\in N} x_{j,i} \qquad \forall i \in N\ , \tag{6}$$

$$x_{i_1,i_2} + x_{i_2,i_3} + ... + x_{i_k,i_{k+1}} \leq k-1 \qquad \forall \{i_1, i_2, ..., i_k, i_{k+1}\} \subseteq N\ . \tag{7}$$

A solution of this problem determines a maximal set of patients who can benefit from two-way,..., and $k$-way exchanges for any $k < |N|$. A maximal set of patients who can benefit from unrestricted exchanges is found by setting $k = |N|$. In this case Constraint 7 becomes redundant. This formulation is used to find the maximal set of unrestricted multi-way exchanges.

Since the problems are NP-complete for $k > 2$, there is no known algorithm that runs in worst-case time that is polynomial in the size of the input. Simulations have shown that for more than a certain number of pairs in the exchange pool, commercial integer programming software programs have difficulty solving these optimization problems. Abraham, Blum, and Sandholm (2007) proposed a tailored integer programming algorithm designed specifically to solve large kidney exchange problems.[16] This algorithm increases the scalability of a computable problem size considerably more than commercial integer programming software capabilities and can solve the problem optimally in less than 2 hours at the full projected scale of the nationwide kidney exchange (10,000 pairs). The US national kidney exchange program whose pilot runs started to be conducted in the late 2010 uses this

[16]There is also a recent strand of literature that deals with different computability issues under various solution concepts for the kidney exchange problem. See e.g. Cechlárová, Fleiner, and Manlove (2005), Biró and Cechlárová (2007), Irving (2007), Biró and McDermid (2008).

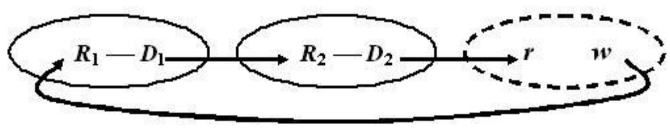
tailored algorithm, while some regional programs continue to use commercial integer programming software versions of the computational implementation.

## 8.2 List Exchange Chains

Another concept that is being implemented in NEPKE is that of **list exchange chains** (Roth, Sönmez, and Ünver 2004; see also Roth et al. 2007). A $k-$way list exchange chain is similar to a $k-$way paired kidney exchange, with the exception that one of the pairs in the exchange is a **virtual pair** with the property that

- the donor of this pair is a priority on the deceased donor waiting list, i.e., whomever is assigned this donor gets priority to receive the next incoming deceased donor kidney; and
- the recipient of this pair is the highest-priority recipient who is waiting for a kidney on the deceased donor waiting list.

Thus, in a list exchange chain, one recipient of a pair receives a priority to receive the next incoming compatible deceased donor kidney (by trading her own paired live donor's kidney); and one donor of a pair in the exchange does not donate to anybody in the exchange pool but donates to the top-priority recipient waiting for a deceased donor kidney.




A three-way list exchange chain. Here $r$ refers to the recipient on the deceased donor waiting list and $w$ refers to priority on the deceased donor waiting list.

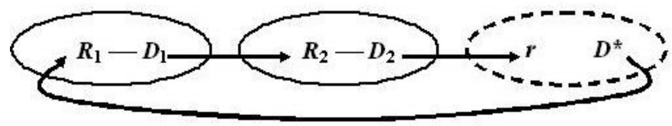
There are two ethical concerns regarding list exchanges in the medical community; therefore, not all regions implement it (Ross et al. 1997, Ross and Woodle 2000).

The first concern regards the imbalance between the blood type of the recipient at the top of the waiting list who receives a kidney and the recipient in the exchange pool who receives top priority on the waiting list. Because of blood-type compatibility requirements, most of the time the recipient who gets a live-donor kidney will be of an inferior type, such as AB, A, or B, while the recipient who is sent to the top of the waiting list will be of O blood-type. Thus, this will increase the waiting time for O blood-type patients on the waiting list.

The second concern regards the inferior quality of deceased-donor kidneys compared to live-donor kidneys. Many medical doctors are not willing to leave such a decision to patients, i.e., whether to exchange a live-donor kidney for a deceased-donor kidney.

