Efficient and Incentive-Compatible Liver Exchange*

Haluk Ergin† Tayfun Sönmez‡ M. Utku Ünver§

May, 2018

Abstract

Liver exchange has been practiced in small numbers, mainly to overcome blood-type incompatibility between patients and their living donors. A donor can donate either his smaller left lobe or the larger right lobe, although the former option is safer. Despite its elevated risk, right-lobe transplantation is often utilized due to size-compatibility requirement with the patient. We model liver exchange as a market-design problem, focusing on logistically simpler two-way exchanges. First, with two patient-donor sizes, we introduce an algorithm when only the safer left-lobe transplantation is feasible. We then introduce an individually rational, Pareto-efficient, and incentive-compatible mechanism that truthfully elicits the right-lobe-donation willingness of donors, and finally extend these results to a general model with any number of patient/donor sizes. The generalization requires new technical tools regarding bilateral exchanges under partial-order-induced preferences. Through simulations we show that not only liver exchange can increase the number of transplants by more than 30%, it can also increase the share of the safer left-lobe transplants.

Keywords: Market design, liver exchange, matching, incentive compatibility, efficiency

JEL Codes: D47, C78

---

*Sönmez acknowledges the research support of Goldman Sachs Gives via the Dalinc Ariburnu - Goldman Sachs Faculty Research Fund. Sönmez and Ünver acknowledge the research support of the NSF via award SES #1729778. We thank the participants at Vanderbilt Market Design Conference, NBER Market Design Working Group Meeting, Matching in Practice Workshop at Budapest, Harvard/MIT Theory Workshop, SAET Conference at Faro, Portugal, Conference on Economic Design at University of York, UK, Southern Methodist University, Penn State University, Montreal-CIREQ Advances in Micro Theory Workshop, Advances in Mechanism-Market Design Conference at NYU-Abu Dhabi, and Tsukuba University Global Science Week, Japan.

†Department of Economics, University of California at Berkeley; hie@berkeley.edu

‡Department of Economics, Boston College and Distinguished Research Fellow, Koç University; sonmezt@bc.edu

§Department of Economics, Boston College and Distinguished Research Fellow, Koç University; unver@bc.edu

1 Introduction

Following the kidney, the liver is the second most common organ for transplantation worldwide. In the case of the US, more than 7000 of nearly 31000 organ transplants in 2015 were liver transplants. Transplantation is the only potential treatment for end-stage liver disease, unlike end-stage kidney disease where there is the alternative (although inferior) treatment of dialysis. As in the case of kidneys, transplants from deceased donors and living donors are both possible (and widespread) for liver transplantation.[1] Unlike kidney transplantation, however, a living donor can donate only a part of his liver —henceforth referred as a *lobe*— going through a liver resection operation called *hepatectomy*. Based on the anatomy of the liver, the main options are donating either the smaller left lobe (normally 30–40% of the liver) with a *left hepatectomy* or the larger right lobe (normally 60–70% of the liver) with a *right hepatectomy*. Following the transplantation, the remnant liver of a living donor typically regenerates within a month. Assuming the donor and the patient are blood-type compatible,[2] which of these two options is preferred (or even feasible) depends on the relative liver volumes of the patient and the donor. In order to provide adequate liver function for the patient, at least 40% of the standard liver volume of the patient is required. The metabolic demands of a larger patient will not be met by the smaller left lobe from a relatively small donor. This phenomenon is known as *small-for-size syndrome*. The primary solution to avoid this syndrome has been harvesting the larger right lobe of the liver for transplantation. This procedure, however, involves considerably higher risks for the donor than harvesting the smaller left lobe. While donor mortality is approximately 0.1% for left hepatectomy, it is in the range of 0.4–0.5% for right hepatectomy (Lee, 2010). Furthermore, other significant risks, referred to as donor *morbidity*, are also much higher under right hepatectomy than left hepatectomy. Mishra et al. (2018) reports that the morbidity rates are 28% for right hepatectomy and 7.5% for left hepatectomy. Hence one of the main challenges for living-donor liver transplantation is that the much safer left-lobe transplantation is not a viable option for a majority of patients with willing donors. As an implication, many patients with potential donors cannot receive a transplant since either their donors hesitate to go through the higher-risk right hepatectomy, or their doctors recommend against this procedure. The high risks associated with the right-lobe liver transplantation also affect the public perception of living-donor liver transplantation. The number of annual living-donor liver transplants in the US peaked in 2001 with 524 transplants, increasing eight-fold in the period from 1996

[1] The attitude towards living-donor liver transplantation differs considerably between western countries and Asian countries. In contrast to western countries, donations for liver transplantation in much of Asia come from living donors. For example, in 2015, while only 359 of 7127 liver transplants were from living donors in the US, 942 of 1398 liver transplants were from living donors in South Korea.

[2] Each individual is of one of the following four blood types: O, A, B, or AB. While a blood-type O donor is blood-type compatible with any blood-type patient, a blood-type A donor is blood-type compatible with patients of blood types A and AB, a blood-type B donor is blood-type compatible with patients of blood types B and AB, and a blood-type AB donor is blood-type compatible with only patients of blood type AB.

to 2001. The highly publicized death of a right-lobe liver donor in the US in 2002 not only brought an end to this remarkable increase, but it also resulted in a 40–50% reduction from its peak US peak since then.[3]

As the worldwide shortage of transplant organs keeps increasing annually, living-donor exchanges emerged as an important source for these potentially life-saving resources, especially in the case of kidneys. In its most basic form, a living-donor organ exchange involves two patients with willing donors who exchange donors either because direct donation is not an option due to an immunological barrier or because one or both patients receive a more favorable outcome through the exchange. The concept was originally proposed for kidneys by Rapaport (1986), and it became widespread over the last 15 years with the introduction of optimization and market-design techniques to kidney exchange (Roth, Sönmez, and Ünver, 2004, 2005, 2007). A vast majority of these exchanges are conducted between incompatible kidney patient-donor pairs, where a donor cannot directly donate to his patient due to immunological barriers.[4] Liver exchanges between incompatible patient-donor pairs are also conducted in modest numbers in several Asian countries, most notably in South Korea. Our focus in this paper is the design of a liver-exchange mechanism that not only includes the incompatible pairs, but also a subset of compatible pairs whose only direct-donation option to their patients is through the higher-risk right hepatectomy. Under an efficient and incentive-compatible mechanism we introduce, these compatible pairs participate in exchange only if they strictly benefit from doing so by reducing the risks to their donors through a left hepatectomy. As such, our proposed mechanism not only increases the number of living-donor liver transplants, but also increases the reliance on the lower-risk left-lobe liver transplantation in the spirit of the central tenet of the hippocratic oath "first do no harm."

While the practice of kidney exchange has flourished worldwide over the last fifteen years, inclusion of compatible pairs in exchange pools has proved to be a challenge since benefits to these pairs from joining kidney-exchange pools are either not present or weak. In contrast, the benefits from joining liver-exchange pools can be considerable for a significant fraction of compatible pairs, if it means their donors can have a left hepatectomy rather than a right hepatectomy. And the welfare gains from their inclusion can be potentially very high. Consider a large, blood-type A liver patient, who in the absence of exchange has to receive a right liver lobe from his small, blood-type O donor. While this is a feasible medical procedure, an alternative arrangement of an exchange of donors with a small, blood-type O patient with a large, blood-type A donor will not only significantly reduce the risks to his donor (by replacing the donor's right hepatectomy with a left hepatectomy), but also enable a second patient to receive a potentially life-saving liver transplant. The possibility of offering a less risky procedure to such pairs provides an opportunity to increase the size

[3]See Grady (2002).

[4]For the case of kidney transplantation, these immunological barriers are blood-type incompatibility and tissue-type incompatibility.

of the liver-exchange pool in a way that includes the much-needed blood-type O donors.

In the above example, the large, blood-type A patient with the small, blood-type O donor would only be willing to participate in exchange if the pair benefits from exchange through an assurance that its donor goes through the much less risky procedure of left hepatectomy. However, not all cases are this straightforward. Consider a blood-type A patient with a blood-type B donor. Since it is blood-type incompatible to start with, not only can it benefit from exchange through a left-lobe donation but also through the less-desired right-lobe donation if the pair is willing to expose the donor to the higher mortality and morbidity risks of a right hepatectomy. This possibility is the primary reason why one cannot directly adopt for liver exchange the mechanisms and techniques developed for kidney exchange, unless the higher-risk right hepatectomy is completely avoided. A liver-exchange mechanism has to determine not only which pairs are to be matched with each other to exchange donors, but it shall also determine which donors have to donate their right lobes rather than their left lobes. Of course, some pairs may not be willing to expose their donors to the more risky procedure of right hepatectomy, but a poorly designed exchange mechanism may also give them incentives to hide their willingness to do so even if they are. As such, our focus is not only the design of an efficient mechanism, but at the same time the design of an *incentive-compatible* liver exchange mechanism where a pair never receives a less favorable outcome by revealing its willingness to go through the less desired right hepatectomy.

For living-donor liver transplantation, size compatibility is a requirement in addition to blood-type compatibility. A patient is size compatible with a donor if the volume of liver tissue transplanted from the donor is at least 40% of the patient's standard liver volume. Based on this requirement, we define the *size of a patient* to be 40% of her standard liver volume, and the *size of a donor* to be the volume of the left lobe of his liver. Hence, assuming they are blood-type compatible, a patient can receive a left-lobe liver transplant from a donor who is at least her own size and a right-lobe liver transplant from a smaller donor. A patient-donor type is thus characterized by the blood types of the patient and donor, along with their respective sizes. The key types in the design of an efficient and incentive-compatible mechanism are those who can participate in exchange both through left-lobe donation as well as through the less-preferred right-lobe donation. The challenge is determining when the donors of a particular type shall be considered for a right-lobe donation rather than a left-lobe donation. We refer to this process as a *transformation* of a type. To assure incentive compatibility, a type should be transformed only after their left-lobe-exchange possibilities are exhausted so that their announcement of whether they are willing for their donors to go through a right hepatectomy does not affect whether or not their donors go through the safer left hepatectomy. One simple approach might be first considering all such pairs for left-lobe donation and then transforming them simultaneously once their left-lobe-donation possibilities are exhausted. There are two problems with this simple approach. First, it is possible that an exchange between two such pairs might be possible with the transformation of only one of these pairs, say pair 1. If so, transforming both pairs and matching them for an exchange will result in a Pareto-inferior outcome. Second, this possibility might encourage pair 1 to hide its willingness for a right-lobe donation of its donor. Hence, key in design is determining the order in which pairs of these critical types shall be transformed. We show that there is a well-defined ordering, which assures that the resulting mechanism is not only Pareto efficient, but also incentive compatible. While we show this result for an arbitrary number of potential sizes of patients and donors, making our mechanism practically relevant, the intuition of the mechanism is more clear for the case of two sizes. Hence we develop first a Pareto-efficient and incentive-compatible mechanism for this simpler case, providing a geometric representation of our mechanism, and then extend our analysis to an arbitrary number of sizes. We also illustrate the potential gains from adopting our proposed mechanism on simulated pools based on South Korean population and transplantation characteristics. We observe that, depending on the willingness of donors to participate in right-lobe donation, 44% to 34% more transplants can occur due to liver exchange than by direct donation.

1.1 Related Literature

Kidney exchange, as an application of market design, was initiated by Roth, Sönmez, and Ünver (2004, 2005, 2007). Recent developments in market design for kidney exchanges include studies on incentivizing compatible pairs to participate in exchange (Nicolò and Rodriguez-Álvarez, 2017; Sönmez, Ünver, and Yenmez, 2018), using kidney exchange along with ABO-blood-type-incompatible kidney transplants (Andersson and Kratz, 2017), and designing an incentive-compatible participation scheme for transplant centers in kidney exchange (Agarwal et al., 2018).

Unlike the growing literature on kidney exchange, there are only a handful papers on liver exchange. These include Hwang et al. (2010) and Chan et al. (2010), both of which demonstrate the proof of concept for liver exchange, and Mishra et al. (2018), which advocates for organized liver exchange in the US. Dickerson and Sandholm (2014) advocates for trans-organ exchange, where a donor associated with a kidney recipient donates a liver lobe and a donor associated with a liver recipient donates a kidney, whereas Samstein et al. (2018) explores some of the ethical concerns this practice might encounter, including unbalanced donor risks. Ergin, Sönmez, and Ünver (2017) studies dual-donor organ exchange, where each patient receives organs from two living donors. Dual-graft liver exchange, where each patient participates in exchange with two left-lobe donating donors, is an application of this model. Although dual-graft liver transplantation is practiced in a few countries, including South Korea and China, overcoming size incompatibility through a right-lobe transplantation is far more common throughout the world. And while the difference between the mortality and morbidity risks of right lobe vs. left lobe donation is well established in the transplantation literature, our main focus, the design implications of these two main liver transplantation technologies, is not considered in any of the papers on liver exchange.

In terms of modeling, there is a conceptual similarity between our liver-exchange model and the "matching with contracts" model of Hatfield and Milgrom (2005), which extends two-sided matching problems (Gale and Shapley, 1962) by allowing various contractual arrangements between the two sides. While left-lobe donation and right-lobe donation can be interpreted as two different contractual arrangements, unlike the matching with contracts model, our model is one sided. Hence the cumulative offer mechanisms introduced for the matching with contracts model by Hatfield and Milgrom (2005) and extended by Hatfield and Kojima (2010) is not applicable in our framework.

More broadly, our paper contributes to a very diverse list of market-design applications, including entry-level labor markets (Roth and Peranson, 1999), spectrum auctions (Milgrom, 2000), internet auctions (Edelman, Ostrovsky, and Schwarz, 2007; Varian, 2007), school choice (Abdulkadiroğlu and Sönmez, 2003), course allocation (Sönmez and Ünver, 2010; Budish and Cantillon, 2012), affirmative action (Kojima, 2012; Hafalir, Yenmez, and Yildirim, 2013; Echenique and Yenmez, 2015), refugee matching (Moraga and Rapoport, 2014; Jones and Teytelboym, 2017; Delacrétaz, Kominers, and Teytelboym, 2017), and assignment of airport landing slots (Schummer and Vohra, 2013; Schummer and Abizada, 2017).

2 A Model of Dual Technology Liver Transplantation

There are two liver transplantation technologies: A donor can donate his *left liver lobe* or his *right liver lobe* for a transplant.

2.1 Size Compatibility

The left lobe can be anywhere in the range from 30% to 40% of the donor's liver volume, and the right lobe makes up the rest. A patient requires a liver graft that is at least 40% of her own liver volume. Based on this constraint, the size of an individual will have a different meaning throughout the paper for a donor and a patient. Formally, the *size of a donor* is the volume of his left liver lobe, whereas the *size of a patient* is the minimum required volume of liver tissue for a transplant, i.e., 40% of the volume of her (dysfunctional) liver. Therefore, for left-lobe transplantation, a donor is **size compatible** with a patient if his (left liver lobe) size is at least as large as the patient's (minimum volume of required liver tissue) size. With a slight abuse of language, a donor who is size compatible with a patient will be referred as a donor who is at least as large as the patient. Let $\mathcal{S} = \{0, 1, \ldots, S-1\}$ denote the set of possible sizes, where $S \geq 1$ is the number of possible sizes. Here the larger numbers correspond to larger sizes.[5]

[5]Set $\mathcal{S}$ can be specific to each liver-exchange pool, allowing for a continuum of sizes generically as long as each pool we analyze is finite.

2.2 Blood-type Compatibility

The blood type of an individual is determined by the availability or the lack of two antigens referred to as antigen A and antigen B. An individual of blood type O has neither antigen, an individual of blood type A has only antigen A, an individual of blood type B has only antigen B, and an individual of blood type AB has both antigens. A donor is **blood-type compatible** with a patient if he does not have a blood antigen the patient lacks. That means a blood-type O donor (having neither antigen) is blood-type compatible with patients of all blood types, a blood-type A donor is blood-type compatible with patients of blood types A and AB, a blood-type B donor is blood-type compatible with patients of blood types B and AB, and a blood type AB donor is blood-type compatible with patients of only blood type AB. Let $\mathcal{B} = \{O, A, B, AB\}$ denote the set of blood types.

2.3 Left-Lobe-Donation Relation & Its Equivalent Representation

We assume that the blood type and the size of each individual are observable physical attributes, and we refer to $\mathcal{B} \times \mathcal{S}$ as the set of individual types. A donor can donate his left lobe to a patient if and only if he is both blood-type compatible with and as large as (or equivalently size compatible with) the patient. Let $\trianglerighteq$ denote the **left-lobe-donation partial order** on $\mathcal{B} \times \mathcal{S}$. Note that for any number of sizes S, the two partially ordered sets $(\mathcal{B} \times \mathcal{S}, \trianglerighteq)$ and $(\{0,1\}^2 \times \{0, 1, \ldots, S-1\}, \geq)$ are order isomorphic, where the order isomorphism associates each individual type $T \in \mathcal{B} \times \mathcal{S}$ with the following vector $X \in \{0,1\}^2 \times \{0, 1, \ldots, S-1\}$:

$$\begin{aligned} X_1 = 0 &\;\Leftrightarrow\; T \text{ has the } A \text{ antigen} \\ X_2 = 0 &\;\Leftrightarrow\; T \text{ has the } B \text{ antigen} \\ X_3 = s &\;\Leftrightarrow\; T \text{ is of size } s \end{aligned}$$

Let type $T \in \mathcal{B} \times \mathcal{S}$ be associated with vector $(X_1, X_2, X_3) \in \{0,1\}^2 \times \{0, 1, \ldots, S-1\}$, and type $T' \in \mathcal{B} \times \mathcal{S}$ be associated with vector $(X_1', X_2', X_3') \in \{0,1\}^2 \times \{0, 1, \ldots, S-1\}$. Due to the above described order isomorphism, we have

$$T \trianglerighteq T' \Leftrightarrow (X_1, X_2, X_3) \geq (X_1', X_2', X_3')$$

For convenience, we will work with the equivalent representation $\mathcal{T} = \{0,1\}^2 \times \{0, 1, \ldots, S-1\}$ and $\geq$ of individual types: For any $X, Y \in \mathcal{T}$, a donor of type Y can **donate his left lobe** to a patient of type X if $X \leq Y$. Observe that this is equivalent to the donor being both blood-type compatible with and as large as the patient. All of our conclusions can be rephrased in terms of blood types and sizes using the order isomorphism described above. For the case of two sizes ($S = 2$), Figure 1 illustrates the left-lobe-donation partial order $\trianglerighteq$ on $\mathcal{B} \times \mathcal{S}$, and the standard partial order $\geq$ over the corners of the three-dimensional cube $\{0,1\}^3$.

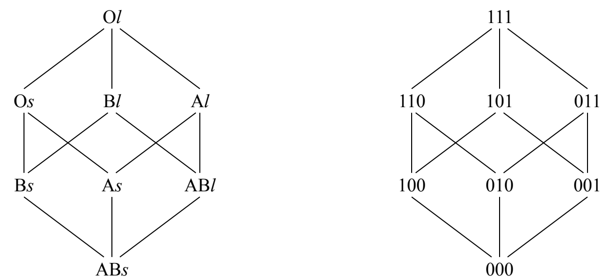


Figure 1: The Partially Ordered Sets $(\mathcal{B} \times \mathcal{S}, \trianglerighteq)$ and $(\{0,1\}^3, \geq)$.

2.4 Right-Lobe Donation

The right liver lobe is considerably larger than the left liver lobe. Therefore the right-lobe-transplantation technology, although involving higher risks for the donor, enables donors to be able to donate to patients who are much larger then them. We will model this technology though a function $\rho(\cdot)$, where $\rho(s)$ denotes the maximum size of a patient who can receive a right-liver-lobe transplant from a donor of size s. Formally, a **right-lobe size function** is a function $\rho : \mathcal{S} \to \mathcal{S}$ such that:

1. ρ is non-decreasing, and
2. $\rho(s) > s$ for all $s \in \mathcal{S} \setminus \{S-1\}$.

Let $\rho(Y) := Y_1 Y_2 \rho(Y_3)$ for any type $Y \in \mathcal{T}$.

