Efficient and Incentive-Compatible Liver Exchange*

Haluk Ergin† Tayfun Sönmez‡ M. Utku Ünver§

May, 2019

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, and introduce an individually rational, Pareto-efficient, and incentive-compatible mechanism. Construction of this mechanism requires novel 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

1 Introduction

Following the kidney, the liver is the second most common organ for transplantation worldwide. In 2018 there were 12,720 new additions in the US to waitlists for liver transplants.

*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 Co-Editor Dirk Bergemann and five anonymous referees of the journal, the participants at MATCH-UP in Ascona - Switzerland, Stanford, Conference in Honor of John Weymark at Vanderbilt, Princeton, INFORMS in Phoenix, Shanghai UFE Theory Conference, UT Sydney Market Design Conference, SCW in Seoul, SAET at Academia Sinica - Taiwan, Melbourne, Deakin, Adelaide, Vanderbilt Market Design Conference, NBER Market Design Working Group Meeting, Matching in Practice Workshop at Budapest, Harvard/MIT, SAET in Faro - Portugal, Conference on Economic Design at University of York - UK, Southern Methodist, Penn State, 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; sonmezt@bc.edu

§Department of Economics, Boston College; unver@bc.edu

While 8,250 patients were removed from waitlists due to receiving liver transplants, 1,159 of them were removed due to death, and 1,315 were removed due to being too sick for a transplant. 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 to 2001. The highly publicized death of a right-lobe liver donor in the US in 2002

[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 2018, while only 401 of 8,250 liver transplants were from living donors in the US, 1,106 of 1,475 liver transplants in South Korea and 1,150 of 1,588 liver transplants in Turkey were from living donors.

[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.

brought an end to this remarkable increase, and resulted in a 40–50% reduction from its peak over the next decade.[3] The number of annual living-donor liver transplants in the US have been mostly increasing again since 2012, with 401 transplants in 2018.

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 incompatible pairs, but also a subset of compatible pairs, such as those whose only direct-donation possibility to their patients is through a much higher-donor-risk right hepatectomy. Under an efficient and incentive-compatible mechanism we introduce, compatible pairs participate in exchange only if they strictly benefit by doing so, most notably 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

[3]See Grady (2002).

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

of offering a less risky procedure to such pairs provides an opportunity to increase the size 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 a small, blood-type O donor would likely be motivated to participate in exchange, if the pair benefits from exchange by reducing the donor risk through a much safer procedure of left hepatectomy. However, not all cases are this straightforward. Consider a blood-type A patient with a blood-type B donor. Since this pair 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 adopt the mechanisms and techniques developed for kidney exchange directly to liver exchange, unless the higher-donor-risk right hepatectomy is completely ruled out. 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 either revealing its willingness to go through the less desired right hepatectomy or by revealing whether it has a direct-transplant bias or not.

The key pairs in the design of an efficient and incentive-compatible mechanism are those who can participate in exchange both through a left-lobe donation as well as through a less-preferred right-lobe donation. The challenge is determining when the donors of a particular pair shall be considered for a right-lobe donation rather than a left-lobe donation. We refer to this process as a *transformation*. To assure incentive compatibility, a pair 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 difficulties 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 results in a Pareto-inferior outcome. Second, this possibility might encourage pair 1 to hide its willingness for a right-lobe donation. Hence, key in our design is determining the order in which pairs are to 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. We also illustrate the potential gains from adopting our proposed mechanism on simulated pools based on

South Korean population and transplantation characteristics. We show that, through liver exchange, the number of living-donor liver transplants can be increased more than 30%.

1.1 Double Equipois & Vancouver Forum

From an ethical perspective, the concept of *double equipois* was proposed by Cronin et al. (2001) to balance the risk of a healthy donor versus the benefit for a high-risk recipient. A well-designed liver exchange system can not only be an effective tool to achieve this balance, but it also complies with the mainstream approach towards living-donor liver transplantation, summarized in the *Report of the Vancouver Forum* (Barr et al., 2006). Two of the main principles of live liver donation, both related to the concept of double equipois, are stated as follows in this reference document:

> Live liver donation should only be performed if the risk to the donor is justified by the expectation of an acceptable outcome in the recipient ...
>
> The estimated risk of mortality and morbidity currently associated with live donor right hepatectomy is 0.4% and 35% respectively. Since the risk to the donor is considerable, programs performing live donor liver transplantation should institute procedures and protocols that insure that donor mortality and morbidity is minimized.