## 8.3 Altruistic Donor Chains

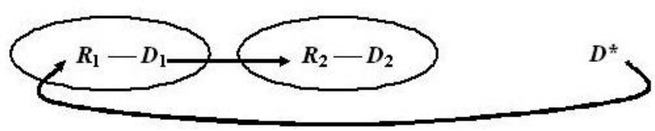
A new form of exchange is finding many applications in the field. In a year, there are about 100 **altruistic donors**, live-donors who are willing to donate one of their kidneys to a stranger, in the US. Such donations are not regulated and traditionally have been treated like deceased-donor donations. However, a recent paradigm suggests that an altruistic donor can donate to a pair in the exchange pool, and in return this pair can donate to another pair, ..., and finally the last pair donates to the top-priority recipient on the waiting list. This is referred to as a **simultaneous altruistic donor chain** (Montgomery et al. 2006; see also Roth et al. 2007).




A simultaneous three-way altruistic donor chain. Here $D^*$ refers to the altruistic donor and $r$ refers to a recipient on the top of the deceased donor waiting list.

Thus, instead of an altruistic donor helping a single recipient on the waiting list, he helps $k$ recipients in a $k-$way closed altruistic donor chain.

A newer paradigm takes this idea one step forward. Instead of the last donor immediately donating a kidney to a recipient on the waiting list, he becomes a **bridge donor**, i.e., he acts as an altruistic donor and may help a future incoming pair to the exchange. The problem with this approach is that the bridge donor can opt out from future donation after his paired-recipient receives a kidney. However, field experimentation suggests that in APD no bridge donor has backed out yet in any of the 6 operational chains. Such an exchange is referred to as a **non-simultaneous altruistic donor chain (**Roth et al. 2007 and Rees et al. 2009**).**




A non-simultaneous two-way altruistic donor chain. Here, $D^*$ refers to the altruistic donor, and $D_2$ is the bridge donor who will act as an altruistic donor in a future altruistic donor chain.

The potential impact of altruistic donor chains is quite large. For example, in APD, 22 transplantations were conducted through 6 non-simultaneous altruistic donor chains in 10 states, all with active bridge donors (at the time this manuscript was drafted).

## 8.4 Exchange with Compatible Pairs

Currently, compatible pairs are not part of the kidney exchange paradigm, since the recipient of the pair receives directly a kidney from her paired-donor. Woodle and Ross (1998) proposed compatible pairs to be included in kidney exchanges, since they will contribute to a substantial increase in the number of transplants from exchanges. Indeed, the simulations by Roth, Sönmez, and Ünver (2005b) show that when compatible pairs are used in exchanges, since the pairs will likely be of overdemanded types, they will increase the gains from exchange tremendously (also see Roth, Sönmez, and Ünver (2004)). Table 3 shows the results of this simulation for efficient two-way exchange mechanisms.

This table shows the dramatic potential impact of including compatible pairs in exchange. When list exchange is not possible for $n = 100$, about 70% of the pairs receive a kidney when only incompatible pairs participate in exchange. This number increases to 91% when compatible pairs also participate in exchange.

Sönmez and Ünver (2010), the authors of this survey, model the two-way kidney exchange problem with compatible pairs. We obtain favorable graph-theoretical results analogous to the problem without compatible pairs (see Roth, Sönmez, and Ünver 2005a). We show that the latter is a special case of the former general model and extend the Gallai–Edmonds decomposition to this domain. We introduce an algorithm that finds a Pareto-efficient matching with polynomial time and space requirements. We generalize the most economically relevant results and the *priority* mechanisms to this domain. Moreover, our results generalize to a domain that includes altruistic donors that are incorporated through simultaneous two-way chains.

## 8.5 False-Negative Crossmatches

Detection of tissue-type incompatibility without a *crossmatch test* is not a perfect science. Since this test, which involves mixing blood samples from the donor and the recipient, is expensive to conduct between all donors and recipients, exchange programs usually rely on a different method to determine whether a donor is tissue-type compatible with a recipient. Using a simple *antibody test*, doctors determine the HLA proteins that trigger antibodies in a recipient. Also taking into account the previous rejection and sensitivity history of the recipient, they determine the HLA proteins that are compatible (or incompatible) with her. Hence, the donors who have the compatible (or incompatible) HLAs are deemed tissue-type compatible (or incompatible) with the recipient. However, this test has a flow: the false-negative crossmatch (false tissue-type compatibility) rate is sometimes high. As a result, some exchanges found by the matching mechanism do not go through. Such cases affect the whole match, since different outcomes could have been found if these incompatibilities had been taken into account. Large kidney exchange programs with an extended history can partially avoid this problem, since many actual crossmatch tests have already been conducted between many donors