A size $s \in \mathcal{S}$ donor can donate his right lobe to blood-type compatible patients who are of size $\rho(s)$ or smaller: Formally, for any $X, Y \in \mathcal{T}$, a donor of type Y can *donate his right lobe* to a patient of type X if $X \leq \rho(Y)$. Since the right lobe is larger than the left lobe, the right-lobe-donation technology increases the set of potential exchanges and direct donations. However, because it involves higher risks for the donor, it is less preferred than left-lobe donation. Therefore, for donors who can feasibly donate their left lobes to a patient, we assume that right-lobe donation is not a viable option. Thus, we say that a patient of type X and a donor of type Y are **left-lobe compatible** if they are blood-type compatible and the donor is at least as large as the patient, i.e., $X \leq Y$. A patient of type X and a donor of type Y are **right-lobe-only compatible** if the donor can donate his right lobe to the patient, but not his left lobe, i.e., $X \leq \rho(Y)$ and $X \nleq Y$.

Throughout the paper, we fix a right-lobe size function ρ.

3 Liver Exchange

As in kidney exchange, the number of living-donor liver transplants can be increased through exchange of donors. Throughout the paper we assume that, in addition to direct transplants, only two-way exchanges are feasible.[6]

3.1 Liver-Exchange Pool

Each patient participates in liver exchange with one donor. A patient and her donor are referred to as a **pair**.[7] The observable characteristics of a pair are summarized by an ordered pair of individual types $X - Y \in \mathcal{T} \times \mathcal{T}$, where X denotes the type of the patient and Y denotes the type of the donor; $X - Y$ is called the **pair type**.

A **liver-exchange pool** is a tuple $\mathcal{E} = (\mathcal{I}, \tau)$ where

1. $\mathcal{I} = \{1, 2, \ldots, I\}$ is a nonempty finite set of patient-donor pairs, and
2. $\tau : \mathcal{I} \to \mathcal{T} \times \mathcal{T}$ is a function such that for every pair $i \in \mathcal{I}$:
   (a) $\tau(i) = \tau_P(i) - \tau_D(i)$.[8]
   (b) $\tau_P(i) \in \mathcal{T}$ is the type of the patient of the pair i.
   (c) $\tau_D(i) \in \mathcal{T}$ is the type of the donor of the pair i.

Throughout the paper, we fix a liver-exchange pool $(\mathcal{I}, \tau)$.

3.2 Preferences & Willingness to Donate a Right Lobe

We interpret a patient-donor pair as a single agent in our model, and thus preferences refer to preferences of the pair. Patients or donors do not have preferences of their own. Preferences only depend on (1) whether the patient receives a transplant or not, (2) whenever the patient receives a transplant from her own donor or through exchange, and (3) which liver lobe the donor has to donate for his patient to receive a transplant. More specifically, we assume a very simple preference structure with the following features:

1. A pair always prefers its donor to donate his left robe rather than his right lobe. This feature precedes any other feature in preference formation.
2. A pair prefers a direct transplant to an exchange, provided that the donor donates the same lobe. Together with feature 1, this means a pair prefers an exchange to a direct transplant if and only if the donor donates his right lobe in the direct transplant and his left lobe in the exchange.
3. A pair is indifferent between all exchanges where the donor donates the same lobe.

There is only one source of potential preference heterogeneity between two pairs of the same type. Right-lobe donation involves significantly higher mortality and morbidity risks for the donor. Hence, some donors may not be willing to donate their right lobes at all.[9] We refer to pairs with such donors as **unwilling ($\boldsymbol{u}$)**. Pairs whose donors are open to the possibility

[6] All liver exchanges reported in the literature as of March 2018 are between two patients and their donors.
[7] We use pronouns "she" for a patient, "he" for a donor, and "it" for a pair.
[8] We refer to a pair type as $X - Y$ instead of (X, Y) as a convention.
[9] If a donor is not willing to donate his safer left lobe, then he never volunteers as a donor.

of right-lobe donation, on the other hand, are referred to as **willing ($\boldsymbol{w}$)**. Whether a pair is willing or unwilling is its private information, and it uniquely determines its preferences. For a willing pair i, the possible outcomes are ranked in the following order under its **willing preference relation** R_i^w:

1. A direct transplant where the donor donates his left lobe.
2. Any exchange where the donor donates his left lobe.
3. A direct transplant where the donor donates his right lobe.
4. Any exchange where the donor donates his right lobe.
5. Patient receives no transplant.

For an unwilling pair i, in contrast, the possible outcomes are ranked in the following order under its **unwilling preference relation** R_i^u**:**

1. A direct transplant where the donor donates his left lobe.
2. Any exchange where the donor donates his left lobe.
3. Patient receives no transplant.
4. A direct transplant where the donor donates his right lobe.
5. Any exchange where the donor donates his right lobe.

3.3 Outcome of the Problem: A Matching

Which direct transplants and exchanges are feasible or not depends on the willingness profile. For a given willingness profile, let $\mathcal{J}_w \subseteq \mathcal{I}$ denote the subset of willing pairs. Then, the set of all feasible matches $E_c[\mathcal{J}_w]$ is given as: For all $i, j \in \mathcal{I}$,[10]

$$\{i,j\} \in E_c[\mathcal{J}_w] \quad \Longleftrightarrow \quad \left\{ \begin{array}{cc} \tau_P(i) \leq \rho(\tau_D(j)) & \text{if } j \in \mathcal{J}_w \\ \tau_P(i) \leq \tau_D(j) & \text{if } j \notin \mathcal{J}_w \end{array} \right\} \& \left\{ \begin{array}{cc} \tau_P(j) \leq \rho(\tau_D(i)) & \text{if } i \in \mathcal{J}_w \\ \tau_P(j) \leq \tau_D(i) & \text{if } i \notin \mathcal{J}_w \end{array} \right\}$$

The **compatibility graph** is defined as $G_c[\mathcal{J}_w] = (\mathcal{I}, E_c[\mathcal{J}_w])$.[11]

A **matching** $M \subseteq E_c[\mathcal{J}_w]$ of the compatibility graph $G_c[\mathcal{J}_w]$ is such that for any $e, e' \in M$, $e \cap e' \neq \emptyset \implies e = e'$, i.e., no pair participates in two distinct exchanges or direct transplants. Let $\mathcal{M}_c[\mathcal{J}_w]$ be the set of matchings given the compatibility graph $G_c[\mathcal{J}_w]$.

We denote the match of pair $i \in \mathcal{I}$ in matching $M \in \mathcal{M}_c[\mathcal{J}_w]$ as $M(i)$. If $M(i) = i$ (i.e., $\{i\} \in M$), then the pair participates in a **direct transplant**. If $M(i) = j$ for some $j \neq i$ (i.e., $\{i, j\} \in M$), then pairs i and j participate in a **(two-way) exchange**. If $M(i) = \emptyset$ (i.e., there is no $e \in M$ such that $i \in e$), then pair i remains **unmatched**.

Consider a match $\{i, j\}$ (with possibly $i = j$) in a matching M. Since a donor only

[10]Observe that this definition allows for a *loop* $\{i, i\} = \{i\}$ to be in $E_c[\mathcal{J}_w]$. This depicts that the donor of pair i can donate to the patient of the pair.

[11]Graph theoretical preliminaries are stated formally in Appendix D.1. Some of our current definitions are restated for general graphs in this appendix, as well.

donates a right lobe to a patient when his left lobe is too small for her, which lobe is donated by either donor is uniquely determined by the match $\{i, j\}$. The same argument also holds for the entire matching M. The following function is helpful to keep track of which lobe is donated in any potential match. Define the **transplant type** function $t : \mathcal{I} \times \mathcal{I} \to \{l, r, \emptyset\}$ as follows: For any $(i, j) \in \mathcal{I} \times \mathcal{I}$,

$$t(i,j) = \begin{cases} l & \text{if } \tau_P(j) \leq \tau_D(i) \\ r & \text{if } \tau_P(j) \nleq \tau_D(i) \ \& \ \tau_P(j) \leq \rho(\tau_D(i)) \\ \emptyset & \text{otherwise} \end{cases}$$

For any two pairs i and j, the transplant type function determines whether the donor of the first pair i and the patient of the second pair j are left-lobe compatible (l), right-lobe-only compatible (r), or incompatible ($\emptyset$). Let $t(i) = t(i, i)$ for all $i \in \mathcal{I}$.

To summarize, each matching M is a collection of direct transplants and exchanges, and together with the function $t(\cdot)$, it also uniquely specifies which liver lobe is donated by each donor: For any $\{i\} \in M$, the pair i engages in a *direct left-lobe transplant* if $t(i) = l$ and in a *direct right-lobe transplant* if $t(i) = r$. Similarly, for any $\{i, j\} \in M$,

- the pairs i and j engage in a two-way exchange,
- the donor of i donates his left lobe if $t(i, j) = l$ and his right lobe if $t(i, j) = r$, and
- the donor of j donates his left lobe if $t(j, i) = l$ and his right lobe if $t(j, i) = r$.

The preferences introduced in Subsection 3.2 can be directly extended to the set of matchings. We will abuse notation and also let R_i denote the induced preference over all matchings $\mathcal{M}_c[\mathcal{I}]$ defined through:

$$M R_i M' \iff M(i) R_i M'(i).$$

Let $\mathcal{R}_i = \{R_i^w, R_i^u\}$ denote the set of possible preferences of pair i. For each $R_i \in \mathcal{R}_i$, let P_i, denote the asymmetric part of R_i. Finally, let $\mathcal{R} = \mathcal{R}_1 \times \ldots \times \mathcal{R}_I$ denote the set of preference profiles.

3.4 Mechanisms and Axioms

Although the types of the participating pairs are observable, their preferences (or equivalently their willingness for a right-lobe donation) are not. A (direct) mechanism determines a matching as a function of the reported preference profile. Given $R \in \mathcal{R}$ and a pair i with type $X - Y$, we will sometimes integrate its preference (or equivalently willingness) information with its type, referring to its type as $X - Yw$ if $R_i = R_i^w$ and as $X - Yu$ if $R_i = R_i^u$.

Since we fix an exchange pool $(\mathcal{I}, \tau)$ throughout, we define a **mechanism** as a function $f : \mathcal{R} \to \mathcal{M}_c[\mathcal{I}]$.

A matching is individually rational if no pair i has an incentive to leave its match in order to stay unmatched nor (if possible) to arrange for a direct transplantation. Formally, a matching $M \in \mathcal{M}_c[\mathcal{I}]$ is **individually rational (IR) at** a preference profile $R \in \mathcal{R}$ if for every $i \in \mathcal{I}$: $M(i) R_i \emptyset$, and if $t(i) \neq \emptyset$ then $M(i) R_i i$. A mechanism f is **individually rational (IR)** if $f(R)$ is individually rational at R for any $R \in \mathcal{R}$.

We next give the definitions of Pareto efficiency and (dominant-strategy) incentive compatibility, which are both standard.

A matching $M \in \mathcal{M}_c[\mathcal{I}]$ is **Pareto efficient (PE) at** a preference profile $R \in \mathcal{R}$ if there does not exist a matching $M' \in \mathcal{M}_c[\mathcal{I}]$ such that $M' R_i M$ for all $i \in \mathcal{I}$ and $M' P_i M$ for some $i \in \mathcal{I}$. A mechanism f is **Pareto efficient (PE)** if $f(R)$ is Pareto efficient at R for any $R \in \mathcal{R}$.

A mechanism f is **incentive compatible (IC)** if for all $i \in \mathcal{I}$, $R_{-i} \in \prod_{j \neq i} \mathcal{R}_j$ and $R_i, \hat{R}_i \in \mathcal{R}_i$:

$$f(R_i, R_{-i}) R_i f(\hat{R}_i, R_{-i}).$$

4 Preliminary Results

In order to ease the presentation of our mechanisms and main results, in this section we present three preliminary results. Our first result characterizes the types of pairs which can exchange donors in an individually rational matching.

Lemma 1 *In any individually rational matching, a pair of type $X{-}Y \in \mathcal{T} \times \mathcal{T}$ can participate in an exchange*

1. *by donating a left lobe only if $X \ngeq Y$ and $Y \ngeq X$, and*
2. *by donating a right lobe only if $X \ngeq \rho(Y)$ and $\rho(Y) \ngeq X$.*

We continue by presenting a lemma that partitions the set of pair types into seven categories. The categorization is organized in terms of whether, in an individually rational matching, a pair can be matched, and, if so, how it can be matched: (a) via a direct left-lobe donation, (b) via a direct right-lobe transplantation, (c) by taking part in an exchange by donating a left lobe, or (d) by taking part in an exchange by donating a right lobe. For pairs in some of these categories (0, I, and II), individual rationality along with the preference profile completely pins down whether or not they are matched, and if so how. For pairs in the remaining categories (III, IV, V, and VI), individual rationality together with the preference profile merely rules out some of the possibilities.

Lemma 2 *Fix a preference profile $R \in \mathcal{R}$ and a matching $M \in \mathcal{M}_c[\mathcal{I}]$. The matching M is individually rational at R if and only if for any pair of any type $X - Y \in \mathcal{T} \times \mathcal{T}$:*

0. *If $X > \rho(Y)$, then the pair remains unmatched.*

I. *If $X \leq Y$, then the pair directly donates a left lobe.*

II. *If $Y < X \leq \rho(Y)$, then either the pair is willing and directly donates a right lobe, or it is unwilling and remains unmatched.*

III. *If $X \not\leq \rho(Y)$, $X \not\geq Y$, & $Y = \rho(Y)$, then either the pair participates in an exchange by donating a left lobe, or it remains unmatched.*

IV. *If $X > Y$, $X \not\geq \rho(Y)$, & $X \not\leq \rho(Y)$, then either the pair is willing and participates in an exchange by donating a right lobe, or it remains unmatched.*

V. *If $X \not\leq \rho(Y)$, $X \not\geq Y$, & $Y < \rho(Y)$, then either the pair participates in an exchange by donating a left lobe, or it is willing and participates in an exchange by donating a right lobe, or it remains unmatched.*

VI. *If $X < \rho(Y)$, $X \not\geq Y$, & $X \not\leq Y$, then the pair participates in an exchange by donating a left lobe, or it is willing and directly donates a right lobe, or it is unwilling and remains unmatched.*

From now on, we will refer to pair types in Lemma 2 as Category 0–VI types. Similarly, any pair whose type lies in a given category will be referred to as a member of that category.

Note that the only pairs which could be part of an exchange by donating either the left lobe or the right lobe are of Category V. This makes the handling of Category V pairs of special importance. Much of the analytical challenges and innovations in our analysis relate to this category.

The next lemma provides a characterization of incentive compatibility for individually rational mechanisms. Specifically, an individually rational mechanism is incentive compatible if and only if, for any pair, whether it participates in an exchange by donating a left lobe is independent of its preferences.

Lemma 3 *Let f be an individually rational mechanism. Then, f is incentive compatible if and only if for all $i \in \mathcal{I}$, and $R_{-i} \in \prod_{j \neq i} \mathcal{R}_j$, the following equivalence holds:*

$$\text{Under } f(R_i^w, R_{-i}),\ i \text{ participates in an exchange by donating the left lobe.}$$

$$\Updownarrow$$

$$\text{Under } f(R_i^u, R_{-i}),\ i \text{ participates in an exchange by donating the left lobe.}$$

Note that by Lemma 2, in an individually rational matching, the only pairs that could be part of an exchange by donating a left lobe are of Category III, V, and VI. As a result, an individually rational mechanism is incentive compatible if for any Category III, V, or VI pair i, whether i participates in an exchange by donating a left lobe is independent of i's willingness announcement.

5 Analysis of a Simplified Model with Two Sizes

Throughout this section, we focus on a simplified model in which there are only two sizes $\mathcal{S} = \{0, 1\}$. While the analysis in Section 6 is more general with an arbitrary number of sizes, much of the intuition can be conveyed in a simpler model that is lighter in both notation and also graph theoretic preliminaries. The algorithms we introduce for this simplified model also have very intuitive graphical representations that we lack for the general model. As such, we believe this version of the model and its results are of independent interest. However, it is possible to skip this section and proceed directly to the general analysis in Section 6, without losing the continuity of the presentation.

5.1 Baseline Model: Two Sizes & Left-Lobe-Only Transplantation

We start our analysis of the two-size model for an even simpler baseline case, when only the less risky left-lobe transplantation is feasible. This model serves as an introduction to our more general model and is also of independent theoretical interest, as it is a symmetric version of our model where all 3 dimensions of an individual type can have at most 2 different, binary values: $\mathcal{T} = \{0, 1\}^3$. This symmetry turns out to simplify many of the graph theoretical complications. As we already emphasized, the compatibility relationship is given as a lattice on the 3-dimensional binary cube $\mathcal{T} = \{0, 1\}^3$. The order isomorphism between $\mathcal{B} \times \mathcal{S}$ and $\mathcal{T}$ are depicted in Figure 1.

Throughout this subsection, we assume that the more risky right-lobe transplantation is not feasible, and thus, the set of pairs that can participate in right-lobe donation is $\mathcal{J}_w = \emptyset$. Therefore, the compatibility graph we consider is $G_c[\emptyset] = (\mathcal{I}, E_c[\emptyset])$.

Since right-lobe donation is ruled out, not only does incentive compatibility become redundant, the notion of individual rationality is also simplified in this subsection. A matching $M \in \mathcal{M}_c[\emptyset]$ is **individually rational** if $\{i\} \in E_c[\emptyset]$ then $\{i\} \in M$. That is, in individually rational left-lobe-only matchings, left-lobe-compatible pairs are matched only through direct transplants. Hence exchange is possible only for pairs that are not left-lobe compatible.

If an individually rational matching in $\mathcal{M}_c[\emptyset]$ is Pareto inefficient, then there exists another individually rational matching in $\mathcal{M}_c[\emptyset]$ that matches a strict superset of pairs with respect to M. In the absence of right-lobe transplantation, the mathematical structure of the problem becomes a special case of a *maximum matching problem*, and for these problems it is well known that every individually rational and Pareto-efficient matching in $\mathcal{M}_c[\emptyset]$ matches the same number of pairs (see for example, *matching matroid* in Korte and Vygen, 2011). Thus, as a corollary, we have an individually rational matching of $\mathcal{M}_c[\emptyset]$ is maximum if and only if it is Pareto efficient. For this reason (and only when right-lobe transplantation is ruled out), we will sometimes refer to an individually rational and Pareto-efficient matching as a **maximum** individually rational matching.

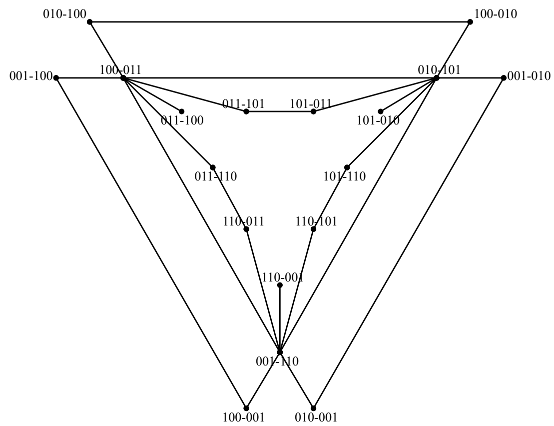


Figure 2: Possible left-lobe-only (two-way) exchanges among pair types when $S = 2$.

Due to Lemma 1, the pair types that could participate in a left-lobe-only exchange in an individually rational matching are $X - Y \in \mathcal{T} \times \mathcal{T}$ such that $X \ngeq Y$ and $X \nleq Y$. All possible exchanges are depicted in Figure 2. The edges of the graph in this figure are all individually rational left-lobe-only exchanges between pair types with two individual sizes.

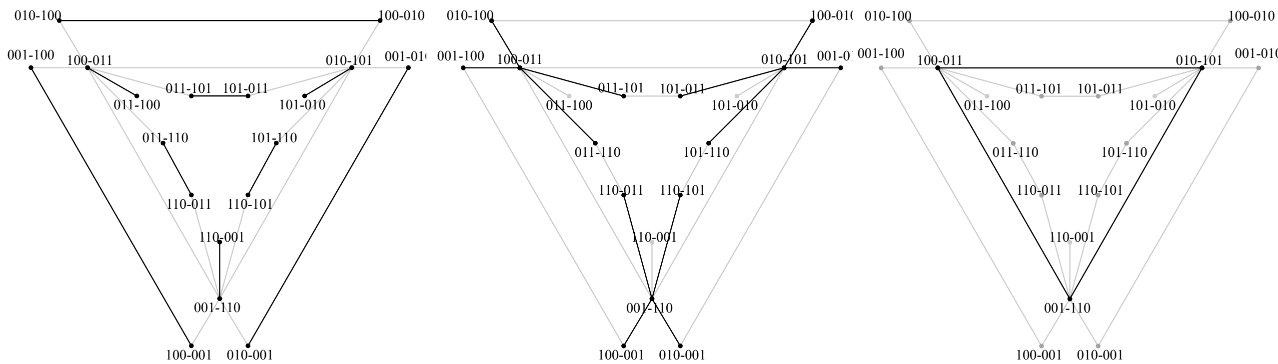
Our goal in this section is to construct an *almost-greedy matching algorithm* that finds a maximum individually rational matching. A greedy algorithm prioritizes exchanges/direct transplants, regardless of which pairs are in the exchange pool, and clears them in order of prioritization, as long as such a direct transplant/exchange is feasible among the remaining pairs.[12] For general graphs, maximum individual matchings cannot always be constructed greedily (see Edmonds, 1965). It turns out this is not the case for left-lobe-only liver exchange with two sizes. This approach has the additional advantage that we can state the number of transplants in a maximum individually rational matching as a simple additive formula of the pool characteristics. Moreover, it turns out to be quite transparent.