Hence, a possible establishment of a liver exchange program is very much in the spirit of the principles outlined by the Vancouver Forum, especially if it gives priority to left-lobe donation.

In part motivated by the theory of double equipois, there has been some renewed interest in the liver transplantation community in finding ways to replace higher-donor-risk right hepatectomy with left hepatectomy. Suggesting that

1. donor complications are not only 4- to 12-fold lower for left-lobe donors than right-lobe donors, but also
2. complications are less severe under left-lobe donation than under right-lobe donation,

Roll et al. (2013) propose shifting the risk from the donor to the patient by lowering the minimum acceptable liver tissue volume to a less conservative level. They state:

> Although using smaller grafts from LL [left lobe] may decrease recipient benefit absolutely, their double-equipoise analysis suggests that LL [left lobe] is more efficient than RL [right lobe] in converting donor risk into recipient benefit.

Establishment of a liver exchange program and adoption of our mechanism can be seen as part of these efforts to reduce donor-risk through increased use of left hepatectomy, although, in contrast to Roll et al. (2013) proposal where the increased utilization of left hepatectomy comes at the expense of an increased average risk to patients, under our proposed liver exchange mechanism it is achieved more naturally without any adverse effect.

Finally, there is one additional benefit of liver exchange, reported in Pomfret et al. (2011). There can be situations where a direct liver transplant from a donor to his patient is ethically unacceptable based on the theory of double-equipoise, for example due to old age of the patient that translates to low patient benefit, but an exchange involving the same patient-donor pair may be ethically acceptable due to the additional benefit to the other patient.

1.2 Other 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, 2019), 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 either his *left liver lobe* or his *right liver lobe* for a transplant, although the latter involves considerably higher risk to the donor. We sometimes refer a liver lobe as a *graft*, when it is donated for a transplant.

2.1 Size Compatibility

The volume of the left liver is generally between 30% to 40% of the liver volume, and the right lobe makes up the rest. Formally, the **size of a liver lobe** (or a graft) is the volume of the liver lobe.

A patient typically requires a liver graft with a volume of at least 40% of her own liver volume, although for some patients this minimum requirement may differ depending on the details of the patient's disease. Formally, the **size of a patient** is the minimum required volume of the liver graft he needs for a transplant. Therefore, a liver lobe is **size compatible** with a patient if and only if it is as large as the size of the patient.

Let $\mathbf{S} = \{0, 1, \ldots, S-1\}$ 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]

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 (and each of his

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

liver lobes) are **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 $\mathbf{B} = \{O, A, B, AB\}$ denote the set of blood types.

2.3 Liver Donation Relation & Its Equivalent Representation

The blood type and the size of each patient are assumed to be observable physical attributes. Similarly, the blood type and the size of each liver lobe (i.e. graft) are also assumed to be observable. The set of patient types, and the set of liver lobe (or equivalently graft) types are both referred as $\mathbf{B} \times \mathbf{S}$. Observe that, the type of a liver lobe only specifies its blood type and size, and not whether it is the left lobe or the right lobe.

Consider a donor who wants to donate a liver lobe to a patient. He can do so if and only if he is blood-type compatible and the intended lobe is size compatible with the patient.

Let $\trianglerighteq$ denote the **liver donation partial order** on $\mathbf{B} \times \mathbf{S}$, where for any $J, J' \in \mathbf{B} \times \mathbf{S}$, $J \trianglerighteq J'$ means a liver lobe of type J is both blood-type compatible and also size-compatible with a patient of type J'. Define $\mathbf{T} \equiv \{0,1\}^2 \times \{0, 1, \ldots, S-1\}$. Note that, the two partially ordered sets $(\mathbf{B} \times \mathbf{S}, \trianglerighteq)$ and $(\mathbf{T}, \geq)$ are order isomorphic, where the order isomorphism associates each patient or graft type $J \in \mathbf{B} \times \mathbf{S}$ with the following vector $X \in \mathbf{T}$:[6]

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

Let type $J \in \mathbf{B} \times \mathbf{S}$ be associated with vector $(X_1, X_2, X_3) \in \mathbf{T}$, and type $J' \in \mathbf{B} \times \mathbf{S}$ be associated with vector $(X'_1, X'_2, X'_3) \in \mathbf{T}$. Due to the above described order isomorphism, we have

$$J \trianglerighteq J' \iff (X_1, X_2, X_3) \geq (X'_1, X'_2, X'_3)$$

That is, a graft of type J can be transplanted to a patient of type J' if and only if $(X_1, X_2, X_3) \geq (X'_1, X'_2, X'_3)$.