| **Compatible Pairs** | **Population Size** | **% wait-list option** | **Total Transplants** | | | |
|---|---|---|---|---|---|---|
| | | | | **Own** | **Exchange** | **w-List** |
| out of the exchange | $n =$ 25 | 0 % | 15.52 | 11.56 | 3.96 | 0 |
| | | 40 % | 21.03 | 11.56 | 5.76 | 3.71 |
| | $n =$ 100 | 0 % | 70.53 | 47.49 | 23.04 | 0 |
| | | 40 % | 87.76 | 47.49 | 28.79 | 11.48 |
| in the exchange | $n =$ 25 | 0 % | 20.33 | 1.33 | 19.00 | 0 |
| | | 40 % | 23.08 | 1.33 | 19.63 | 2.12 |
| | $n =$ 100 | 0 % | 91.15 | 1.01 | 90.14 | 0 |
| | | 40 % | 97.06 | 1.01 | 91.35 | 4.70 |

Table 3: A Pareto-efficient two-way exchange mechanism outcome for $n$ pairs randomly generated using national population characteristics (including compatible and incompatible pairs) when compatible pairs are in/out of exchange, when $n = 25/100$, when list exchanges are impossible/possible and 40% of the pairs are willing to use this option. *Own* refers to the patients receiving their own-donor kidneys (i.e., when compatible pairs are out, this is the number of compatible pairs generated in the population). *Exchange* refers to the number of patients who receive a kidney through exchange. *w-List* refers to the number of patients who get priority on the waiting list when list exchange is possible.

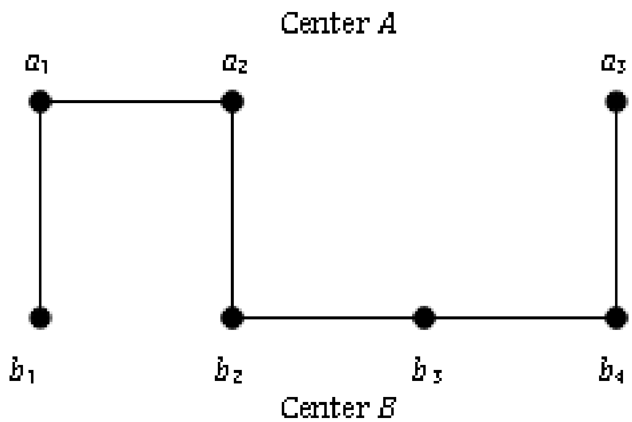
and recipients over the years. They can simply use this data in matching instead of the simple test results. Morrill (2008) introduces a mechanism for the two-way matching problem (aka *roommates* problem) to find a Pareto-efficient matching starting from a Pareto-inefficient matching. His model's preference domain is strict preferences. An application of this mechanism is as follows: after a set of kidney exchanges are fixed, if some of these fail to go through for some reason, we can use Morrill's mechanism to find a matching that Pareto-dominates the initial one. This mechanism has a novel polynomial time algorithm that synthesizes the intuition from Gale's top-trading cycles algorithm (used to find the core for strict preferences with unrestricted multi-way exchanges) with Edmonds' algorithm (used to find a Pareto-efficient matching for 0-1 preferences with two-way exchanges).

## 8.6 Transplant Center Incentives

Transplant centers decide voluntarily whether to participate in a larger exchange program, such as the APD or the forthcoming national program. Moreover, if they do, they are free to determine which recipients of their center will be matched through the larger program. Thus, centers can strategically decide which of their patients will be matched through the larger program. If centers care about maximizing the number of recipients to be matched through exchanges, the following result shows that no efficient mechanism is immune to manipulation:

**Theorem 10** ***(Roth, Sönmez, and Ünver 2005c):*** *Even if there is no tissue-type incompatibility between recipients and donors of different pairs, there exists no Pareto efficient mechanism where full participation is always a dominant strategy for each transplant center.*

The proof is through an example: There are two transplant centers $A, B$, three pairs $a_1, a_2, a_3 \in I_A$ in center $A$, and four pairs $b_1, b_2, b_3, b_4 \in I_B$ in center $B$. Suppose that the list of feasible exchanges are as follows: $(a_1, a_2)$, $(a_1, b_1)$, $(a_2, b_2)$, $(a_3, b_4)$, $(b_2, b_3)$, $(b_3, b_4)$. The following figure shows all feasible exchanges among the pairs:




In all Pareto efficient matchings 6 pairs receive transplants (an example is $\{(a_1, b_1), (a_2, b_2), (b_3, b_4)\}$). Since there are 7 pairs, one of the pairs remains unmatched under any Pareto-efficient matching. Let $\phi$ be a Pareto-efficient mechanism. Since $\phi$ chooses a Pareto-efficient matching, there is a single pair that does not receive a transplant. This pair is either in Center $A$ or in Center $B$.