We introduce two additional concepts to this end.

The **value** of a pair type $X - Y \in \mathcal{T} \times \mathcal{T}$ is $v(X - Y) = \sum_{p=1}^{3}(Y_p - X_p)$. We refer to this variable as a *value* since it gives the number of valuable donor characteristics net of patient's, and these can be interpreted as the net asset of a pair in finding a match in an exchange. For example, a small patient with a large donor, or a patient who has the A antigen (i.e., a blood-type A or AB patient) with a donor who lacks this antigen (i.e., a blood-type O or B donor) are both bringing a net positive asset to an exchange. An exchange can be interpreted as a trade of these three assets between the pairs. By Lemma 1, only pairs with a value of -1, 0, or 1 can be part of an exchange.[13]

[12]The algorithm will be *almost* greedy, as in the last step we will additionally need the information regarding the number of remaining pairs of certain types to execute it.

[13]More specifically, of the three characteristics, at least one of them shall have a net value of -1 and at least one shall have a net value of 1. If the former fails, the pair is compatible, and if the latter fails, the pair has no feature of value to offer for an exchange. The sum of the net values of these two characteristics is 0,



(a) Step 1 **(b)** Step 2 **(c)** Step 3

Figure 3: Exchanges carried out in Steps 1,2,3

To have a feasible exchange of types $U - V$ & $X - Y$, we need $V \geq X$ and $Y \geq U$. Thus, as a necessary condition for a feasible exchange, the sum of values of the two pairs needs to be at least zero. And when it is exactly zero, each of the three features is fully utilized in the exchange. That means no small patient receives a transplant from a large donor, no patient who has antigen A receives a transplant from a donor who lacks antigen A, or no patient who has antigen B receives a transplant from a donor who lacks antigen B. Thus, any total value excess of zero can be interpreted as the *waste of the exchange*. Formally, the **waste** of an exchange of pair types $U - V$ & $X - Y$ is $v(U - V) + v(X - Y)$. As seen in Figure 2, all possible exchanges are of waste 0, 1, or 2. We refer to an exchange with a waste of $k \in \{0, 1, 2\}$ as an exchange of **k-waste**.

Through these variables, value and waste, we reduce the six dimensions of the type of a pair and twelve dimensions in an exchange, respectively, to two single-dimensional summary variables. Thus, in theory, we are losing some important information regarding the pair types and exchanges that could be valuable in determining maximum individually rational matchings. However, it turns out that the waste of each exchange is all we need to determine a maximum individually rational matching: We can order the wastes of exchanges from the smallest to the largest and clear them in order.

Thus, the main innovation of the algorithm introduced below is the use of the wastes of exchanges as sufficient information to prioritize among them.

The two-size left-lobe-only sequential matching algorithm:

Step 0a. Fix a priority order over all pairs.

Step 0b. Clear direct transplants: All compatible pairs (i.e., pairs with types $X - Y$ such that $X \leq Y$) participate in direct transplants.

and the third one is flexible with possible values of -1, 0, or 1.

Step 1. Clear 0-waste exchanges: For each $X - Y \in \mathcal{T} \times \mathcal{T}$, match the highest-priority pair of type $X - Y$ with the highest-priority pair of its reciprocal type $Y - X$. Proceed in a similar way, matching the next highest-priority pairs of types $X - Y$ and $Y - X$ with each other, and so on, until one of the two types is exhausted.

Step 2. Clear 1-waste exchanges: For each of the three value 1 types $X - Y \in \{100 - 011, 010 - 101, 001 - 110\}$, match the highest-priority pair of type $X - Y$ with the highest-priority pair of its value 0 neighbor. Proceed in a similar way, matching the next highest-priority pair of type $X - Y$ with the next highest-priority pair of its value 0 neighbor, and so on, until either type $X - Y$ or their 0 value neighbors are exhausted.

Step 3. Clear 2-waste exchanges: Match the maximum number of pairs of value 1 types $100-011$, $010-101$, and $001-110$ among each other, following the given priority order.[14]

Figure 3 graphically illustrates the exchanges that are maximized at each step of the sequential matching algorithm. The next theorem states the optimality of this algorithm. It is straightforward to compute the maximum number of transplants as a simple formula through the algorithm.

Theorem 1 *The two-size left-lobe-only sequential matching algorithm maximizes the number of exchanges and the number of transplants in any individually rational matching.*

Prioritizing 0-waste exchanges over 1-waste exchanges and 1-waste exchanges over 2-waste exchanges is intuitively very plausible. Recall that each patient and each donor are a collection of three features: their size and the presence of antigens A and B. In a 0-waste exchange, no patient receives a transplant with any feature that is more valuable than the transplant she needs. For example, no small patient receives a transplant from a large donor, or no patient who has antigen A receives a transplant from a donor who lacks this antigen. Hence,

[14] The maximum matching is found following the given priority order in this case as follows: Let type set $\{X_1 - Y_1, X_2 - Y_2, X_3 - Y_3\} = \{100 - 011, 010 - 101, 001 - 110\}$ be such that $n^*(X_1 - Y_1) \geq n^*(X_2 - Y_2) \geq n^*(X_3 - Y_3)$ where $n^*(X_k - Y_k)$ is the number of pairs of type $X_k - Y_k$ remaining at the beginning of Step 3 for each $X_k - Y_k$. Two cases are possible: Case (1) $n^*(X_1 - Y_1) > n^*(X_2 - Y_2) + n^*(X_3 - Y_3)$: Then the algorithm matches all pairs of types $X_2 - Y_2$ and $X_3 - Y_3$ exclusively with those of type $X_1 - Y_1$ according to the priority order such that $n^*(X_1 - Y_1) - \big(n^*(X_2 - Y_2) + n^*(X_3 - Y_3)\big)$ lowest-priority pairs of type $X_1 - Y_1$ remain unmatched. Case (2) $n^*(X_1 - Y_1) \leq n^*(X_2 - Y_2) + n^*(X_3 - Y_3)$: Then all remaining pairs of these three types are matched if their total number is even, and the lowest-priority one among them remains unmatched if their total number is odd. A way to implement the matching in this case is that, after determining which pairs will be matched, (a) match type $X_k - Y_k$ pairs for any k with the pairs of the other two types such that an equal number of the other two type pairs are yet to be matched, and then (b) match the remaining two groups with each other.

no valuable feature is unutilized in a 0-waste exchange. In contrast, one valuable feature is unutilized in a 1-waste exchange and two valuable features are unutilized in a 2-waste exchange.

5.2 Two Sizes & Left-Lobe or Right-Lobe Transplantation

We are ready to proceed with our analysis maintaining the assumption of two sizes, but including both transplantation technologies. This version of the model is no longer a special case of the maximum matching problem, and thus, it is in need of a novel analysis.

5.2.1 Transformation of Willing Pairs

With the introduction of right-lobe transplantation, a small donor can donate his right lobe to a large patient (i.e., $\rho(0) = 1$). A large donor, on the other hand, is in no need to donate his right lobe, whether he donates to a large patient or a small patient. Whenever matched, he donates his left lobe. Hence, without loss of generality, we can assume that all pairs with large donors are unwilling.

Observe that when a type $X_1X_2X_3 - Y_1Y_20w$ pair donates a right lobe, it *mimics* a pair of the type $X_1X_2X_3 - Y_1Y_21$: The only difference between these two types is the size of the donors, with a small donor for the former and a large donor for the latter. But with only two individual sizes, a small donor is able to donate to any size patient (as if he were a large donor) provided that he is willing to donate his right lobe. And once he donates a right lobe, he becomes indistinguishable from a large donor. Thus we treat a type $X_1X_2X_3-Y_1Y_20w$ pair as if it were of type $X_1X_2X_3-Y_1Y_21$ once its left-lobe-donation opportunities are exhausted, and refer to this operation as a **transformation**. A transformed pair of type $X_1X_2X_3 - Y_1Y_20w$ is only considered for right-lobe donation, and, for description purposes, the pair is treated as if it were of type $X_1X_2X_3 - Y_1Y_21$. When such a transformation is carried out, the type $X_1X_2X_3 - Y_1Y_21$ includes both original pairs of type $X_1X_2X_3 - Y_1Y_21$ and transformed pairs of type $X_1X_2X_3 - Y_1Y_20w$, and it is referred to as an **auxiliary type**.

With the addition of right-lobe donation to left-lobe donation as a feasible technology, the set of possible exchanges expands. The expanded set of exchanges are depicted in Figure 4. As in Figure 2, each solid line in Figure 4 represents a left-lobe-only exchange. In addition, any pair of a type $X_1X_2X_3 - Y_1Y_20w$ can mimic a pair of type $X_1X_2X_3 - Y_1Y_21$ by donating a right lobe, and through the resulting transformation, it can participate in any exchange the latter pair qualifies for. Transformations of Category IV and Category V willing pairs, as defined in Lemma 2, can result in new exchanges in this way, and they are represented by a dashed line ending with an arrow in Figure 4. Category IV willing pairs have no left-lobe-exchange option, while Category V willing pairs do. The flexibility to take part in exchange both through a left-lobe donation and a right-lobe donation renders the treatment of Category V willing pairs of particular importance, and it means the timing of their transformations

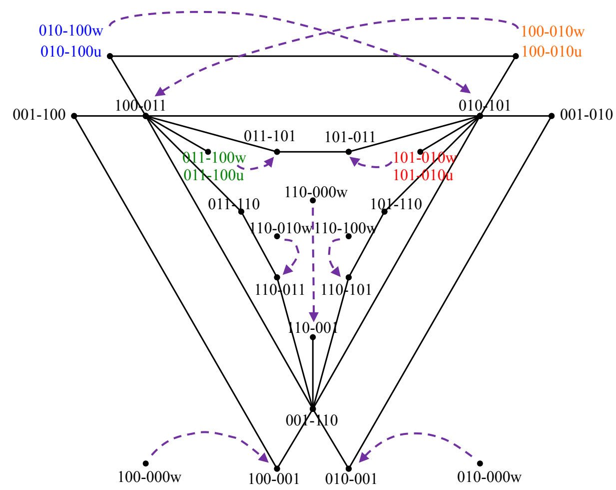


Figure 4: Possible (two-way) exchanges with left- and right-lobe transplants among pair types when $S = 2$

have efficiency and incentive-compatibility implications.

There are additional transformations that do not lead to any new exchange, and hence, they are not depicted on Figure 4. These are transformations of Category II and Category VI willing pairs, and they render a pair to be right-lobe-only compatible. Category VI willing pairs can participate in exchange by donating a left lobe, whereas Category II willing pairs cannot. Hence, for Category II willing pairs, the only transplantation option is a direct right-lobe transplantation (as indicated in Lemma 2).

5.2.2 The Two-Size Left&Right-Lobe Sequential Matching Algorithm

As our main contribution of this section, we introduce an individually rational, Pareto-efficient, and incentive-compatible mechanism. Our starting point is the baseline left-lobe-only algorithm presented in Section 5.1.

In the absence of right-lobe donation, the intuitive idea of clearing 0-waste exchanges, 1-waste exchanges, and 2-waste exchanges in sequence results in a Pareto-efficient outcome. While the inclusion of the right-lobe-donation technology considerably complicates the analysis, a number of key observations are instrumental to formulate a natural modification of the baseline algorithm presented in Section 5.1. Rather than presenting the entire algorithm in an uninterrupted way, we introduce its steps one step at a time, emphasizing how each step relates and differs from the corresponding step of the baseline algorithm. As we introduce each step, we make a number of observations, leading our way to a natural next step.

While pairs of four categories (Categories II, IV, V, and VI) are potentially affected by the adoption of the right-lobe-transplantation technology, the key incremental innovation in our modified algorithm pertains to the timing of the transformations of Category V willing pairs. There are four Category V types, each denoted with a different color in this section to visually emphasize their role in the modified algorithm: $010 - 100$, $100 - 010$, $011 - 100$, $101 - 010$.

We are ready to proceed with the introduction of the modified algorithm:

Step 0a. Fix a priority order over all pairs, and a preference (or equivalently willingness) profile R.

In addition to each compatible pair (i.e., pairs of Category I) that immediately receives a direct left-lobe transplantation under the modified algorithm, each willing pair of Category II immediately receives a direct right-lobe transplantation. Pairs of Category II lack any exchange possibility, and a direct right-lobe transplantation is their only transplantation option. As such, willing pairs of Category II are to be directly matched right away.

Step 0b. Match each Category I pair by a direct left-lobe transplant.

Step 0c. Match each Category II willing pair by a direct right-lobe transplant.

Another category that completely relies on right-lobe donation for a transplant is Category IV. And just as for the members of Category II, its members have a single means of transplantation. For Category II, that single means was a direct right-lobe transplantation. For Category IV, it is an exchange through a right-lobe donation. Since willing pairs of Category IV have no means of transplantation through left-lobe donation, they are to be transformed right away.

Step 1a. Transform each Category IV willing pair, constructing a modified pool.

Since some pairs are already transformed, a type in the rest of the algorithm refers to the associated auxiliary type that includes both the original pairs of the type as well as the transformed pairs. We are ready to clear 0-waste exchanges.

Step 1b. Clear 0-waste exchanges in the modified pool: For each $X - Y \in \mathcal{T} \times \mathcal{T}$, match the highest-priority pair of (auxiliary) type $X - Y$ with the highest-priority pair of its reciprocal (auxiliary) type $Y - X$. Proceed in a similar way, matching the next highest-priority pairs of types $X - Y$ and $Y - X$ with each other, and so on, until one of the two types is exhausted.

So far the modifications to the algorithm have been straightforward. The next few modifications, however, rely on a number of critical observations on Category V types.

First observe that, any remaining pair of Category V may have left-lobe-donation possibilities, although not through exchanges of 0-waste. Therefore, by incentive compatibility,

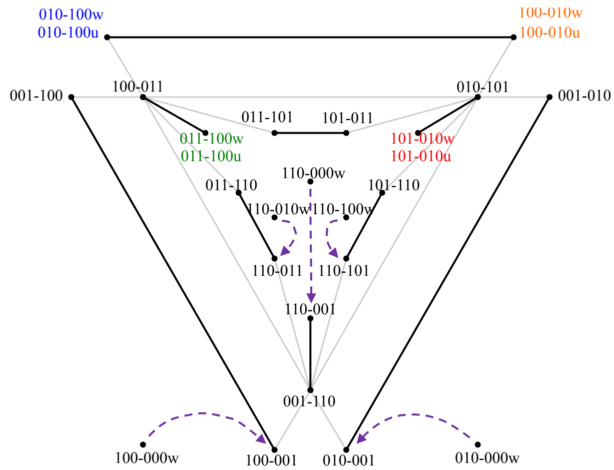


Figure 5: Transformations in Step 1(a), followed by cleared exchanges in Step 1(b)

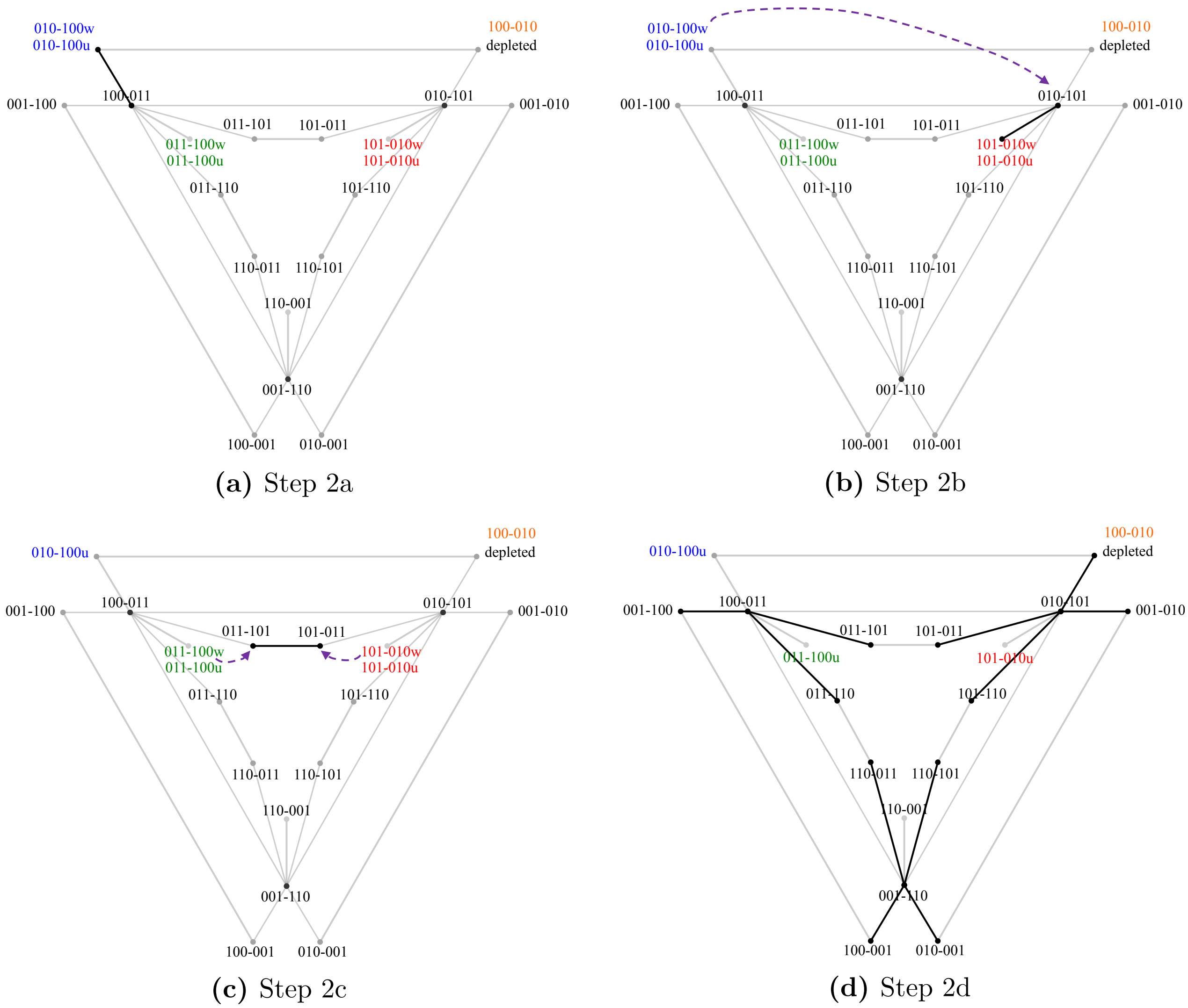
no pair of Category V can be transformed just yet. But since all available 0-waste exchanges are already cleared in Step 1b, at least some of the 1-waste exchanges are to be cleared.

Next observe that, once some Category V pairs are transformed, new potential exchanges may form, including some 0-waste exchanges. Since 0-waste exchanges are associated with higher welfare, not all 1-waste exchanges shall be cleared right away. A given 1-waste exchange is a candidate at this stage only if it is blocking a transformation that could generate new 0-waste exchanges.

Third, while transformation of willing pairs of Category V type $010-100$ may potentially lead to new left-lobe-donation possibilities for pairs of Category V type $101-010$, left-lobe-donation possibilities for pairs of type $010-100$ are completely exhausted, even if all other Category V pairs are transformed. Similarly, while transformation of willing pairs of Category V type $100-010$ may potentially lead to new left-lobe-donation possibilities for pairs of Category V type $011-100$, left-lobe-donation possibilities for pairs of type $100-010$ are exhausted regardless of which pairs of Category V are transformed. Therefore, it is sensible to transform willing pairs of Category V types $010-100$ and $100-010$ before the willing pairs of Category V types $101-010$ and $011-100$.

Finally, observe that, since the two Category V types $010-100$ and $100-010$ are reciprocal, at least one of these two types is exhausted at the end of Step 1b. Together with our third observation, this observation suggests that the willing pairs of the surviving type between $010-100$ and $100-010$ will be the first Category V type to be transformed.