- The pair that does not receive a transplant is in Center $A$. In this case, if Center $A$ does not submit pairs $a_1$ and $a_2$ to the centralized match, and instead matches them internally to each other, then there is a single multi-center Pareto-efficient matching $\{(a_3, b_4), (b_2, b_4)\}$, and $\phi$ chooses this matching. As a result, Center $A$ succeeds in matching its all three pairs.

- The pair that does not receive a transplant is in Center $B$. In this case, if Center $B$ does not submit pairs $b_3$ and $b_4$ to the centralized match, and instead matches them internally to each other, then there is a single multi-center Pareto-efficient matching $\{(a_1, b_1), (a_2, b_2)\}$, and $\phi$ chooses this matching. As a result, Center $B$ succeeds in matching its all four pairs.

In either case, we showed that there is a center that can successfully manipulate the Pareto-efficient multi-center matching mechanism $\phi$.

Future research in this area involves finding mechanisms that have good incentive and efficiency properties for centers, using different solution and modeling concepts. A recent example of this line of research is by Ashlagi and Roth (2011), who investigate the participation problem using computer science techniques for large populations.

# References


[1] Abdulkadiroğlu, Atila and Tayfun Sönmez (1999) "House Allocation with Existing Tenants." *Journal of Economic Theory*, 88, 233-260.

[2] Abecassis, M. , M. Adams, P. Adams, R. M. Arnold, C. R. Atkins, M. L. Barr, W. M. Bennett, M. Bia, D. M. Briscoe, J. Burdick, R. J. Corry, J. Davis, F. L. Delmonico, R. S. Gaston, W. Harmon, C. L. Jacobs, J. Kahn, A. Leichtman, C. Miller, D. Moss, J. M. Newmann, L. S. Rosen, L. Siminoff, A Spital, V. A. Starnes, C. Thomas, L. S. Tyler, L. Williams, F. H. Wright, and S. Youngner (2000) "Consensus statement on the live organ donor." *Journal of the American Medical Association*, 284, 2919-2926.

[3] Abraham, David J., Avrim Blum, and Tuomas Sandholm (2007) "Clearing Algorithms for Barter Exchange Markets: Enabling Nationwide Kidney Exchanges." In *Proceedings of ACM-EC 2007: the Eighth ACM Conference on Electronic Commerce.*

[4] Ashlagi, Itai and Alvin E. Roth (2011) "Individual rationality and participation in large scale, multi-hospital kidney exchange." Working paper.

[5] Awasthi, P. and Tuomas Sandholm (2009) "Online Stochastic Optimization in the Large: Application to Kidney Exchange." In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI).*

[6] Biró, Peter, and Katarína Cechlárová (2007) "Inapproximability of the kidney exchange problem." *Information Processing Letters*, 101, 199-202.

[7] Biró, Peter and Eric McDermid (2008) "Three-sided Stable Matchings with Cyclic Preferences and the Kidney Exchange." Ulle Endriss and Paul W. Goldberg (eds.) *COMSOC-2008: Proceedings of the 2nd International Workshop on Computational Social Choice*, 97-108.

[8] Bogomolnaia, Anna and Hervé Moulin (2001) "A new solution to the random assignment problem." *Journal of Economic Theory*, 100 (2001), 295-328.

[9] Bogomolnaia, Anna and Hervé Moulin (2004) "Random matching under dichotomous preferences." *Econometrica*, 72, 257-279.

[10] Cechlárová, Katarína, Tamás Fleiner, and David F. Manlove (2005) "The kidney exchange game." *Proceedings of SOR'05: the 8th International Symposium on Operations Research in Slovenia*, 77-83.

[11] Delmonico, Francis L. (2004) “Exchanging Kidneys-Advances in Living-Donor Transplantation.” *New England Journal of Medicine*, 350, 1812-1814.