Without loss of generality, let us assume that pairs of type $100-010$ are exhausted. In that case, the target type to be transformed is $010-100$. Since pairs of this type still have remaining left-lobe-donation possibilities, they cannot be transformed just yet. But their only remaining left-lobe donations involve 1-waste exchanges with the type $100-011$. So these exchanges are to be cleared next. Moreover, once willing pairs of type $010-100$ are



(a) Step 2a **(b)** Step 2b **(c)** Step 2c **(d)** Step 2d

Figure 6: Transformations followed by cleared exchanges in Step 2

transformed to type $010 - 101$, a new set of 0-waste exchanges becomes available between pairs of auxiliary type $010 - 101$ and type $101 - 010$. Hence, next two steps have emerged:

Step 2a. Assume without loss of generality, pairs of type $100 - 010$ are exhausted.[15] Following the given priority order, clear 1-waste exchanges between the remaining pairs of type $010 - 100$ and pairs of type $100 - 011$.

Step 2b. Transform any remaining pairs of type $010 - 100w$ to auxiliary type $010 - 101$, obtaining an exchange pool modified for a second time.

Clear 0-waste exchanges in the new pool: Match the highest-priority pair of (auxiliary) type $010 - 101$ with the highest-priority pair of its reciprocal type

[15]Otherwise, all remaining steps are symmetrically defined below, swapping 1st and 2nd binary digits for both patient and donor in all relevant auxiliary types.

$101-010$. Proceed in a similar way, matching the next highest-priority pairs of types $010-101$ and $101-010$ with each other, and so on, until one of the two types is exhausted.

At this point, left-lobe-donation possibilities are exhausted for pairs of type $101-010$. In addition, since new left-lobe-donation possibilities for pairs of type $011-100$ depend on transformations of pairs of type $100-010w$, and because the latter type are exhausted by assumption at the end of Step 1b, left-lobe-donation possibilities are exhausted for pairs of type $011-100$ as well. Thus, willing pairs of types $101-010$ and $011-100$ are to be transformed to auxiliary types $101-011$ and $011-101$, respectively. These transformations, in turn, result in a new set of 0-waste exchanges between auxiliary types $101-011$ and $011-101$. Hence our next step has also emerged:

Step 2c. Transform any remaining pairs of types $011-100w$ and $101-010w$, to auxiliary types $011-101$ and $101-011$, respectively, obtaining a new exchange pool modified for a third (and the last) time.

Clear 0-waste exchanges in the new pool: Match the highest-priority pair of (auxiliary) type $101-011$ with the highest-priority pair of its reciprocal (auxiliary) type $011-101$. Proceed in a similar way, matching the next highest-priority pairs of types $101-011$ and $011-101$ with each other, and so on, until one of the two types is exhausted.

Having transformed all willing pairs of Category V, all potential 0-waste exchanges are exhausted. Thus, we are ready to clear 1-waste exchanges, followed by 2-waste exchanges.

Step 2d. Clear the remaining 1-waste exchanges: For each of the three value 1 types $X-Y \in \{100-011, 010-101, 001-110\}$, match the highest-priority pair of type $X-Y$ with the highest-priority pair of its value 0 neighbor. Proceed in a similar way, matching the next highest-priority pair of type $X-Y$ with the next highest-priority pair of its value 0 neighbor, and so on, until either type $X-Y$ or its 0 value neighbors are exhausted.

Step 3. Clear 2-waste exchanges in the new pool: Match the maximum number of $100-011$, $010-101$, and $001-110$ auxiliary types with each other, following the given priority order.

As in our baseline algorithm, all exchanges are exhausted by the end of Step 3. However, different than the baseline algorithm, one last set of possible transplants remain. Recall that pairs of Category VI can not only take part in exchange through left-lobe donation but also

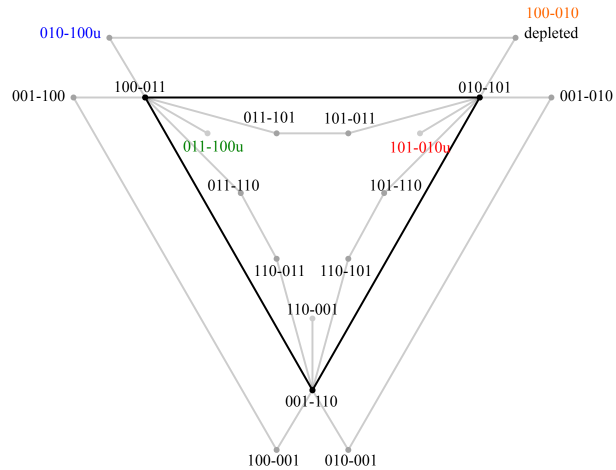


Figure 7: Cleared exchanges in Step 3

receive direct right-lobe transplantation. Since they all prefer left-lobe donation to right-lobe donation, willing pairs of Category VI have not been transformed until all exchanges are exhausted. Since their left-lobe-donation possibilities are exhausted, willing pairs of Category VI are to be matched by direct right-lobe transplants at the termination of the modified algorithm.

> **Step 4.** Match each remaining Category VI willing pair by a direct right-lobe transplant.

Steps 1a-b, 2a-d, 3 of the modified algorithm are depicted in Figures 5, 6, and 7, respectively, for the case when pairs of Category V type $100-010$ are exhausted in Step 1.

We are ready to present the main result of this section:

Theorem 2 *The two-size left&right-lobe sequential matching mechanism is individually rational, Pareto efficient, and incentive compatible.*

5.2.3 An Impossibility for Two Sizes & Left-Lobe or Right-Lobe Transplantation

Although the left&right-lobe sequential matching mechanism is Pareto efficient, it may not maximize the number of transplants. The example below shows that one has to sacrifice incentive compatibility in order to maximize the number of transplants or even the number of left-lobe transplants.

Example 1 *Consider an exchange pool with* $\mathcal{I} = \{i_1, i_2, i_3, i_4\}$ *and*

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

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

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

Observe that $t(i_3, i_1) = t(i_4, i_1) = r$ while $t(i_3, i_2) = t(i_4, i_2) = l$. Any (probabilistic) mechanism that chooses a matching with the maximum number of transplants or the maximum number of left-lobe transplants chooses at least one of these two matchings in its support. Without loss of generality, suppose M is that matching. Then i_3 has an incentive to announce its type as unwilling by revealing $R'_{i_3} = R^u_{i_3}$, as the mechanism will choose M', which is the unique left-lobe-donation- and total-transplant-maximizing matching in this case, with probability 1. Hence, there is no incentive-compatible mechanism that maximizes the total number of transplants or left-lobe transplants.

Example 1 also serves as a proof for the following impossibility result:

Proposition 1 *There is no incentive-compatible mechanism that maximizes the number of transplants or the number of left-lobe transplants.*

6 An Efficient & IC Mechanism for the General Model

We are ready to analyze the general model, dropping the restriction of two individual sizes. Thus, $\mathcal{S} = \{0, 1, ..., S\}$ with $S \geq 2$ for the rest of the paper. The symmetry of the simplified model on the binary representation of each characteristic, the concepts of value and waste, and the interpretation of exchange as a trade of the three characteristics paved our way to a very natural algorithm when there are only two individual sizes. While each exchange still involves a trade of the three characteristics under the general model, the full symmetry of the simplified model is lost, and the concepts of value or waste are less natural in this more general context. With only two sizes, each pair can have a "deficit" or "excess" of each characteristic without any cardinality consideration. And pairs that have two characteristics with an excess are more valuable for trade, as reflected in their roles under our algorithms presented in Section 5. If excess or deficit instead had a cardinal measure in our simplified model, the roles of these pairs would have been less clear, because they could have a "large" deficit in the third characteristic. The challenge of the general model with multiple sizes is that pairs can have different levels of excess or deficit in their size characteristics. This complication makes comparisons of pair values or exchange wastes less natural in this context. As such, we abandon these concepts for the general model. Certain features of our simplified model, however, such as the critical role of Category V pairs and their hierarchy for the timing of their transformations, still persist under the general model. We build on these features to design a mechanism with the desired properties.

There are two main technical challenges for this more ambitious design. First, whether an efficient matching can be obtained through a straightforward sequential algorithm is an open problem for the general model, even if transplantation is restricted to left-lobe only. Thus, we introduce a non-sequential (i.e., non-greedy) algorithm. Second, unlike the simplified model where a natural transformation sequence for Category V pairs emerges, it is not immediately clear how these pairs are to be transformed to maintain efficiency and incentive compatibility. Fortunately, it is possible to overcome both challenges by extending the tools we introduced for the simplified model and supplementing them with standard techniques from combinatorial optimization.

The first challenge is well analyzed in the combinatorial optimization literature. In a general graph (not necessarily the compatibility graph of a liver-exchange pool), one has to recursively expand the set of simultaneously *matchable* pairs to find an efficient matching (e.g., as in the cardinality matching algorithm of Edmonds, 1965). In the absence of the right-lobe transplantation, incentive compatibility would be redundant and efficient liver exchange would be a special case of the maximal matching problem. As a starting point of our general algorithm, we follow this approach. We next define matchability formally.

Suppose pairs in $\mathcal{J}_w \subseteq \mathcal{I}$ are willing to participate in both left- and right-lobe donation. We say that a subset of pairs $\mathcal{J} \subseteq \mathcal{I}$ is **matchable** in compatibility graph $G_c[\mathcal{J}_w]$ if there exists a matching $M \in \mathcal{M}_c[\mathcal{J}_w]$ such that $M(j) \neq \emptyset$ for all $j \in \mathcal{J}$.[16] Whether a subset of pairs is matchable or not can be checked in polynomial time.[17]

6.1 The Precedence Digraph

We next develop the tools that will help us to overcome the second challenge and pave our way to design a Pareto-efficient and incentive-compatible mechanism for the general model.

The timing and the sequence of Category V transformations play key roles in assuring Pareto efficiency and incentive compatibility of the mechanism we introduced for the simplified model. Incentive compatibility is assured simply by transforming each Category V pair only after their left-lobe-donation possibilities are exhausted. We adopt a similar strategy for the general model. Assuring Pareto efficiency for the simplified model, in contrast, relies on a deeper observation on a hierarchy between Category V types, in terms of the timing of their transformations. The key to our design for the general model is the observation that similar hierarchies also exist for the general model.

Consider two pairs, one each from Category V types $X - Y$ and $U - V$. Suppose that while the two pairs cannot form a left-lobe-only exchange, they can form an exchange where the donor of type Y donates his right lobe to the patient of type U, and the donor of type V

[16] Our definition of matchability differs from the standard definition in graph theory, which requires the subgraph induced by $\mathcal{J}$ to have a perfect matching. See, for instance, Schrijver (2003, Vol I, p59).

[17] We explain a polynomial-time method for checking matchability in Appendix D.2.3.

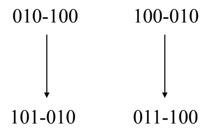


Figure 8: Precedence digraph on Category V types with two sizes ($S = 2$)

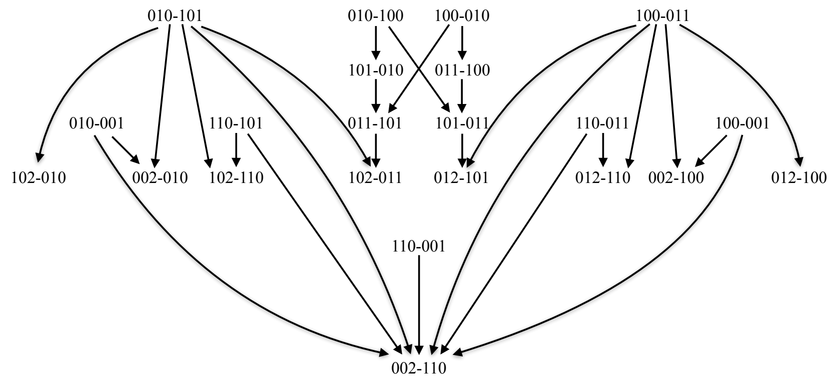


Figure 9: Precedence digraph on Category V types with three sizes ($S = 3$)

donates his left lobe to the patient of type X. In such a case, it is plausible to transform the pair of type $X - Y$ before the pair of type $U - V$, because the left-lobe-exchange possibilities of the pair of type $U - V$ may expand after transforming the type $X - Y$ pair. This motivates the following definition.

Define a directed graph on the set of Category V types, that we call the **precedence digraph**, where for any two Category V types $X - Y$ and $U - V$:

$$X - Y \longrightarrow U - V \quad \Longleftrightarrow \quad X \leq V,\ U \not\leq Y\ \&\ U \leq \rho(Y).$$

If $X - Y \longrightarrow U - V$, we will also say that $X - Y$ **precedes** $U - V$. For any two Category V pairs i and j, we will also write $i \longrightarrow j$ whenever $\tau(i) \longrightarrow \tau(j)$, and say that i **precedes** j.

In the above definition, type $X - Y$ precedes type $U - V$ if a donor of type V can donate his left lobe to a patient of type X, whereas a donor of type Y can donate his right lobe but not his left lobe to a patient of type U. In this case, two pairs fail to form a left-lobe-only exchange, but they can form an exchange following a transformation of the type $X - Y$ pair. Figure 8 depicts the precedence digraph over Category V types in the case of two sizes ($\mathcal{S} = \{0, 1\}$, $\rho(0) = \rho(1) = 1$). Figure 9 depicts the precedence digraph over Category V types in the case of three sizes ($\mathcal{S} = \{0, 1, 2\}$, $\rho(0) = 1$, $\rho(1) = \rho(2) = 2$).

In defining our mechanism more generally, we would like to transform pairs of Category V in a sequence such that for any two Category V types $X - Y$ and $U - V$, where $X - Y$ precedes $U - V$, each pair of type $X - Y$ is transformed before each pair of type $U - V$. It is indeed possible to find such a transformation order. Note that the precedence digraphs in

Figure 8 and Figure 9 are acyclic. This observation is not specific to these two examples. As stated by the next lemma, the precedence digraph for any liver-exchange pool is acyclic.

Lemma 4 *The precedence digraph on Category V types is acyclic.*

Together with Lemma 8 in Appendix D.1, Lemma 4 implies that the precedence digraph is consistent with a linear order, which can then be used to construct a linear order, called *topological order*, over all pairs of Category V. Our general mechanism uses the topological order to determine the transformation sequence of Category V pairs.

6.2 The Left&Right-Lobe Priority Algorithm for the General Model

We are ready to present an iterative algorithm, which can be used to find the outcome of our mechanism. An expanded equivalent definition, used in the proof of our main result in this section, is given in Appendix D.2.

> **Step 0:** Fix a priority order over all pairs, a topological order of Category V pairs induced by the precedence digraph, and a preference profile R.
>
> Match each Category I pair by a direct left-lobe transplant.
>
> Match each willing Category II pair by a direct right-lobe transplant.
>
> Leave unwilling Category II and IV pairs and all Category 0 pairs unmatched.
>
> **Step 1:** Let $\mathcal{I}_0$ be the set of pairs not handled in Step 0. Transform willing Category IV pairs, the set of which we denote as $\mathcal{I}_{IVw}$, and obtain a new auxiliary pool. Thus, $G_0 := (\mathcal{I}_0, E_c[\mathcal{I}_{IVw}])$ is our initial compatibility graph.[18] Let $\mathcal{J}_0 := \emptyset$ and $\widetilde{\mathcal{J}}_0 := \emptyset$. Step 1's substeps proceed inductively.
>
>> *Step* 1.(k)*:* Let i be the k'th highest-priority Category V pair under the topological order. If $\mathcal{J}_{k-1} \cup \{i\}$ is matchable in G_{k-1}, then let $\mathcal{J}_k := \mathcal{J}_{k-1} \cup \{i\}$, $\widetilde{\mathcal{J}}_k := \widetilde{\mathcal{J}}_{k-1}$. Otherwise let $\mathcal{J}_k := \mathcal{J}_{k-1}$, and
>>
>> - if i is not willing: let $\widetilde{\mathcal{J}}_k := \widetilde{\mathcal{J}}_{k-1}$.
>> - if i is willing: let $\widetilde{\mathcal{J}}_k := \widetilde{\mathcal{J}}_{k-1} \cup \{i\}$.
>>
>> Define $G_k := (\mathcal{I}_0, E_c[\mathcal{I}_{IVw} \cup \widetilde{\mathcal{J}}_k])$, i.e., the graph obtained by transforming all willing Category IV pairs along with willing Category V pairs in $\widetilde{\mathcal{J}}_k$. Proceed with Step 1.(k+1).
>
> Step 1 ends at substep K, where K is the number of Category V pairs. Set $\mathcal{J}_K$ of Category V pairs is matchable in G_K, where each pair in $\mathcal{J}_K$ donates a left lobe, while $\widetilde{\mathcal{J}}_K$ is the set of all willing Category V pairs that are transformed.

[18] This means, at this point only pairs in the set $\mathcal{I}_{IVw}$ are available for a right-lobe transplant.

Step 2: By this point, we are committed to match pairs in $\mathcal{J}_K$, but not the others. Inductively, we continue with the remaining pairs $\mathcal{I}_0 \setminus \mathcal{J}_K$. Let $\mathcal{J}_0^* := \emptyset$.

> *Step* 2.(n)*:* Let i be the n'th highest-priority pair in $\mathcal{I}_0 \setminus \mathcal{J}_K$ according to the priority order over $\mathcal{I}$. If $\mathcal{J}_K \cup \mathcal{J}_{n-1}^* \cup \{i\}$ is matchable in G_K, then let $\mathcal{J}_n^* := \mathcal{J}_{n-1}^* \cup \{i\}$. Otherwise, $\mathcal{J}_n^* := \mathcal{J}_{n-1}^*$. Proceed with Step 2.(n+1).

When Step 2 ends at substep $N = |\mathcal{I}_0 \setminus \ \mathcal{J}_K|$, the set of pairs $\mathcal{J}_K \cup \mathcal{J}_N^*$ is a matchable set in G_K.

Step 3: Match each willing Category VI pair in $\mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$ by a direct right-lobe transplant.

Together with the direct transplants determined in Steps 0 and 3, the algorithm determines $\mathcal{J}_K \cup \mathcal{J}_N^*$ as the set of pairs to participate in exchange, where pairs in $\mathcal{I}_{IVw} \cup \widetilde{\mathcal{J}}_K$ participate via right-lobe donation and the rest via left-lobe donation. Each pair is indifferent between any such matching, and the mechanism picks any one of them as its outcome.[19]

The starting point of the above-described algorithm is the priority matching mechanism analyzed in Roth, Sönmez, and Ünver (2005). In the absence of right-lobe donation, the two mechanisms are identical. And in the absence of Category V pairs, modifying this mechanism to allow for right-lobe donation is straightforward: In addition to pairs of Category V, pairs of Categories II, IV, and VI potentially benefit from right-lobe donation. For pairs of Category II, there are no left-lobe-donation possibilities, and hence willing pairs of this category are to directly receive a right-lobe transplant. For pairs of Category IV, an exchange through right-lobe donation is the only possibility for a transplant, and therefore, willing pairs of this category are to be transformed right away for a possible exchange. For each willing pair of Category VI, on the other hand, a direct right-lobe transplant is a last course of action in case it fails to be part of an exchange donating a left lobe. Dealing with Category V pairs, in contrast, is more delicate. That is because these pairs can be part of an exchange donating either a left lobe or a right lobe, and while each Category V pair prefers the former, the exchange options for other pairs might expand if they do the latter. Thanks to the topological order among Category V pairs induced by the precedence digraph, we are able to integrate these pairs into the general mechanics of the priority mechanism without compromising its Pareto efficiency and incentive compatibility. Since these pairs influence the underlying compatibility graph, they are to be handled before the other pairs (in Step 1). Following the topological order for the transformation of these pairs assures that no new left-lobe-donation possibilities emerge for any Category V pair at a later step of the algorithm once they are exhausted; this in turn assures that these pairs do not engage in

[19] We explain a polynomial-time method for how to find such a matching in Appendix D.2.3.

less-preferred exchanges due to a premature transformation. Once which Category V pairs are to be matched and which are to be transformed is determined at the end of Step 1, our mechanism adopts at Step 2 the standard priority mechanism dynamics given for a fixed compatibility graph. There is one additional trivial step at the conclusion of the algorithm, to directly right-lobe transform the remaining willing Category VI pairs.

We are ready to present our last result.

Theorem 3 *The general left&right-lobe priority mechanism is individually rational, Pareto efficient, and incentive compatible.*

7 Simulations

In this section, we report the results of computer simulations using South Korean aggregate statistics to determine the potential gains of liver exchange over direct transplantation alone.

Calibration Statistics for Simulations from South Korean Population				
	Live-Donation Recipients	**Live Donors**	**Height (cm)**	
Female	1492 (34.55%)	1149 (26.61%)	*Mean:* 157.40	*Std Dev:* 5.99
Male	2826 (64.45%)	3169 (73.39%)	*Mean:* 170.70	*Std Dev:* 6.40
Total	4318 (100.0%)	4318 (100.0%)		
Blood-Type Distribution				
O	**A**	**B**	**AB**	**Total**
37%	33%	21%	9%	100%

Table 1: Calibration statistics from South Korea for liver-exchange simulations. Blood-type distribution is obtained from `http://bloodtypes.jigsy.com/East_Asia-bloodtypes` on 04/10/2016. Mean and standard deviation for South Korean adult height distribution are obtained from the Korean Agency for Technology and Standards (KATS) website `http://sizekorea.kats.go.kr` on 04/10/2016. The transplant data is obtained from the Korean Network for Organ Sharing (KONOS) 2014 Annual Report, retrieved from `http://www.konos.go.kr/konosis/index.jsp` on 04/10/2016 and contains the years 2010–2014.

Table 1 summarizes the calibration parameters used in our simulations. Each patient is assumed to be paired with a donor. Blood type, gender, and height characteristics for patients and their donors are determined independently and randomly.[20]

[20] We use the following weight determination formula as a function of height (also see Ergin, Sönmez, and Ünver, 2017): $w = a\ h^b$, where w is weight in kilograms, h is height in meters, and constants a and b are set as $a = 26.58, b = 1.92$ for males and $a = 32.79, b = 1.45$ for females (Diverse Populations Collaborative Group, 2005). The body surface area (BSA in m^2) of an individual is determined through the Mostellar formula given in Um et al. (2015) as $BSA = \frac{\sqrt{h\ w}}{6}$, and the liver volume ($l_v$ in ml) of Korean adults is determined through the estimated formula in Um et al. (2015) as $l_v = 893.485\ BSA - 439.169$. Each patient and donor have a height drawn independently from the truncated normal distribution using the mean and std. dev. reported in this table with the support [mean - 3 std. dev., mean + 3 std. dev.]. We assume that the left lobe of each donor is 35% of all his liver, as this is reported as the mean of the left-lobe volume in Korea (Um et al., 2015).

A donor and patient are deemed left-lobe compatible if they are blood-type compatible and the donor's left lobe volume is at least 40% of the total liver volume of the patient. A donor and patient are deemed right-lobe-only compatible if they are not left-lobe compatible, they are blood-type compatible, and the donor's right-lobe volume is at least 40% of the total liver volume of the patient.

We generate $I = 50$, 100, and 250 patient-donor pairs in three sets of simulations. Since we do not have empirical statistics on the willingness of donors for right-lobe donation, we consider 6 scenarios for each population size in which on average $0, 20, 40, 60, 80$, and 100% of all pairs are willing. For each given willingness rate, we randomly determine each pair's willingness. We consider two treatments:

- *No exchange*: Only left-lobe-compatible and right-lobe-only-compatible, willing pairs participate in direct transplants.
- *Proposed Mechanism*: An outcome of our general left&right-lobe priority mechanism is determined for arbitrary topological and priority orders.

The results of the simulations are given in Table 2. About 12.5% of all pairs are left-lobe compatible and their patients receive a direct left-lobe transplant. Up to 45.5% of all pairs become right-lobe-only compatible as a linear, increasing function of the willingness rate (see the no exchange treatment in the table). Therefore, in the absence of liver exchange, 12.5% to 58.0% of patients with living donors receive a direct transplant as a linear, increasing function of the willingness rate. Our mechanism, on the other hand, matches from 18% to 77.75% of all pairs, in a seemingly concave, increasing function of the willingness rate for $I = 100$ (see exchange treatment in the table). Thus, the increase in the number of transplants due to exchange is in the range of 44% to 34%, higher for the lower values of the willingness rate. Our proposed mechanism not only increases the number of living-donor liver transplants, but also increases the reliance on the lower-risk left-lobe liver transplantation in the spirit of the central tenet of the hippocratic oath "first do no harm." For example, when all pairs are willing, the share of left-lobe transplants increases from 21.5% to 31.1%. In general for any willingness rate, the rate of increase in left-lobe transplants is higher than the rate of increase in right-lobe transplants.

8 Conclusion

We introduced a liver-exchange model where the donor of each pair can donate either the smaller and safer-to-donate left liver lobe or the larger and riskier-to-donate right liver lobe. While liver exchange is inspired by the increasingly widespread kidney exchange, analytically it is a more challenging problem due to its dual-donation possibility. On the one hand, right-lobe donation expands the set of feasible exchanges, increasing the number of patients who can receive a transplant. On the other hand, it is a considerably higher-risk procedure for the donor, thereby possibly discouraging some of the donors from this option. And since

Population & Cat. V Pairs	w Rate	No exchange: Left Lobe Direct	No exchange: Right Lobe Direct	No exchange: *Total*	Proposed Mechanism: Left Lobe Direct	Proposed Mechanism: Right Lobe Direct	Proposed Mechanism: Left Lobe Exchange	Proposed Mechanism: Right Lobe Exchange	Proposed Mechanism: *Total*
50 & 10.044 (2.824)	0	6.204	0	6.204	6.204	0	1.918	0	8.122
				(2.280)					(2.831)
	0.2	6.204	4.529	10.733	6.204	4.218	2.670	1.822	14.914
				(2.912)					(3.598)
	0.4	6.204	9.081	15.285	6.204	8.201	3.286	3.532	21.223
				(3.330)					(3.961)
	0.6	6.204	13.665	19.869	6.204	11.948	3.816	5.028	26.996
				(3.491)					(3.930)
	0.8	6.204	18.301	24.505	6.204	15.636	4.336	6.404	32.580
				(3.581)					(3.840)
	1	6.204	22.870	29.074	6.204	19.126	4.767	7.521	37.618
				(3.507)					(3.581)
100 & 20.003 (4.026)	0	12.497	0	12.497	12.497	0	5.498	0	17.995
				(3.367)					(4.526)
	0.2	12.497	9.097	21.594	12.497	8.372	7.031	4.425	32.325
				(4.255)					(5.540)
	0.4	12.497	18.196	30.693	12.497	16.126	8.303	8.303	45.229
				(4.605)					(5.681)
	0.6	12.497	27.277	39.774	12.497	23.267	9.543	11.611	56.918
				(4.814)					(5.570)
	0.8	12.497	36.402	48.899	12.497	30.127	10.693	14.447	67.764
				(5.032)					(5.350)
	1	12.497	45.561	58.058	12.497	36.703	11.691	16.869	77.760
				(5.062)					(5.226)
250 & 49.628 (6.314)	0	31.031	0	31.031	31.031	0	19.652	0	50.683
				(5.236)					(7.681)
	0.2	31.031	22.895	53.926	31.031	20.523	23.315	13.141	88.010
				(6.572)					(9.000)
	0.4	31.031	45.500	76.531	31.031	39.045	26.512	23.792	120.38
				(7.263)					(8.967)
	0.6	31.031	68.387	99.418	31.031	56.391	29.478	32.736	149.636
				(7.639)					(8.543)
	0.8	31.031	91.294	122.325	31.031	72.679	32.249	40.423	176.382
				(7.777)					(8.322)
	1	31.031	114.084	145.115	31.031	88.376	34.656	46.546	200.609
				(7.744)					(7.805)

Table 2: Simulation results for population sizes $I = 50$, 100, 250 and willingness (w) rates 0, 0.2, 0.4, 0.6, 0.8, 1. Standard deviations of the populations for the total number of transplants are reported below the averages in parentheses for 1000 simulations. Numbers of Category V pairs are given below I in the first column.

some donors will be willing to donate their left lobes but not their right lobes, the liver-exchange problem harbors a novel incentive compatibility consideration that is not present in kidney exchange. Exploiting the acyclicity of a certain directed graph among pairs which can participate in exchange both through left-lobe donation and right-lobe donation, we introduced a novel exchange mechanism that is Pareto efficient and incentive compatible. The welfare gains from adopting our mechanism are considerable, and depending on the ratio of donors who are willing to donate a right lobe, it increases the number of living-donor liver transplants by 34–44%.

Recently Mishra et al. (2018) advocated for organized liver exchange in the US, emphasizing the choice of a matching algorithm as one of the most difficult issues to be resolved. We believe our proposed mechanism is a viable solution for this important problem.

Appendix A Proofs of Lemmas in Section 3.4

Proof of Lemma 1: Suppose a pair $i \in \mathcal{I}$ of type $X - Y$ participates in an exchange with a pair $j \in \mathcal{I} \setminus \{i\}$ of type $U - V$ in an individually rational matching $M \in \mathcal{M}_c[\mathcal{I}]$. Then $X \not\leq Y$, as otherwise, i would have received a direct left-lobe transplant. For the proof of the first claim, suppose $t(i,j) = l$. Suppose, to the contrary of the claim, $X > Y$. Two cases are possible, $t(j,i) = l$ or $t(j,i) = r$.

- Case 1. $t(j,i) = l$: Then $U - V$ has to satisfy $V \geq X > Y \geq U$. But then, j has to participate in a direct transplant in M since $t(j) = l$. A contradiction.
- Case 2. $t(j,i) = r$: Then $U - V$ has to satisfy $\rho(V) \geq X > Y \geq U$. But then j has to participate in a direct transplant in M as $t(j) = r$ or $t(j) = l$, a contradiction again.

Thus, $X \not> Y$ should also hold. The second claim, i.e., the case when $t(i,j) = r$, is proven using $\rho(Y)$ instead of Y in the above proof for the first claim. ■

Proof of Lemma 2: Let $R \in \mathcal{R}$ be a willingness profile and $M \in \mathcal{M}_c[\mathcal{I}]$ be an individually rational matching under R. All types $X - Y \in \mathcal{T} \times \mathcal{T}$ fall into one of the following seven mutually exclusive categories:

0. $X > \rho(Y)$: This also implies $X > Y$. Thus, a type $X - Y$ pair cannot participate in exchange in M by a donating a left or right lobe by Lemma 1.

I. $X \leq Y$: A type $X - Y$ pair is left-lobe compatible, and in M, it receives a direct left-lobe transplant. On the other hand, in none of the remaining cases is a direct left-lobe transplant possible for an $X - Y$ pair.

II. $Y < X \leq \rho(Y)$: A type $X - Y$ pair can receive a direct right-lobe transplant in M. Moreover, as $X \geq Y$, it cannot participate in an exchange in M by donating a left lobe by Lemma 1. Thus, it can only remain unmatched in M if it is unwilling.

III. $X \not\leq \rho(Y)$, $X \not\geq Y$, & $Y = \rho(Y)$: Since also $X \not\leq \rho(Y)$, the pair cannot participate in a direct left-lobe or right-lobe transplant. Since $Y = \rho(Y)$, in all exchanges it can participate by donating a right lobe, it can also participate by donating a left lobe. By Lemma 1, it can participate in an exchange in M by donating a left lobe, or it remains unmatched.

IV. $X > Y$, $X \not\geq \rho(Y)$, & $X \not\leq \rho(Y)$: By Lemma 1, a type $X - Y$ pair cannot participate in an exchange in M by donating a left lobe. On the other hand, again by Lemma 1, it can participate in an exchange in M to donate a right lobe. So this is its only exchange option in M if it is willing, or it remains unmatched in M.

V. $X \not\leq \rho(Y)$, $X \not\geq Y$, & $Y < \rho(Y)$: By Lemma 1, a type $X - Y$ pair can participate in exchange by donating a left lobe. On the other hand, by Lemma 1, it can participate in an exchange in M by donating a right lobe as well, if it is willing. Thus, it has two exchange options in M. It can also remain unmatched in M.

VI. $X < \rho(Y)$, $X \not\gtrsim Y$, & $X \not\lesssim Y$: In M, it can participate in a direct right-lobe transplant if it is willing. As an exchange needs to be individually rational, it will never participate in an exchange to donate a right lobe. We also have $X \not\gtrsim Y$. Therefore, by Lemma 1, it can participate in an exchange in M by donating a left lobe. In M, it will only be unmatched if it is unwilling.

Conversely, suppose M is a matching consisting entirely of the types of transplants depicted in I–VI above. Then, by definition, M is individually rational, concluding the proof. ■

Proof of Lemma 3: Throughout the proof, let f be an individually rational mechanism.

We start by proving that if f is incentive compatible, then the equivalence holds. Let $i \in \mathcal{I}$ and $R_{-i} \in \prod_{j \neq i} \mathcal{R}_j$. To see the "$\Leftarrow$" direction, suppose for a contradiction that i participates in an exchange by donating a left lobe at $f(R_i^u, R_{-i})$, but not at $f(R_i^w, R_{-i})$. By individual rationality, the pair i is not left-lobe compatible (i.e., not Category I); otherwise, it would directly donate a left lobe at $f(R_i^u, R_{-i})$. Therefore, i cannot directly donate a left lobe also at $f(R_i^w, R_{-i})$. The only feasible match possibilities of i at $f(R_i^w, R_{-i})$ are: i directly donates its donor's right lobe; i takes part in an exchange by donating a right lobe; or i is unmatched. By the definition of preferences, at R_i^w, i strictly prefers being part of an exchange by donating a left lobe to all the three match possibilities at $f(R_i^w, R_{-i})$, i.e.: $f(R_i^u, R_{-i}) P_i^w f(R_i^w, R_{-i})$, contradicting the incentive compatibility of f. The proof of the "$\Rightarrow$" direction is symmetric by switching the roles of R_i^u and R_i^w.

Now assume that the equivalence holds. We will show that this implies incentive compatibility of f. Take any $i \in \mathcal{I}$ and $R_{-i} \in \prod_{j \neq i} \mathcal{R}_j$. If i is left-lobe compatible (i.e., Category I), then individual rationality implies that i directly donates a left lobe at $f(R_i^u, R_{-i})$ and $f(R_i^w, R_{-i})$. Therefore, whether i is willing or unwilling, i is indifferent between $f(R_i^u, R_{-i})$ and $f(R_i^w, R_{-i})$, implying the incentive compatibility condition for i. Suppose next that i is not left-lobe compatible. By the equivalence, there are two cases to consider:

Case 1: In both $f(R_i^u, R_{-i})$ and $f(R_i^w, R_{-i})$, the pair i is part of an exchange by donating a left lobe. Therefore, whether i is willing or unwilling, i is indifferent between $f(R_i^u, R_{-i})$ and $f(R_i^w, R_{-i})$, implying the incentive-compatibility condition for i holds.

Case 2: In neither $f(R_i^u, R_{-i})$ nor $f(R_i^w, R_{-i})$, the pair i is part of an exchange by donating a left lobe. By individual rationality, this implies that i is unmatched at $f(R_i^u, R_{-i})$. Thus, $f(R_i^u, R_{-i}) R_i^u f(R_i^w, R_{-i})$, by the specification of preference R_i^u. At $f(R_i^w, R_{-i})$, the only individually rational match possibilities of i are: i directly donates a right lobe; i takes part in an exchange by donating a right lobe; or i is unmatched, which is the worst of these three options under R_i^w. Therefore, we also have $f(R_i^w, R_{-i}) R_i^w f(R_i^u, R_{-i})$ by the individual rationality of f, as desired. ■

Appendix B Proof of Theorem 1

Fix an exchange pool. In the rest of the paper, given an exchange pool, let $n(X-Y)$ be the number of $X-Y$ pairs for any $X-Y \in \mathcal{T} \times \mathcal{T}$. A maximum individually rational matching M_0 exists by finiteness of the problem.

Step 0: In all individually rational matchings all compatible pairs participate in direct transplants.

Step 1: Suppose that M_0 does not maximize the exchanges between $100-011$ and $011-100$ types. That is, the number of $100-011$ and $011-100$ types matched in M_0 is strictly less than $\min\{n(100-011), n(011-100)\}$. Let $\Delta > 0$ be the difference. Since $011-100$ types can only be matched to $100-011$ types, at least Δ many $011-100$ types are unmatched in M_0 and at least Δ many $100-011$ types are matched to other types. We can define a new matching M_0' by unmatching Δ many of those $100-011$ types and rematching them to the unmatched $011-100$'s. Then, M_0' continues to be maximum and maximizes the exchanges between $100-011$ and $011-100$ types.

Suppose that M_0' does not maximize the exchanges between $010-100$ and $100-010$ types. Then, the number of $010-100$ and $100-010$ types matched in M_0' is strictly less than $\min\{n(010-100), n(100-010)\}$. Let $\Delta > 0$ be the difference. From Figure 2 and the optimality of M_0', at least Δ many $010-100$ types are matched to $100-011$ types, and at least Δ many $100-010$ types are matched to $010-101$ types. We can define a new matching M_0'' by undoing these matches, rematching those $010-100$ and $100-010$ types to each other, and rematching those $100-011$ and $010-101$ types to each other. Then, M_0'' continues to be maximum and also maximizes the exchanges between $010-100$ and $100-010$ types.

By applying the above arguments to the other exchanges in Figure 3(a), we obtain a maximum matching M_1 that maximizes the exchanges in Step 1 of the matching algorithm.

Step 2: Fix the matches maximized in Step 1, and consider any submatching M^* among remaining types that maximizes the exchanges between the remaining $100-011$ types and the remaining types in $T = \{010-100, 001-100, 011-101, 011-100\}$. Let k and m denote the number of exchanges between $100-011$ types and types in T at the submatching M^* and at the matching M_1, respectively. Since the submatching M^* maximizes these exchanges after Step 1, $k \geq m$. At the matching M_1, unmatch the m matches between $100-011$ types and types in T, and unmatch an additional $k-m$ many $100-011$ types matched to $010-101$ or $001-110$ types. Then, rematch those k $100-011$ types with types in T as in the submatching M^*. The new matching M_1' obtained in this way, is maximum and maximizes the exchanges between the remaining $100-011$ types and the remaining types in T.

By applying the above argument to the other exchanges in Figure 3(b), we obtain a maximum matching M_2 that sequentially maximizes the exchanges in Steps 1 and 2 of the matching algorithm.

Step 3: Take any matching M that agrees with M_2 in the matches created in the first two steps. Since the only remaining exchanges after Steps 1 and 2 of the algorithm are those in Figure 3(c), the matching M is maximum if and only if it maximizes the exchanges in Figure 3(c) given the matches in the first two steps.

Appendix C Proof of Theorem 2

We prove Theorem 2 through three lemmas.

In the following proofs, by a *neighbor* of a type $X-Y$ (or *neighbor* of a pair i), we mean a type whose pairs can participate in an individually rational exchange with an $X-Y$ pair (or i).

Lemma 5 *The two-size left&right-lobe sequential matching mechanism is individually rational.*

Proof of Lemma 5: Let M be the outcome of the mechanism for a given willingness profile R and priority order. First, observe that no unwilling pair donates a right lobe in M. Category 0 pairs cannot be matched in any individually rational matching (by Lemma 2) and are not matched in M, either. All Category I pairs receive direct left-lobe transplants in M in Step 0. Category II and VI pairs cannot participate in individually rational direct left-lobe donation (by Lemma 2). All willing Category II pairs and Category VI pairs that cannot be matched in feasible exchanges by donating a left lobe receive direct right-lobe transplants in M. Category III, V, and VI pairs cannot get individually rational direct transplants (by Lemma 2). Matching M does not match them in direct transplants, either. Given these individually rational direct transplants in M, the algorithm conducts only individually rational exchanges by construction, and thus, M is individually rational. ■

Lemma 6 *The two-size left&right-lobe sequential matching mechanism is Pareto efficient.*

Proof of Lemma 6: Let $R \in R$ be a willingness profile, and let a priority order be fixed. Let M be the outcome of the mechanism and $M' \in \mathcal{M}_c[\mathcal{I}]$ be a Pareto-efficient matching such that for all $i \in \mathcal{I}$, $M' R_i M$, i.e., M' weakly Pareto dominates M. We will show that all pairs are indifferent between M and M'.

Our proof strategy will be sequentially constructing a set of matchings $M_0 := M'$, M_{1b}, M_{2a}, M_{2b}, M_{2c}, and M_{2d} such that M_s agrees with M for all pairs matched in steps up to and in Step s of the algorithm, and all pairs are indifferent between $M' = M_0$ and M_s.