[12] Dutta, Bhaskar and Debraj Ray (1989) “A Concept of Egalitarianism under Participation Constraints.” *Econometrica*, 57, 615-635.

[13] Edmonds, Jack (1965) “Paths, Trees, and Flowers.” *Canadian Journal of Mathematics,* 17, 449-467.

[14] Gallai, Tibor (1963) “Kritische Graphen II.” *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 8, 373-395.

[15] Gallai, Tibor (1964) “Maximale Systeme unabhängiger kanten.” *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 9, 401-413.

[16] Gjertson, David W. and J. Michael Cecka (2000) “Living Unrelated Donor Kidney Transplantation.” *Kidney International*, 58, 491-499.

[17] Hatfield, John W. (2005) “Pairwise Kidney Exchange: Comment.” *Journal of Economic Theory*, 125, 189-193.

[18] Irving, Robert W. (2007) “The cycle roommates problem: a hard case of kidney exchange.” *Information Processing Letters*, 103, 1-4.

[19] de Klerk, M., K. M. Keizer, F. H. J. Claas, M. Witvliet, B. J. J. M. Haase-Kromwijk, and W. Weimar (2005). The Dutch national living donor kidney exchange program. *American Journal of Transplantation 5*, 2302-2305

[20] Lovász, László and Michael D. Plummer (1986) *Matching Theory*, North-Holland, Amsterdam, New York, Oxford, Tokyo.

[21] Ma, Jinpeng (1994) “Strategy-Proofness and the Strict Core in a Market with Indivisibilities.” *International Journal of Game Theory,* 23, 75-83.

[22] Mandal, A. K., J. J. Snyder, D. T. Gilbertson, A. J. Collins, and J. R. Silkensen (2003) “Does cadavaric donor renal transplantation ever provide better outcomes than live-donor renal transplantation?” *Transplantation*, 75, 494-500.

[23] Montgomery, RA, Gentry SE, Marks WH, Warren DS, Hiller J, Houp J, Zachary AA, Melancon JK, Maley WR, Rabb H, Simpkins CE, Segev DL. (2006) ”Domino paired kidney donation: a strategy to make best use of live non-directed donation.” *Lancet*, 368, 419-421.

[24] Morrill, Thayer (2008) "The Roommates Problem Revisited." *Working paper.*

[25] Opelz, Gerhard (1997) "Impact of HLA Compatibility on Survival of Kidney Transplants from Unrelated Live Donors." *Transplantation,* 64, 1473–1475.

[26] Overbeck, I., M. Bartels, O. Decker, J. Harms, J. Hauss, and J. Fangmann (2005) "Changes in quality of life after renal transplantation." *Transplantation Proceedings 37*, 1618–1621.

[27] Pápai, Szilvia (2000) "Strategyproof Assignment by Hierarchical Exchange." *Econometrica*, 68, 1403-1433.

[28] Park, K., J. H. Lee, K. H. Huh, S. I. Kim, and Y. S. Kim (2004) "Exchange living donor kidney transplantation: Diminution of donor organ shortage." *Transplantation Proceedings 36*, 2949–2951.

[29] Pycia, Marek, and M. Utku Ünver (2009) "A Theory of House Allocation and Exchange Mechanisms." *Working paper.*

[30] Rapaport, F. T. (1986) "The case for a living emotionally related international kidney donor exchange registry." *Transplantation Proceedings*, 18, 5-9.

[31] Rawls, John (1971) A Theory of Justice, Harvard University Press, Cambridge.

[32] Rees, Michael A., Jonathan E. Kopke, Ronald P. Pelletier, Dorry L. Segev, Matthew E. Rutter, Alfredo J. Fabrega, Jeffrey Rogers, Oleh G. Pankewycz, Janet Hiller, Alvin E. Roth, Tuomas Sandholm, M. Utku Ünver, and Robert A. Montgomery (2009) "A Non-simultaneous Extended Altruistic Donor Chain." *The New England Journal of Medicine*, 360, 1096-1101.

[33] Ross, L. F., D. T. Rubin, M. Siegler, M. A. Josephson, J. R. Thistlethwaite, Jr., and E S. Woodle (1997) "Ethics of a paired-kidney-exchange program." *The New England Journal of Medicine*, 336, 1752-1755.

[34] Ross, Laine Friedman, and E. Steve Woodle (2000) "Ethical issues in increasing living kidney donations by expanding kidney paired exchange programs." *Transplantation*, 69, 1539-1543.

[35] Roth, Alvin E. (1982) "Incentive compatibility in a market with indivisibilities." *Economics Letters*, 9, 127-132.

[36] Roth, Alvin E. and Andrew Postlewaite (1977) "Weak versus Strong Domination in a Market with Indivisible Goods." *Journal of Mathematical Economics*, 4, 131-137.

[37] Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2004) "Kidney Exchange." *Quarterly Journal of Economics,* 119, 457-488.

[38] Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2005a) "Pairwise Kidney Exchange." *Journal of Economic Theory,* 125, 151-188.

[39] Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2005b) "A Kidney Exchange Clearinghouse in New England." *American Economic Review Papers and Proceedings*, 95(2): 376-380.

[40] Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2005c) "Transplant Center Incentives in Kidney Exchange." *Unpublished.*

[41] Roth, Alvin E., 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.

[42] Roth, Alvin E., Tayfun Sönmez, M. Utku Ünver, Francis L. Delmonico, and Susan L. Saidman (2006) "Utilizing List Exchange and Nondirected Donation through 'Chain' Paired Kidney Donations." *American Journal of Transportation*, 6, 2694-2705.

[43] Saidman, Susan L., 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, 773—782.

[44] Segev, Dorry, Sommer Gentry, Daniel S. Warren, Brigette Reeb, and Robert A. Montgomery (2005) "Kidney paired donation: Optimizing the use of live donor organs." *Journal of the American Medical Association*, 293, 1883-1890.

[45] Shapley, Lloyd and Herbert Scarf (1974) "On Cores and Indivisibility." *Journal of Mathematical Economics.* 1, 23-28.

[46] Sönmez, Tayfun and M. Utku Ünver (2005) "House Allocation with Existing Tenants: An equivalence." *Games and Economic Behavior*, 52: 153-185.

[47] Sönmez, Tayfun and M. Utku Ünver (2010a) "House Allocation with Existing Tenants: A Characterization." *Games and Economic Behavior*, 69(2), 425-445.

[48] Sönmez, Tayfun and M. Utku Ünver (2011) "Matching, Allocation, and Exchange of Discrete Resources." *J. Benhabib, A. Bisin, and M. Jackson (eds.) Handbook of Social Economics*, Vol 1A. The Netherlands: North-Holland, pp. 781-852.

[49] Sönmez, Tayfun and M. Utku Ünver (2010b) "Altruistic Kidney Exchange." *Working paper.*

[50] Su, Xuanming and Stefanos A. Zenios (2006) "Recipient Choice Can Address the Efficiency-Equity Trade-off in Kidney Transplantation: A Mechanism Design Model" *Management Science*, 52, 1647-1660.

[51] Schweitzer, E. J., A. Wiland, D. Evans, M. Novak, I. Connerny, L. Norris, J. O. Colonna, B. Philosophe, A. C. Farney, B. E. Jarrell, and S. T. Bartlett (1998) "The shrinking renal replacement therapy break-even point." *Transplantation*, 107, 1702–1708.

[52] Ünver, M. Utku (2010) "Dynamic Kidney Exchange." *Review of Economic Studies*, 77 (1), 372-414

[53] Votruba, M. (2002) "Efficiency-equity tradeoffs in the allocation of cadaveric kidneys." Working paper.

[54] Woodle, E. Steve and Lainie Friedman Ross (1998) "Paired Exchanges Should Be Part of the Solution to ABO Incompatibility in Living Donor Kidney Transplantation." *Transplantation*, 66(3), pp. 406–7.

[55] Yilmaz, Özgür (2011) "Kidney Exchange: An Egalitarian Mechanism." *Journal of Economic Theory*, 146(2), 592-618.

[56] Zenios, Stefanos (1996) *Health care applications of optimal control theory.* Ph.D. Thesis, Massachusetts Institute of Technology.

[57] Zenios, Stefanos A. (2002) "Optimal Control of a Paired-Kidney Exchange Program." *Management Science*, 48, 328-342.

[58] Zenios, Stefanos A., Glenn M. Chertow , and Lawrence M. Wein (2000) "Dynamic Allocation of Kidneys to Candidates on the Transplant Waiting List." *Operations Research*, 48, 549-569.

[59] Zenios, Stefanos A., E. Steve Woodle, 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.