Since M' and M are both individually rational (the individual rationality of M is implied by Lemma 5, and thus, the individual rationality of M' is implied by the fact that M' weakly

Pareto dominates M), all Category I pairs have to receive direct left-lobe donation, and all willing Category II pairs have to receive direct right-lobe donation in both matchings by Lemma 2. Thus, pairs matched in Step 0 of the algorithm are indifferent between $M_0 = M'$ and M.

Without loss of generality assume that $n(010 - 100) \geq n(100 - 010)$ in the rest of this proof.

Construction of M_{1b}: We construct M_{1b} as follows, by changing the matches of pairs matched in Step 1b. Recall that 0-waste exchanges are conducted in this step. These are either (a) value -1 & value +1 exchanges or (b) value 0 & value 0 exchanges. Suppose we clear them in this order in Step 1b, value -1 & value +1 first and value 0 & value 0 next.

(a) Value -1 & value +1 exchanges are cleared: Consider an exchange between pair i with $\tau(i) = 101 - 010$ and pair j with $\tau(j) = 010 - 101$ chosen by M but not by $M_0 = M'$. Since i is matched to donate a left lobe in M, it is also matched to donate a left lobe in M', as M' weakly Pareto dominates M. So is j.
Let $M'(i) = h$ and $M'(j) = \ell$ for some pairs h and ℓ. Either (i) $\tau(h) = 010 - 101$, or (ii) $\tau(h) = 010 - 100$, and h is willing and matched to donate a right lobe to i in M' (see Figure 4, $010 - 101$ is the only auxiliary neighbor of $101 - 010$). Thus, pair j and pair h are either of the same type or pair h can be transformed to pair j's type. Since $M'(j) = \ell$, pairs h and ℓ can also be matched with each other.
Define
$$M'' := \Big[M' \setminus \Big\{\{i,h\},\{j,\ell\}\Big\}\Big] \cup \Big\{\{i,j\},\{h,\ell\}\Big\}.$$
For pair h, $M''\ R_h\ M'$. To see this: If $\tau(h) = \tau(j) = 010 - 101$ then h is matched to donate a left lobe in both M'' and M'. If $\tau(h) = 010 - 100$, then h is willing and matched to donate a right lobe in M'.
For pair ℓ, $M''\ R_\ell\ M'$. To see this: If ℓ donates a left lobe in M' to j, which is of type $010-101$, then it can donate feasibly a left lobe to $h = M''(\ell)$ as well, as $\tau(h) = 010-101$, or $\tau(h) = 010 - 100$ and h is willing.
Now all pairs weakly prefer M'' to $M_0 = M'$. Since M' is Pareto efficient, then they should be Pareto indifferent.
We use a similar construction to obtain a new matching from M'' (instead of M') for any other value -1 & value +1 exchange cleared in Step 1 and not chosen in M' (i.e., those exchanges of auxiliary types $110 - 001$ & $001 - 110$ and $011 - 100$ & $100 - 011$), and iteratively continue.[21] Suppose M'_{1b} is the outcome of this (a) part. Now, all pairs are indifferent between M'_{1b} and $M' = M_0$, and all value -1 & value 1 exchanges in M also exist in M'_{1b}.

(b) Value 0 & value 0 exchanges are cleared: Consider an exchange between pair i with

[21] The only difference in the argument is that an auxiliary $110 - 001$ type can only be matched with a pair of its reciprocal type in M'.

$\tau(i) = 010 - 100$ and pair j with $\tau(j) = 100 - 010$ chosen by M but not M'_{1b}. As M' weakly Pareto dominates M, both i and j are matched, and they both donate left lobes in M' as well. Suppose $M'(i) = h$ and $M'(j) = \ell$ for some pairs h and ℓ. Then $\tau(h) = 100 - 011$ or $\tau(h) = 100 - 010$, and $\tau(\ell) = 010 - 100$ or $\tau(\ell) = 010 - 101$ (see Figure 4). Observe that patients of h and ℓ are left-lobe compatible with each other's donors. Therefore, we can form a matching

$$M''' := \Big[M'_{1b} \setminus \Big\{\{i,h\},\{j,\ell\}\Big\}\Big] \cup \Big\{\{i,j\},\{h,\ell\}\Big\}$$

such that all pairs are indifferent between M''' and M'_{1b} (and hence, M' by Part (a)). We use a similar construction to obtain a new matching from M''' (instead of M'_{1b}) for any other value 0 & value 0 exchange cleared in Step 1 (i.e., exchanges involving pairs of types $001 - 100$, $001 - 010$, $011 - 101$, $011 - 110$, and $101 - 110$) and not picked in M', and iteratively continue. Suppose M_{1b} is the outcome of this step. Now, all pairs are indifferent between M_{1b} and $M' = M_0$, and all Step 1b exchanges in M also exist in M_{1b}.

Construction of M_{2a}: The remaining highest-priority type $010 - 100$ pairs are matched with the remaining highest-priority type $100 - 011$ pairs in M in Step 2a.

Take the highest-priority pair i of type $010 - 100$ such that $M(i) = j$ for some pair j with $\tau(j) = 100 - 011$ and $M_{1b}(i) \neq j$ if such a pair exists. Now $M_{1b}(i) = h$ and $M_{1b}(j) = \ell$ for some pairs h and ℓ such that i and j donate left lobes because M_{1b} weakly Pareto dominates M. Since pair i is still available in Step 2, all pairs of i's reciprocal type $100 - 010$ should have been exhausted in Step 1b. Because M_{1b} and M coincide for all pairs matched in Step 1 by construction of M_{1b}, no type $100 - 010$ pairs could be matched with i in M_{1b}. Thus, the only possible type for h is $100 - 011$ (h is not of type $100 - 010w$, because such pairs are exhausted in Step 1; see Figure 6(a)), i.e., the same type of j. Thus the following is a feasible matching:

$$M'' := \Big[M_{1b} \setminus \Big\{\{i,h\},\{j,\ell\}\Big\}\Big] \cup \Big\{\{i,j\},\{h,\ell\}\Big\}$$

such that all pairs are indifferent between M'' and M_{1b} (and thus, $M_0 = M'$), and M'' agrees with M for all pairs matched before i and j in the algorithm and also for pairs i and j.

We repeat the above procedure for all such remaining i, starting with M'' instead of M_{1b}. Let M_{2a} be the final outcome of this procedure.

Construction of M_{2b}: In Step 2b, if willing type $010 - 100$ pairs are left they are transformed to donate a right lobe and matched with type $101 - 010$ pairs. Suppose that some willing pair i with $\tau(i) = 010 - 100$ is matched to donate its right lobe to some pair j with $\tau(j) = 101 - 010$ in Step 2b of the algorithm so that $M(i) = j$ but $M_{2a}(i) = h \neq j$ (see Figure 6(b)). Moreover, let i be the highest-priority pair with this property. By construction of M_{1b} and M_{2a}, all pairs, with which i can form an exchange by donating its donor's left lobe, are matched with other pairs in M_{2a}. Thus, i has to be matched to donate a right lobe in M_{2a}, by the construction of M_{1b} and M_{2a} and that M_{2b} weakly Pareto dominates M. Hence, $\tau(h)$ is a type that can be matched with $010-101$.

Suppose $M_{2a}(j)=\ell$ for some pair ℓ. Since j is matched to donate a left lobe in M, it should also be matched to donate a left lobe in M_{2a}, which weakly Pareto dominates M. Thus, the only options for ℓ's type are $010-101$ or $010-100w$, since pair of type $\tau(j)=101-010$ can donate a left lobe in an exchange with a pair of only one of these types (see Figure 4). Hence, pairs h and ℓ can be matched with each other. We construct

$$M'' := \Big[M_{2a} \setminus \Big\{\{i,h\},\{j,\ell\}\Big\}\Big] \cup \Big\{\{i,j\},\{h,\ell\}\Big\}$$

as a feasible matching such that all pairs are indifferent between M_{2a} (and thus, M') and M'', and M'' agrees with M for all pairs matched before i and j in the algorithm and also for pairs i and j.

We repeat the above procedure for all such remaining i, starting with M'' instead of M_{2a}. Let M_{2b} be the final outcome of this procedure.

Construction of M_{2c}: Observe that pairs of both types $101-011$ and $011-101$ could not have remained in the pool simultaneously until Step 2c of the algorithm (as such pairs would have been matched in Step 1b with each other). Thus, either $101-011$ or $011-101$ type pairs (or both) have been depleted in Step 1.

Take the highest-priority pair i matched in Step 2c such that $j=M(i)\neq M_{2b}(i)=h$ for some pairs j and h. Let $M_{2b}(j)=\ell$ for some pair ℓ. Such a pair ℓ exists as M_{2b} weakly Pareto dominates M and j is matched in M.

Without loss of generality, let i be of auxiliary type $011-101$ and j be of auxiliary type $101-011$ such that at least one is matched to donate a right lobe. Because i and j remained unmatched until Step 2c and because we showed that all pairs matched in Steps 0, 1, 2a, and 2b have the same matches in both M and M_{2b}, we have the following cases:

1. i can be of two types:
   (a) $011-101$: Since i is not matched until Step 2c, all pairs of its neighbor $101-011$ are matched in Step 1b in M. In particular, these pairs are matches different from i in M_{2b}. Hence, h has to belong to one of the two remaining neighbors of i, $101-010w$ or $100-011$.
   (b) $011-100w$: Since i is not matched until Step 2c, all pairs of its neighbors $100-011$ and $100-010w$ are matched in Step 1b in M. In particular, these pairs are not matched with i in M_{2b}. Thus, h has to belong to one of the two remaining neighbors of i, $101-010w$ or $101-011$.
2. j can be of two types:

(a) $101-011$: Since j is not matched until Step 2c, all pairs of its neighbor $011-101$ are matched in Step 1b in M. In particular, these pairs are matched in matches different from j in M_{2b}. Hence, ℓ has to belong to one of the three remaining neighbors of j, $011-100w$, $010-100w$, or $010-101$.

(b) $101-010w$: Since j is not matched until Step 2c, all pairs of its neighbor $010-101$ are matched in Step 1b, and those of $010-100w$ are matched in Steps 1b and 2b in M. In particular, these neighboring-type pairs with j are not matched with j in M_{2b}. Thus, ℓ has to belong to type $011-101$ or type $011-100w$, the only remaining neighbors of j.

Now, regarding i and j together, 3 of the above possible 4 combinations can occur at the same time: 1(a) and 2(b), 1(b) and 2(a), or 1(b) and 2(b). Thus, in all cases $\{h, \ell\}$ is a feasible exchange such that either h and ℓ each improve or remain indifferent with respect to M_{2b}. Since M_{2b} is Pareto efficient,

$$M'' := \Big[M_3 \setminus \Big\{\{i,h\},\{j,\ell\}\Big\}\Big] \cup \Big\{\{i,j\},\{h,\ell\}\Big\}$$

leaves every pair indifferent with respect to M_{2b} (and thus, M'). Moreover, M'' agrees with M for all pairs matched before i and j in the algorithm and also for pairs i and j. We repeat the above procedure for all such remaining i starting with M'' instead of M_{2b}. Let M_{2c} be the final outcome of this procedure.

Construction of M_{2d}: Take pair i of one of the auxiliary types $100-011$, $010-101$, or $001-110$ such that $M(i) = j \neq h = M_{2c}(i)$ and i is the highest-priority pairs among such pairs. Let $M_{2c}(j) = \ell$. Such a pair ℓ exists, as M_{2c} weakly Pareto dominates M and j is matched in M. Observe that by construction of M_{1b}, M_{2a}, M_{2b}, and M_{2c}, all pairs matched in steps before Step 2d in the algorithm have the same matches in M_{2c} and M. Three cases are possible (See Figure 6(d)):

1. $\tau(i) = 100-011$: Then j belongs to one of the following types: $011-101/011-100w$ (i.e., originally of type $011-101$ or transformed to $011-101$ and originally of type $011-100w$), $011-110$, and $001-100$. If j belongs to $011-101/011-100w$, then pairs of types $101-010w$ and $101-011$, which are the other two neighbors with j besides i's type, are exhausted in Steps 1b, 2b, and 2c. If j is of type $011-110$ or $001-100$, then the other auxiliary neighbor with j besides i's type, which are $110-011$ and $100-001$, are exhausted in Step 1b. Thus, $\tau(\ell) = \tau(i)$.
2. $\tau(i) = 010-100$ and i is willing, or $\tau(i) = 010-101$: Then j belongs to one of the types (i) $001-010$, (ii) $101-110$, or (iii) $101-010w/101-011$. In each of these cases, observe, in Figure 6(d), that pairs of all neighbors of j have been exhausted in (i) Step 1b, (ii) Step 1b, or (iii) Steps 1b, 2b, and 2c, respectively, except those of auxiliary type $010-101$. Thus, ℓ's and i's auxiliary types have to be the same.
3. $\tau(i) = 001-110$: Then j belongs to one of the four auxiliary types, $110-011$, $110-101$,

$100-001$, and $010-001$. Thus, all of j's potential matches except those with pairs of type $001-110$ have been exhausted in Step 1b. Thus, $\tau(\ell)=\tau(i)$.

In each case, $\{h,\ell\}$ is a feasible exchange, and

$$M'' := \Big[M_{2c} \setminus \Big\{\{i,h\},\{j,\ell\}\Big\}\Big] \cup \Big\{\{i,j\},\{h,\ell\}\Big\}$$

leaves every pair indifferent with respect to M_{2c} (and thus, M'). Moreover, M'' agrees with M for all pairs matched before i and j in the algorithm and also for pairs i and j. We repeat the above procedure for all such remaining pairs starting with M'' instead of M_{2c}. Let M_{2d} be the final outcome of this procedure.

We are ready to finish the proof of the theorem. By the algorithm, if a pair i of auxiliary type $X-Y \in \mathcal{V}_1 = \{100-011, 010-101, 001-110\}$ (i.e., the set of value $+1$ auxiliary types) was not matched in Steps 1b, 2a, 2b, 2c, and 2d, then no pairs of its neighbors except the ones belong to the auxiliary types in $\mathcal{V}_1 \setminus \{X-Y\}$ have remained available until Step 3. Thus, to create a Pareto-efficient matching among the pairs remaining in this step, we need to maximize the number of exchanges among them. As the algorithm exactly does this and M_{2d} weakly Pareto dominates M while all the matches prior to Step 3 are identical between M_{2d} and M, M_{2d} should be Pareto indifferent to M for the pairs matched until and in Step 3. Since in Step 4, unmatched and willing Category VI pairs are matched to have direct right-lobe transplants, M is Pareto efficient. ■

Lemma 7 *The two-size left&right-lobe sequential matching mechanism is incentive compatible.*

Proof of Lemma 7: Let $R \in \mathcal{R}$ be a willingness profile, and let a priority order be fixed.

First observe the following: In each step of the algorithm, we clear exchanges involving a given (auxiliary) type according to the exogenously fixed priority order. Moreover, for each pair, the lobe it will donate is uniquely defined once we clear them in an exchange or a direct transplant. Thus, a pair's willingness type revelation does not affect whether it will be matched by donating a left lobe in a step given that it is available at that step in the algorithm after all required transformations are done in that step. We refer to this argument as (*) in the proof below.

We consider the incentives faced by each pair with respect to their categories:

Category 0 pairs have no possibility of being matched by Lemma 2. Each Category I pair is matched in Step 0a by left-lobe donation in a direct transplant. Thus, pairs of these three types are indifferent between being and not being truthful about their willingness types.

Each Category II pair can only be matched through a direct right-lobe transplant by Lemma 2. Since we match the willing Category II pairs in Step 0c of the algorithm in direct right-lobe transplants, it is a strictly dominant strategy for each Category II pair to be truthful about its willingness type.

Each Category III pair is either matched in an exchange by donating a left lobe or left unmatched by Lemma 2. By Argument (*), it is indifferent between being and not being truthful about its willingness type.

Each Category IV pair can only be matched through an exchange and only by donating its right lobe by Lemma 2. Since we transform the willing Category IV pairs in Step 1a and make them available for exchange through a right-lobe donation, no Category IV pair will be worse off by being truthful about its willingness type.

Each Category V pair can only be matched through an exchange and, in particular, by donating a left lobe or, if it is willing, by donating a right lobe by Lemma 2. We explore the 4 Category V types, $100-010$, $010-100$, $011-100$, and $101-010$ individually. Without loss of generality, assume that $n(010-100) \geq n(100-010)$.

- Each type $100-010$ pair is matched in an exchange with a type $010-100$ pair in Step 1b by donating a left lobe regardless of its willingness type as $n(010-100) \geq n(100-010)$ (see Figure 5).
- A type $010-100$ pair can donate a left lobe only to a type $100-011$ or $100-010$ pair in an exchange. Type $010-100$ pairs are matched in Step 1b and Step 2a with these two types, respectively. By Argument (*), a type $100-010$ pair is matched in Step 1b or 2a while reporting willing if and only if it is matched in Step 1b or 2a while reporting unwilling. The remaining type $010-100w$ pairs are transformed for a right-lobe donation in Step 2b only after all of their left-lobe-donation possibilities are exhausted (see Figures 5 and 6). Thus, type $010-100$ pairs that are still unmatched at the beginning of Step 2b are weakly better off being truthful about their willingness type.
- A type $011-100$ pair can donate a left lobe only to a type $100-011$ or $100-010w$ pair in an exchange. In Step 1b, $011-100$ pairs are matched with $100-011$ type pairs. On the other hand, all type $100-010$ pairs are exhausted in Step 1b; therefore, none of them are transformed and matched with $011-100$ pairs. By Argument (*), a type $011-100$ pair is matched in Step 1b while reporting willing if and only if it is matched in Step 1b while reporting unwilling. The remaining type $011-100w$ pairs are transformed in Step 2c, only after their all left-lobe-donation possibilities are exhausted (see Figures 5 and 6). Thus, type $011-100$ pairs that are still unmatched at the beginning of Step 2c are weakly better off being truthful about their willingness type.
- A type $101-010$ pair can donate a left lobe only to a type $010-101$ or $010-100w$ pair in an exchange. By Argument (*), a type $101-010$ pair is matched in Step 1b or 2b while reporting willing if and only if it is matched in Step 1b or 2b while reporting unwilling. The remaining type $101-010w$ pairs are transformed in Step 2c, only after their left-lobe-donation possibilities are exhausted (see Figures 5 and 6). Thus, type

101 − 010 pairs that are still unmatched at the beginning of Step 2c are weakly better off being truthful about their willingness type.

Each Category VI pair can only be matched in an individually rational exchange by donating a left lobe or, if it is willing, it can be matched in a direct right-lobe transplant by Lemma 2. Category VI type pairs are only transformed after all exchanges are cleared in the algorithm in Step 4. Therefore, they are best off by being truthful as well. ■

References

Abdulkadiroğlu, Atila and Tayfun Sönmez (2003). "School choice: A mechanism design approach." *American Economic Review*, 93 (3), 729–747.

Agarwal, Nikhil, Eduardo Azevedo, Itai Ashlagi, Clayton Featherstone, and Ömer Karaduman (2018). "Market failure in kidney exchange." Working paper.

Andersson, Tommy and Jörgen Kratz (2017). "Kidney exchange over the blood group barrier." Working paper.

Budish, Eric and Estelle Cantillon (2012). "The multi-unit assignment problem: Theory and evidence from course allocation at Harvard." *American Economic Review*, 102 (5), 2237–71.

Chan, See Ching, Chung Mau Lo, Boon Hun Yong, et al. (2010). "Paired donor interchange to avoid ABO-incompatible living donor liver transplantation." *Liver Transplantation*, 16 (4), 478–481.

Delacrétaz, David, Scott Duke Kominers, and Alexander Teytelboym (2017). "Refugee resettlement." Working paper.

Dickerson, John P. and Tuomas Sandholm (2014). "Multi-organ exchange: The whole is greater than the sum of its parts." *Proceedings of the Association for the Advancement of Artificial Intelligence*.

Diverse Populations Collaborative Group (2005). "Weight-height relationships and body mass index: Some observations from the diverse populations collaboration." *American Journal of Physical Anthropology*, 128 (1), 220–229.

Echenique, Federico and M. Bumin Yenmez (2015). "How to control controlled school choice." *American Economic Review*, 105 (8), 2679–2694.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz (2007). "Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords." *American Economic Review*, 97 (1), 242–259.

Edmonds, Jack (1965). "Paths, trees, and flowers." *Canadian Journal of Mathematics*, 17, 449–467.

Ergin, Haluk, Tayfun Sönmez, and M. Utku Ünver (2017). "Dual-donor organ exchange." *Econometrica*, 85, 1645–1671.

Gale, David and Lloyd S. Shapley (1962). “College admissions and the stability of marriage.” *American Mathematical Monthly*, 69, 9–15.

Grady, Dennise (2002). “Donor’s death at hospital halts some liver surgeries.” *New York Times*, January, 16.

Hafalir, Isa E., M. Bumin Yenmez, and Muhammed A. Yildirim (2013). “Effective affirmative action in school choice.” *Theoretical Economics*, 8 (2), 325–363.

Hatfield, John W. and Paul Milgrom (2005). “Matching with contracts.” *American Economic Review*, 95 (4), 913–935.

Hatfield, John William and Fuhito Kojima (2010). “Substitutes and stability for matching with contracts.” *Journal of Economic Theory*, 145 (5), 1704 –1723.

Hwang, Shin, Sung-Gyu Lee, Deok-Bog Moon, et al. (2010). “Exchange living donor liver transplantation to overcome abo incompatibility in adult patients.” *Liver Transplantation*, 16 (4), 482–490.

Jones, Will and Alexander Teytelboym (2017). “The local refugee match: Aligning refugees’ preferences with the capacities and priorities of localities.” *Journal of Refugee Studies*, fex022.

Kojima, Fuhito (2012). “School choice: Impossibilities for affirmative action.” *Games and Economic Behavior*, 75, 685–693.

Korte, Bernhard and Jens Vygen (2011). *Combinatorial Optimization: Theory and Algorithms*. Springer-Verlag Berlin Heidelberg.

Lee, Sung-Gyu (2010). “Living-donor liver transplantation in adults.” *British Medical Bulletin*, 94 (1), 33–48.

Milgrom, Paul (2000). “Putting auction theory to work: The simultaneous ascending auction.” *Journal of Political Economy*, 108 (2), 245–272.

Mishra, Ashish, Alexis Lo, Grace S. Lee, et al. (2018). “Liver paired exchange: Can the liver emulate the kidney?” *Liver Transplantation*, 24 (5), 677–686.

Moraga, Jesús Fernández-Huertas and Hillel Rapoport (2014). “Tradable immigration quotas.” *Journal of Public Economics*, 115, 94–108.

Nicolò, Antonio and Carmelo Rodriguez-Álvarez (2017). “Age-based preferences: Incorporating compatible pairs into paired kidney exchange.” *Games and Economic Behavior*, forthcoming.

Okumura, Yasunori (2014). “Priority matchings revisited.” *Games and Economic Behavior*, 88, 242–249.

Rapaport, F. T. (1986). “The case for a living emotionally related international kidney donor exchange registry.” *Transplantation Proceedings*, 18 (3), 5–9.

Roth, Alvin E. and Elliot Peranson (1999). “The redesign of the matching market for American physicians: Some engineering aspects of economic design.” *American Economic Review*, 89 (4), 748–780.

Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2004). "Kidney exchange." *Quarterly Journal of Economics*, 119 (2), 457–488.

Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2005). "Pairwise kidney exchange." *Journal of Economic Theory*, 125 (2), 151–188.

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.

Samstein, Benjamin, Inmaculada Melo-Martin, Sandip Kapur, Lloyd Ratner, and Jean Emond (2018). "A liver for a kidney: Ethics of trans-organ paired exchange." *American Journal of Transplantation*, 18 (5), 1077–1082.

Schrijver, Alexander (2003). *Combinatorial Optimization: Polyhedra and Efficiency*. Springer-Verlag Berlin Heidelberg.

Schummer, James and Azar Abizada (2017). "Incentives in landing slot problems." *Journal of Economic Theory*, 170, 29 –55.

Schummer, James and Rakesh V. Vohra (2013). "Assignment of arrival slots." *AEJ: Microeconomics*, 5 (2), 164–185.

Sönmez, Tayfun and M. Utku Ünver (2010). "Course bidding at business schools." *International Economic Review*, 51 (1), 99–123.

Sönmez, Tayfun, M. Utku Ünver, and M. Bumin Yenmez (2018). "Incentivized kidney exchange." Working paper.

Um, Eun-Hae, Shin Hwang, Gi-Won Song, et al. (Nov. 2015). "Calculation of standard liver volume in Korean adults with analysis of confounding variables." *Korean Journal of Hepatobiliary Pancreatic Surgery*, 19 (4), 133–138.

Varian, Hal R. (2007). "Position auctions." *International Journal of Industrial Organization*, 25 (6), 1163 –1178.

Intended for Online-Only Publication

Appendix D Proofs of Results in Section 6

D.1 Mathematical Preliminaries

In this section, we will state some definitions and a result from graph theory that will be used in subsequent proofs.

A tuple $G = (\mathcal{V}, E)$ is a **graph** if $\mathcal{V}$ is a nonempty set and $E \subseteq \Big\{\{x, y\} : x, y \in \mathcal{V}\Big\}$. The elements of $\mathcal{V}$ are called **vertices**. The elements of E are called **edges**.

Note that in the definition of a graph, we are allowing for loops, i.e., edges $\{x, y\}$ such that $x = y$.[22]

A **matching** in a graph $G = (\mathcal{V}, E)$ is a subset $M \subseteq E$ of pairwise disjoint edges, i.e., $e, e' \in M$ such that $e \cap e' \neq \emptyset \implies e = e'$. Given a matching M in G, we will abuse notation and also define the function $M : \mathcal{V} \to \mathcal{V} \cup \{\emptyset\}$ by:

$$M(x) = \begin{cases} y & \text{if there exists } y \in \mathcal{V} \text{ such that } \{x, y\} \in M \\ \emptyset & \text{otherwise} \end{cases}$$

for all $x \in \mathcal{V}$. We call $M(x)$ the **match of x in M**. We will say that a subset $\mathcal{W} \subseteq \mathcal{V}$ is **matchable in** G, if there is a matching M in G such that $M(x) \neq \emptyset$ for all $x \in \mathcal{W}$.

In a graph, the vertices corresponding to each edge $e = \{x, y\}$ are unordered. We will also need the notion of a *directed graph* where the order of the vertices does matter.

A tuple $G = (\mathcal{V}, E)$ is a **directed graph (digraph)** if $\mathcal{V}$ is a nonempty set and $E \subseteq \{(x, y) \in \mathcal{V} \times \mathcal{V} : x \neq y\}$. When the digraph is understood, we will also use $x \to y$ to denote $(x, y) \in E$.

Note that as opposed to our definition of a compatibility graph, in the definition of a digraph, we are ruling out loops, i.e., directed edges (x, y) such that $x = y$.[23]

Given a digraph $G = (\mathcal{V}, E)$, a **topological order on** G is a linear order $\mathbb{L}$ on $\mathcal{V}$ such that: $x \to y$ implies $x \mathbb{L} y$, for all $x, y \in \mathcal{V}$.

[22] In some texts, a *simple undirected graph with loops* is what we call a graph here. See for example Korte and Vygen (2011, p13-14).

[23] In some texts, a *simple directed graph without loops* is what we call a digraph here. See again Korte and Vygen (2011, p13-14).

A digraph $G = (\mathcal{V}, E)$ is **acyclic** if there does not exist an integer $n \geq 2$ and $v_1, \ldots, v_n \in \mathcal{V}$ such that: $v_1 \rightarrow v_2 \rightarrow \ldots \rightarrow v_n \rightarrow v_1$.

The following lemma is a standard result in graph theory.[24]

Lemma 8 *Given a digraph $G = (\mathcal{V}, E)$, there exists a topological order on G if and only if G is acyclic.*

We continue with the proof of Lemma 4 in Subection 6.1.

Proof of Lemma 4: Suppose for a contradiction that the precedence digraph has a cycle:

$$X^0 - Y^0 \longrightarrow X^1 - Y^1 \longrightarrow \ldots \longrightarrow X^{n-1} - Y^{n-1} \longrightarrow X^0 - Y^0$$

where $n \geq 2$.

Note that for each $k \in \{0, 1, \ldots, n-1\}$:

$$X^k - Y^k \longrightarrow X^{\text{ mod }_n(k+1)} - Y^{\text{ mod }_n(k+1)} \longrightarrow X^{\text{ mod }_n(k+2)} - Y^{\text{ mod }_n(k+2)}$$

implies that $X_3^k \leq Y_3^{\text{mod }_n(k+1)}$. It also implies that $Y_3^{\text{mod }_n(k+1)} < X_3^{\text{mod }_n(k+2)}$ since $Y^{\text{ mod }_n(k+1)} \not\geq X^{\text{mod }_n(k+2)}$ and $\rho(Y^{\text{ mod }_n(k+1)}) \geq X^{\text{mod }_n(k+2)}$. Therefore, $X_3^k < X_3^{\text{mod }_n(k+2)}$. That is, a patient along the cycle has a smaller size than the patient two steps ahead in the cycle. This can be used to obtain a contradiction in two separate cases:

Case 1 "n is even": $X_3^0 < X_3^2 < \ldots < X_3^{n-2} < X_3^0$.

Case 2 "n is odd": $X_3^0 < X_3^2 < \ldots < X_3^{n-1} < X_3^1 < X_3^3 < \ldots < X_3^{n-2} < X_3^0$. ■

D.2 Expanded Description of the Mechanism

We will now give an expanded description of the algorithm for the general sizes that will define an individually rational, Pareto-efficient, and incentive-compatible mechanism. While giving the definition, we will prove certain results and provide some discussion, which will explain why we chose these particular steps in this order.

In the rest of the section, fix a liver-exchange pool $(\mathcal{I}, \tau)$, a right-lobe size function ρ, a preference profile $R \in \mathcal{R}$, and an arbitrary linear order over $\mathcal{I}$ that we will interpret as the priority ranking of the pairs. Let $\mathcal{I}_V \subseteq \mathcal{I}$ denote the set of Category V pairs. Extend the precedence digraph on the types of Category V pairs to the pairs themselves as follows: Suppose (τ_V, ε_V) is the precedence digraph of the existing Category V types in the pool. Construct a digraph $(\mathcal{I}_V, E_V)$ such that for any two types $X - Y, U - V \in \tau_V$ and for any two

[24] For example, see Proposition 2.9 in Korte and Vygen (2011, p20).

pairs $i, j \in \mathcal{I}_V$ such that $\tau(i) = X-Y$ and $\tau(j) = U-V$: $(i,j) \in E_V \iff (X-Y, U-V) \in \varepsilon_V$. We refer to $(\mathcal{I}_V, E_V)$ as the precedence digraph on $\mathcal{I}_V$. By Lemma 4, the precedence digraph on $\mathcal{I}_V$ is also acylic. By Lemma 8, we can fix a topological order of the precedence digraph on $\mathcal{I}_V$. Let $\mathcal{I}_V = \{i_1, \ldots, i_K\}$ be an enumeration of Category V pairs with respect to the topological order.

The algorithm is separated into four main steps in four subsections. The second one is the key step. The succinct version of the algorithm presented in Section 6 and the algorithm below are equivalent.

D.2.1 Step 0

One of our objectives is to ensure that the outcome of the algorithm is individually rational. As an implication, we can immediately conclude, using Lemma 2, that some categories of pairs have to be unmatched and some have to be involved in direct donation, independently of the rest of the liver-exchange pool. In Step 0 of the algorithm, we determine these matches:

1. Leave all Category 0 pairs unmatched.
2. Let all Category I pairs directly donate a left lobe.
3. For any Category II pair i
   (a) If i is unwilling, leave i unmatched.
   (b) If i is willing, let i directly donate a right lobe.
4. Leave all unwilling Category IV pairs unmatched.

Let M_0 be the matching in $G_c[\mathcal{I}]$ that corresponds to the above direct donations. Formally:

$$M_0 = \Big\{\{i\} : i \in \mathcal{I} \text{ is Category I}\Big\} \cup \Big\{\{i\} : i \in \mathcal{I} \text{ is Category II and } R_i = R_i^w\Big\}.$$

D.2.2 Step 1

Let $\mathcal{I}_0$ denote the set of pairs whose match (possibly $\emptyset$) is not determined in Step 0. That is, $\mathcal{I}_0$ consists of all Category III pairs, all willing Category IV pairs, all Category V pairs, and all Category VI pairs. Suppose $\mathcal{I}_{IVw} \subseteq \mathcal{I}$ is the set of willing Category IV pairs.

In this part of the algorithm, we will define a sequence of graphs $G_0 = (\mathcal{I}_0, E_0)$, $G_1 = (\mathcal{I}_0, E_1)$,. . ., $G_K = (\mathcal{I}_0, E_K)$ on $\mathcal{I}_0$, where the edge sets are nondecreasing: $E_0 \subseteq E_1 \subseteq \ldots \subseteq E_K$. We will also define two nondecreasing sequences of subsets $\emptyset = \mathcal{J}_0 \subseteq \mathcal{J}_1 \subseteq \ldots \subseteq \mathcal{J}_K$ and $\emptyset = \widetilde{\mathcal{J}}_0 \subseteq \widetilde{\mathcal{J}}_1 \subseteq \ldots \subseteq \widetilde{\mathcal{J}}_K$ of Category V pairs.

The initial graph G_0 has edges that represent exchanges in which all participating Category III, V, and VI pairs donate a left lobe, and all (willing) Category IV pairs donate a right lobe. Formally,

$$E_0 := E_c[\mathcal{I}_{IVw}] = \left\{ \{i,j\} \in E_c[\mathcal{I}] : \begin{array}{c} i,j \in \mathcal{I}_0,\ i \neq j \\ t(i,j) = r \ \Rightarrow\ i \in \mathcal{I}_{IVw} \\ t(j,i) = r \ \Rightarrow\ j \in \mathcal{I}_{IVw} \end{array} . \right\}$$

We will construct a maximal subset $\mathcal{J}_K$ of $\mathcal{I}_V$ matchable in G_K, the final graph of this step. We will commit to every pair in $\mathcal{J}_K$ that it will participate in an exchange by donating a left lobe, without specifying until later (the end of Step 2) which pair it enters the exchange with. The maximality of $\mathcal{J}_K$ means that, in addition to pairs in $\mathcal{J}_K$, we cannot match any pair $i \in \mathcal{I}_V \setminus \mathcal{J}_K$ via an exchange where i donates a left lobe. Therefore, given our commitment to pairs in $\mathcal{J}_K$, the only match possibility for the willing Category V pairs in $\mathcal{I}_V \setminus \mathcal{J}_K$ is when they donate a right lobe. We let $\widetilde{\mathcal{J}}_K \subseteq \mathcal{I}_V \setminus \mathcal{J}_K$ be such willing Category V pairs. We iteratively transform each of these pairs to include the exchanges where it can donate a right lobe to the compatibility graph. We proceed inductively for this construction:

We next explain the algorithm for Step 1. Below, k runs though $1, \ldots, K$.

Step $1.k$**:** Is $\mathcal{J}_{k-1} \cup \{i_k\}$ matchable in G_{k-1}?

$$\begin{array}{llllllll} \textbf{YES} & \text{Let} & \mathcal{J}_k & := & \mathcal{J}_{k-1} \cup \{i_k\} & \text{and} & \widetilde{\mathcal{J}}_k & := \ \widetilde{\mathcal{J}}_{k-1} \\ \textbf{NO} & \text{Let} & \mathcal{J}_k & := & \mathcal{J}_{k-1} & \text{and} & \widetilde{\mathcal{J}}_k & := \left\{ \begin{array}{ll} \widetilde{\mathcal{J}}_{k-1} \cup \{i_k\} & \text{if } R_{i_k} = R^w_{i_k} \\ \widetilde{\mathcal{J}}_{k-1} & \text{otherwise.} \end{array} \right. \end{array}$$

Define the graph $G_k = (\mathcal{I}_0, E_k)$ by:

$$E_k := E_c[\mathcal{I}_{IVw} \cup \widetilde{\mathcal{J}}_k] = \left\{ \{i,j\} \in E_c[\mathcal{I}] : \begin{array}{c} i,j \in \mathcal{I}_0,\ i \neq j \\ t(i,j) = r \ \Rightarrow\ i \in \mathcal{I}_{IVw} \cup \widetilde{\mathcal{J}}_k \\ t(j,i) = r \ \Rightarrow\ j \in \mathcal{I}_{IVw} \cup \widetilde{\mathcal{J}}_k \end{array} \right\}$$

If $k < K$, go to Step $1.(k+1)$. If $k = K$, Step 1 of the algorithm is over.

Note that the sets $\mathcal{J}_k$, $\widetilde{\mathcal{J}}_k$, E_k, defined inductively by the above algorithm, are indeed nondecreasing in k as stated earlier in the section.

D.2.3 Step 2

In this step, we will determine a maximal subset of $\mathcal{I}_0$ that contains $\mathcal{J}_K$ and is matchable in G_K. Let N denote the number of pairs in $\mathcal{I}_0 \setminus \mathcal{J}_K$. Enumerate those pairs with respect to the priority order as $\{i^*_1, \ldots, i^*_N\}$. We will define an increasing sequence of subsets of this set $\mathcal{J}^*_0 \subseteq \mathcal{J}^*_1 \subset \ldots \mathcal{J}^*_N$. Let $\mathcal{J}^*_0 := \emptyset$. Below, n runs through $1, \ldots, N$.

Step $2.n$**:** Is $\mathcal{J}_K \cup \mathcal{J}^*_{n-1} \cup \{i^*_n\}$ matchable in G_K?

YES Let $\mathcal{J}_n^* := \mathcal{J}_{n-1}^* \cup \{i_n^*\}$
NO Let $\mathcal{J}_n^* := \mathcal{J}_{n-1}^*$

If $n < N$, go to Step 2.$(n+1)$. If $n = N$, Step 2 of the algorithm is over.

The algorithm above returns a set $\mathcal{J}_N^*$, with the property that $\mathcal{J}_K \cup \mathcal{J}_N^*$ is a maximal subset of $\mathcal{I}_0$ that contains $\mathcal{J}_K$ and is matchable in G_K.

Let M_2 be any matching in G_K such that $M_2(i) \neq \emptyset$ for all $i \in \mathcal{J}_K \cup \mathcal{J}_N^*$. The following lemma summarizes how different pairs in $\mathcal{I}$ are matched in such an M_2:

Lemma 9 *Suppose that*

$$M_2 \textit{ is a matching in } G_K \textit{ s.t. } \forall i \in \mathcal{J}_K \cup \mathcal{J}_N^* : M_2(i) \neq \emptyset. \tag{1}$$

Then, in M_2, the pairs in $\mathcal{I}_0$ are matched as follows:

1. $\forall i \in \mathcal{J}_K$: *$i$ takes part in an exchange, donating a left lobe.*
2. $\forall i \in \mathcal{J}_N^*$, *one of the following holds:*
   (a) *i is a Category III pair and takes part in an exchange, donating a left lobe.*
   (b) *i is a willing Category IV pair and takes part in an exchange, donating a right lobe.*
   (c) *$i \in \widetilde{\mathcal{J}}_K$ and takes part in an exchange, donating a right lobe.*
   (d) *i is a Category VI pair and takes part in an exchange, donating a left lobe.*
3. $\forall i \in \mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$: *$i$ remains unmatched.*

There could be multiple matchings that satisfy Condition (1), but as implied by Lemma 9, all of them match pairs in $\mathcal{I}_0$ in the same way. An implication is that every pair i is indifferent among all matchings that satisfy Condition (1). Thus, we can select any matching M_2 that satisfies Condition (1).

One method to find such a matching M_2 and also to check matchability of a set of pairs $\mathcal{J} \subseteq \mathcal{I}_0$ on a graph $G = (\mathcal{I}_0, E)$, as we do in each Step 1.k and Step 2.n, in polynomial time is as follows: We do not have any direct transplants (i.e., loops) on any graph G that we consider in substeps of Steps 1 or 2 in the algorithm. Thus, we can use the (polynomial-time) edge-weighted matching algorithm of Edmonds (1965) to find an optimal edge-weighted matching with the following edge weights. We assign each pair $i \in \mathcal{I}_0$ a priority number $\pi(i) > 0$ such that if $j \in \mathcal{J}$ and $i \in \mathcal{I}_0 \setminus \mathcal{J}$, then $\pi(j) > \pi(i)$. For each $\{i, j\} \in E$, we define the edge weight $\Pi(\{i, j\}) = \pi(i) + \pi(j)$. Then a matching of G that is the outcome of Edmonds' edge-weighted matching algorithm for edge weights Π solves the integer-programming problem $\max_{M \in \mathcal{M}} \sum_{\{i,j\} \in M} \Pi(\{i, j\}) = \max_{M \in \mathcal{M}} \sum_{i: M(i) \neq \emptyset} \pi(i)$, where $\mathcal{M}$ denotes the set of matchings in the compatibility graph G.[25] Since all pairs in $\mathcal{J}$ have higher priority numbers than

[25] The equality follows from Okumura (2014).

the pairs in $\mathcal{I}_0 \setminus \mathcal{J}$, by Proposition 2 in Roth, Sönmez, and Ünver (2005), this matching maximizes the number of pairs matched in $\mathcal{J}$. Thus, all pairs in $\mathcal{J}$ are matched in this matching if and only if $\mathcal{J}$ is matchable in G. In the final substep of Step 2, Substep 2.N, by setting $\mathcal{J} := \mathcal{J}_K \cup \mathcal{J}_N^*$ and $G := G_K$, we can use the outcome of this above procedure as M_2.

Another implication of Lemma 9 is that in such an M_2, all Category V pairs in $\mathcal{J}_K$ take part in exchanges donating a left lobe, some Category V pairs in $\widetilde{\mathcal{J}}_K$ take part in exchanges donating a right lobe, and all other Category V pairs are unmatched.

D.2.4 Step 3

In this step, we let any Category VI pair that we left unmatched in previous steps directly donate a right lobe to themselves. Let M_3 be a matching in $G_c[\mathcal{I}]$ that corresponds to these direct donations. Formally:

$$M_3 = \Big\{\{i\} : i \in \mathcal{I}_0 \setminus \mathcal{J}_N^*, i \text{ is Category VI, and } R_i = R_i^w\Big\}.$$

The algorithm is over at the end of this step, returning the matching:

$$M = M_0 \cup M_2 \cup M_3$$

Given a priority order over pairs, and a topological order over pairs induced by the precedence digraph over Category V types, we define the **general left&right-lobe priority mechanism** $f^{l\&r} : \mathcal{R} \to \mathcal{M}_c[\mathcal{I}]$ by: For any preference profile $R \in \mathcal{R}$, $f^{l\&r}(R)$ is a matching computed by the algorithm in Section D.2.

D.3 Proofs

D.3.1 Proof of Lemma 9

We start by proving a lemma that we will use in proving Lemma 9.

Lemma 10 *Let $j \in \mathcal{I}_V \setminus \mathcal{J}_K$. Then, there is no matching M in G_K such that $t(i, M(i)) = l$ for all $i \in \mathcal{J}_K \cup \{j\}$.*

Proof: Take any $j \in \mathcal{I}_V \setminus \mathcal{J}_K$. Since j is Category V, $j = i_k$ for some $k \in \{1, \ldots, K\}$. Since $j \notin \mathcal{J}_K$, by Step 1 of the algorithm, $\mathcal{J}_{k-1} \cup \{j\}$ is not matchable in G_{k-1}.

Suppose for a contradiction that there exists a matching M in G_K such that $t(i, M(i)) = l$ for all $i \in \mathcal{J}_K \cup \{j\}$. Define a (smaller) matching M' in G_K by:

$$M' := \Big\{\{i, M(i)\} : i \in \mathcal{J}_{k-1} \cup \{j\}\Big\}.$$

Note that M' is not a matching in G_{k-1}, otherwise $\mathcal{J}_{k-1} \cup \{j\}$ would be matchable in G_{k-1}, contradicting the previous paragraph. Therefore, there exists $i \in \mathcal{J}_{k-1} \cup \{j\}$ such that:

$$\{i, M(i)\} \in E_K \setminus E_{k-1}.$$

Since $t(i, M(i)) = l$, this is only possible if $t(M(i), i) = r$ and $M(i) \in \widetilde{\mathcal{J}}_K \setminus \widetilde{\mathcal{J}}_{k-1}$.

Note that $i \in \mathcal{J}_{k-1} \cup \{j\}$ and $M(i) \in \widetilde{\mathcal{J}}_K \setminus \widetilde{\mathcal{J}}_{k-1}$ imply that $i = i_l$ and $M(i) = i_m$ for some $l, m \in \{1, \ldots, K\}$ such that $l \leq k \leq m$.

Note that $t(i, M(i)) = l$ and $t(M(i), i) = r$ imply that $M(i) \to i$ in the precedence digraph over Category V pairs. Then, $M(i) = i_m$ must be ranked higher than $i = i_l$ with respect to the topological order, i.e., $m < l$, a contradiction. ■

In the following, fix a matching M_2 that satisfies Condition (1) in Lemma 9. We prove Lemma 9 in five parts.

Proof of Parts 1 and 2(d) in Lemma 9: By definition of E_K, all edges in $G_K = (\mathcal{I}_0, E_K)$ correspond to exchanges where the only pairs that might donate a right lobe are those in $\widetilde{\mathcal{J}}_K$ and those that are (willing) Category IV. Since M_2 is a matching in G_K, and since $\mathcal{J}_K$ and $\widetilde{\mathcal{J}}_K$ are disjoint, all pairs in $\mathcal{J}_K$ must be matched through exchanges where they donate a left lobe. Similarly, all matched Category VI pairs must be matched through exchanges where they donate a left lobe. ■

Proof of Parts 2(a) and 2(c) in Lemma 9: To see part 2(a), suppose i is a Category III pair. Since $i \in \mathcal{J}_N^*$, and i can only be matched through an exchange donating a left lobe in the algorithm (as Category III pairs are never transformed and do not receive direct transplants), $t(i, M_2(i)) = l$.

To see part 2(c), let $j \in \mathcal{J}_N^* \cap \widetilde{\mathcal{J}}_K$. Since $j \in \widetilde{\mathcal{J}}_K$, we have that $j \in \mathcal{I}_V \setminus \mathcal{J}_K$. By Part 1 of Lemma 9, M_2 is a matching in G_K such that $t(i, M_2(i)) = l$ for all $i \in \mathcal{J}_K$. So by Lemma 10, $t(j, M_2(j)) \neq l$. Since $j \in \mathcal{J}_N^*$, j is matched in M_2, so j must be part of an exchange by donating a right lobe. ■

Proof of Part 2(b) in Lemma 9: Suppose for a contradiction that i is a willing Category IV pair, and that in M_2, i takes part in an exchange with some pair j by donating a left lobe. Let $X - Y = \tau(i)$ and $U - V = \tau(j)$. Since i is Category IV, $X > Y$. Since $j \in \mathcal{I}_0$, it is enough to consider two cases:

Case 1 "j is a Category III pair or a (willing) Category IV pair or a Category V pair": Since i and j take part in an exchange where i donates a left lobe, and j donates a left or right lobe, we have $\rho(V) \geq X > Y \geq U$. This contradicts $\rho(V) \not\geq U$ since j is Category III, IV, or V.

Case 2 "j is a Category VI pair": Since i and j take part in an exchange where both i and j donate left lobes, we have $V \geq X > Y \geq U$. This contradicts $V \ngeq U$ since j is Category VI. ■

Proof of "Part 2 ⇒ 2(a), 2(b), 2(c) or 2(d)" in Lemma 9: It is enough to show that there is no $j \in \mathcal{J}_N^*$ such that j is Category V and $j \notin \widetilde{\mathcal{J}}_K$. Suppose for a contradiction that there exists such a j. Then, $j \in \mathcal{I}_V \setminus \mathcal{J}_K$. By Part 1 of Lemma 9, M_2 is a matching in G_K such that $t(i, M_2(i)) = l$ for all $i \in \mathcal{J}_K$. So by Lemma 10, $t(j, M_2(j)) \neq l$. Since $j \in \mathcal{J}_N^*$, j is matched in M_2, so j must be part of an exchange by donating a right lobe, i.e.: $t(j, M(j)) = r$. Since $\{j, M(j)\} \in E_K$ and $j \in \mathcal{I}_V \setminus \widetilde{\mathcal{J}}_K$, this contradicts the definition of E_K. ■

Proof of Part 3 in Lemma 9: Take any $i \in \mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$. Since $i \in \mathcal{I}_0 \setminus \mathcal{J}_K$, there exists $n \in \{1, \ldots, N\}$ such that $i = i_n^*$. Since $i = i_n^* \notin \mathcal{J}_N^*$, at Step 2.$n$ of the algorithm, $\mathcal{J}_K \cup \mathcal{J}_{n-1}^* \cup \{i_n^*\}$ was not matchable in G_K. Since $\mathcal{J}_{n-1}^* \subseteq \mathcal{J}_N^*$, this also implies that $\mathcal{J}_K \cup \mathcal{J}_N^* \cup \{i_n^*\}$ is not matchable in G_K. Therefore, M_2 must leave $i = i_n^*$ unmatched. ■

D.3.2 Proof of Theorem 3

Given a priority order over all pairs, a topological order of the precedence digraph of Category V pairs, and a willingness profile, the Pareto indifference of all pairs in all possible outcomes of $f^{l\&r}$ follows from Lemma 9. We will prove the other three properties of $f^{l\&r}$ in three separate lemmas.

Lemma 11 *The general left&right-lobe priority mechanism is individually rational.*

Proof: Take any willingness profile $R \in \mathcal{R}$ and let $M = f^{l\&r}(R)$. By Lemma 2, the individual rationality condition is satisfied for all pairs whose match is determined in Step 0 of the algorithm, i.e., Category 0, I, and II pairs and unwilling Category IV pairs.

Next, take any Category III, willing Category IV, or Category V pair i such that $M(i) \neq \emptyset$. Note that i must be matched in Step 2 of the algorithm, so $M(i) R_i \emptyset$ by parts 1, 2(a), 2(b), and 2(c) of Lemma 9. This is enough to conclude that the individual rationality condition is satisfied for all Category III, IV, and V pairs, since they cannot directly donate.

Finally, consider any Category VI pair i. Note that i is either matched in Step 2 by being part of an exchange donating a left lobe (by part 2(d) of Lemma 9), or is willing and directly donates a right lobe in Step 3, or is unwilling and left unmatched in Step 3. Therefore, the individual rationality condition is also satisfied for i. ■

Lemma 12 *The general left&right-lobe priority mechanism is Pareto efficient.*

Proof: Take any willingness profile $R \in \mathcal{R}$ and let $M = f^{l\&r}(R)$. Let $\mathcal{J}_w$ be the set of pairs i such that $R_i = R_i^w$, i.e., the set of willing pairs. Let M' be a matching in $G_c[\mathcal{J}_w]$ such that $M' R_i M$ for all $i \in \mathcal{I}$. To conclude that $f^{l\&r}$ is Pareto efficient, it is enough to show that all pairs are indifferent between M and M'. Below, we will prove this indifference separately for pairs in $\mathcal{I} \setminus \mathcal{I}_0$, $\mathcal{J}_K$, $\mathcal{J}_N^*$, and $\mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$. Note first that since all pairs weakly prefer M' to M, and M is individually rational at R (by Lemma 11), the matching M' is also individually rational at R.

Claim 1. Pairs in $\mathcal{I} \setminus \mathcal{I}_0$ are indifferent between M and M'.

Proof: By individual rationality of M' and Lemma 2, M' matches Category 0, I, and II pairs and unwilling Category IV pairs in exactly the same way as M, as determined in Step 0 of the algorithm. □

Claim 2. Pairs in $\mathcal{J}_K$ are indifferent between M and M'.

Proof: Take any pair $i \in \mathcal{J}_K$. By part 1 of Lemma 9, in M, i takes part in an exchange donating a left lobe. Since $M' R_i M$, and direct donation is not possible for Category V pairs, i must also take part in an exchange donating a left lobe in M'. □

Let M'' be the edges in M' corresponding to the *exchanges* among pairs in $\mathcal{I}_0$, i.e.:

$$M'' = \Big\{ \{i, j\} \in M' : i, j \in \mathcal{I}_0 \text{ and } i \neq j \Big\}.$$

Since Category VI pairs are the only pairs in $\mathcal{I}_0$ that can directly donate, we have:

$$M''(i) = \begin{cases} \emptyset & \text{if } i \text{ is Category VI and } M'(i) = i, \\ M'(i) & \text{otherwise} \end{cases}$$

for all $i \in \mathcal{I}_0$. Note that since $M'' \subseteq M'$, M'' is also a matching in $G_c[\mathcal{J}_w]$.

Claim 3. M'' is a matching in G_K.

Proof: To see that $M'' \subseteq E_K$, take any $\{i, j\} \in M''$. Then, $i \neq j$ by definition of M''; and $\{i, j\} \in E_c[\mathcal{J}_w]$, since M'' is a matching in $G_c[\mathcal{J}_w]$. To show that $\{i, j\} \in E_K$, by Lemma 9 we need to show that for the two cases such that $\{k, \ell\} = \{i, j\}$, if $t(k, \ell) = r$ then either k is Category IV or $h \in \widetilde{\mathcal{J}}_K$. Suppose that $t(i, j) = r$, i.e., in M'', i is part of an exchange by donating a right lobe. Then, $i \notin \mathcal{J}_K$ because, as argued in Claim 2, the pairs in $\mathcal{J}_K$ take part in exchanges donating a left lobe in M', so also in M''. Also $i \notin \mathcal{I}_V \setminus (\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K)$, because by Step 1 of the algorithm any pair in $\mathcal{I}_V \setminus (\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K)$ is unwilling, so individual rationality of M' implies that in M', hence in M'', it cannot be part of an exchange donating a right lobe. Finally, note that i is not Category VI since by individual rationality of M' and Lemma 2, in M', hence in M'', Category VI pairs cannot be part of an exchange donating a right lobe. Therefore, i is a (willing) Category IV pair or belongs to $\widetilde{\mathcal{J}}_K$. We proved that

$$t(i,j) = r \;\Rightarrow\; i \text{ is Category IV or } i \in \widetilde{\mathcal{J}}_K.$$

The proof of the other implication:

$$t(j,i) = r \;\Rightarrow\; j \text{ is Category IV or } j \in \widetilde{\mathcal{J}}_K$$

is exactly symmetric, hence omitted. So $\{i,j\} \in E_K$. □

Claim 4. Pairs in $\mathcal{J}_N^*$ are indifferent between M and M'.

Proof: We will prove the Claim separately for Category III, IV, V, and VI pairs:

- *Claim 4.III*: Take any Category III pair $j \in \mathcal{J}_N^*$. By part 2 of Lemma 9, j is taking part in an exchange donating a left lobe in M. By individual rationality of M' and Lemma 2, j is unmatched or takes part in an exchange donating a left lobe in M'. Therefore, $M' R_j M$ implies that j is part of an exchange donating a left lobe also in M'.
- *Claim 4.IV*: Take any Category IV pair $j \in \mathcal{J}_N^*$. By part 2 of Lemma 9, j is a willing Category IV pair taking part in an exchange donating a right lobe in M. By individual rationality of M' and Lemma 2, j is unmatched or takes part in an exchange donating a right lobe in M'. Therefore, $M' R_j M$ implies that j is part of an exchange donating a right lobe also in M'.
- *Claim 4.V*: Take any Category V pair $j \in \mathcal{J}_N^*$. By part 2 of Lemma 9, $j \in \widetilde{\mathcal{J}}_K \cap \mathcal{J}_N^*$ and j is part of an exchange donating a right lobe in M. As argued in Claim 2, the pairs in $\mathcal{J}_K$ take part in exchanges donating a left lobe in M', hence also in M''. By Claim 3, M'' is a matching in G_K, so by Lemma 10 and $j \in \mathcal{I}_V \setminus \mathcal{J}_K$, j is not part of an exchange donating a left lobe in M'', hence also in M'. Being Category V, j cannot directly donate, so $M' R_j M$ implies that j is part of an exchange donating a right lobe also in M'.
- *Claim 4.VI*: Take any Category VI pair $j \in \mathcal{J}_N^*$. By part 2 of Lemma 9, j is part of an exchange donating a left lobe in M. Being Category VI, j cannot directly donate a left lobe, so $M' R_j M$ implies that j is part of an exchange donating a left lobe also in M'. □

The argument of Claim 4 also establishes that M'' is a matching in G_K such that $M''(i) \neq \emptyset$ for all $i \in \mathcal{J}_K \cup \mathcal{J}_N^*$. So by Lemma 9, $M''(i) = \emptyset$ for all $i \in \mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$.

Claim 5. Pairs in $\mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$ are indifferent between M and M'.

Proof: Take any $i \in \mathcal{I}_0 \setminus (\mathcal{J}_K \cup \mathcal{J}_N^*)$. If i is not Category VI, then i cannot directly donate so $M''(i) = \emptyset$ implies $M'(i) = \emptyset$. Note that i is also unmatched in M, since i is not matched in Step 2 (by Lemma 9) nor Step 3 of the algorithm (since i is not Category VI). Therefore, such pairs are indifferent between M and M'. Next suppose that i is Category VI. Then, if i is unwilling, $M''(i) = \emptyset$ and individual rationality of M' imply that $M'(i) = \emptyset$. Note that i is also unmatched in M, since i is not matched in Step 2 (by Lemma 9) or Step 3 of the algorithm (since i is not willing). Therefore, such pairs are also indifferent between M and M'. Finally, suppose that i is a willing Category VI pair. Then, $M''(i) = \emptyset$ and individual rationality of M' imply that $M'(i) = i$, i.e., i directly donates a right lobe. Note that i also directly donates a right lobe in M, since i is not matched in Step 2 (by Lemma 9), but is matched in Step 3 of the algorithm (since i is willing). Therefore, such pairs are also indifferent between M and M'. □ ■

Lemma 13 *The general left&right-lobe priority mechanism is incentive compatible.*

Proof: Note that all exchanges are determined at Step 2 of the algorithm. By Lemma 9, only Category III, V, and VI pairs may be part of an exchange by donating a left lobe. We will show that for any Category III, V, or VI pair i and $R_{-i} \in \prod_{j \neq i} \mathcal{R}_j$ (which we will refer to Condition (**) below in the proof):

$$\text{Under } f(R_i^w, R_{-i}),\ i \text{ participates in an exchange by donating a left lobe}$$

$$\Updownarrow$$

$$\text{Under } f(R_i^u, R_{-i}),\ i \text{ participates in an exchange by donating a left lobe,}$$

which will prove that $f^{l\&r}$ is incentive compatible by Lemmas 3 and 11. In the rest of the proof, denote the objects defined by the algorithm under the willingness profile (R_i^w, R_{-i}) by using the superscript w, and those defined by the algorithm under the willingness profile (R_i^u, R_{-i}) by using the superscript u.

First consider the case where i is Category III. Pair i is never transformed in the algorithm regardless of its willingness announcement. Thus, the algorithm runs exactly the same manner regardless of its willingness announcement, implying $\mathcal{J}_K^u \cup \mathcal{J}_N^{*u} = \mathcal{J}_K^w \cup \mathcal{J}_N^{*w}$, and thus, Condition (**) above trivially holds for i.

Next consider the case where i is Category V. Then $i = i_k$ for some $k \in \{1, \ldots, K\}$. Up to the end of Step 1.$(k-1)$, both versions of the algorithm run in exactly the same way, since they do not depend on the willingness announcement of $i = i_k$. This implies that $\mathcal{J}_{k-1}^w = \mathcal{J}_{k-1}^u$ and $G_{k-1}^w = G_{k-1}^u$. Then,

$$t(i, M_2^w(i)) = l \;\Leftrightarrow\; i \in \mathcal{J}_K^w \;\Leftrightarrow\; \mathcal{J}_{k-1}^w \cup \{i_k\} \text{ is matchable in } G_{k-1}^w$$

$$\Leftrightarrow\; \mathcal{J}_{k-1}^u \cup \{i_k\} \text{ is matchable in } G_{k-1}^u \;\Leftrightarrow\; i \in \mathcal{J}_K^u \;\Leftrightarrow\; t(i, M_2^u(i)) = l$$

where the first and last equivalences above follow from Lemma 9 and i being Category V; the second and fourth equivalences follow from the definition of Step 1.k of the algorithm and $i = i_k$; and the middle equivalence follows from $\mathcal{J}^w_{k-1} = \mathcal{J}^u_{k-1}$ and $G^w_{k-1} = G^u_{k-1}$.

Finally consider the case where i is Category VI. Since the algorithm is independent of the willingness announcenment of Category VI pairs until the end of Step 2, we have $\mathcal{J}^{*w}_N = \mathcal{J}^{*u}_N$. Then,

$$t(i, M^w_2(i)) = l \iff i \in \mathcal{J}^{*w}_N \iff i \in \mathcal{J}^{*u}_N \iff t(i, M^u_2(i)) = l$$

where the the first and last equivalences above follow from Lemma 9 and i being Category VI; and the middle equivalence follows from $\mathcal{J}^{*w}_N = \mathcal{J}^{*u}_N$. ■