For notational convenience, we use the equivalent representation $(\mathbf{T}, \geq)$ to represent the liver donation partially ordered set. Therefore, from now on, $\mathbf{T}$ represents both the set of patient types, and the set of liver lobe types. The description of a donor, on the other hand,

[6] Relation $\geq$ in $(\mathbf{T}, \geq)$ is the usual partial order over integer vectors in Euclidean n-dimensional space: for all $a = (a_1, \ldots a_n)$, $b = (b_1, \ldots, b_n) \in \mathbb{Z}^n_+$ we say $a \geq b$ if $a_k \geq b_k$ for all k. Its asymmetric part is denoted as $>$: $a > b$ if $a_k \geq b_k$ for all k and $a_k > b_k$ for some k; its symmetric part is denoted as $=$: $a = b$ if $a_k = b_k$ for all k.

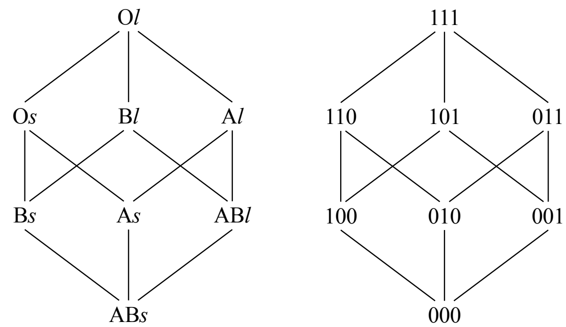

Figure 1: The Partially Ordered Sets $(\mathbf{B} \times \mathbf{S}, \trianglerighteq)$ and $(\{0,1\}^3, \geq)$ for $\mathbf{S} = \{s, l\}$ such that s refers to small size and l refers to large size.

includes the sizes of both his liver lobes, and hence the set of donor types is represented by

$$\mathbf{T^D} \equiv \{0,1\}^2 \times \{0,1,\ldots,S-1\}^2.$$

Given a donor of type $Y = (Y_1, Y_2, Y_{3\ell}, Y_{3r}) \in \mathbf{T^D}$,

- $Y^\ell \equiv (Y_1, Y_2, Y_{3\ell}) \in \mathbf{T}$ denotes the type of his left liver lobe, and
- $Y^r \equiv (Y_1, Y_2, Y_{3r}) \in \mathbf{T}$ denotes the type of his right liver lobe.

Given the human anatomy, the right liver lobe is always larger than the left liver lobe. Therefore, $Y_{3\ell} < Y_{3r}$ for any $Y \in \mathbf{T^D}$.[7] For any pair $X, Y \in \mathbf{T} \times \mathbf{T^D}$, a patient of type X can receive a transplant from a donor of type Y

- through a left-lobe donation (i.e. left hepatectomy) if $X \leq Y^\ell$, and
- through a right-lobe donation (i.e. right hepatectomy) if $X \leq Y^r$.

Observe that, these vector inequalities are equivalent to the donor being blood-type compatible with and (2) his liver lobe for transplantation being size compatible with 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 liver donation partial order $\trianglerighteq$ on $\mathbf{B} \times \mathbf{S}$, and the standard partial order $\geq$ over the corners of the three-dimensional cube $\{0,1\}^3$.

A patient of type $X \in \mathbf{T}$ and a donor of type $Y \in \mathbf{T^D}$ are **left-lobe compatible** if $X \leq Y^\ell$. 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

[7] The only exception is when $Y_{3\ell} = S - 1$, i.e., when left lobe of a donor of type Y is the largest size, we assume $Y_{3r} = Y_{3\ell}$ for notational convenience of defining one size set $\mathbf{S}$ for both patient needs and donor lobe sizes and terminological convenience of saying there are S sizes. For this donor type Y, right lobe is never donated as its left lobe is large enough for all patient types. This assumption is made for notational simplicity as well in order to avoid introducing two separate sets of sizes, one for the patients and other for the liver lobes.

who can feasibly donate their left lobes to a patient, we assume that right-lobe donation is not a viable option. A patient of type $X \in \mathbf{T}$ and a donor of type $Y \in \mathbf{T^D}$ are **right-lobe-only compatible** if the donor can donate his right lobe to the patient, but not his left lobe, i.e., $X \leq Y^r$ and $X \nleq Y^\ell$.

3 Liver Exchange

As in kidney exchange, the number of living-donor liver transplants can be increased through exchange of donors. Living-donor liver transplantation is a more complex medical procedure than living-donor kidney transplantation, in part because only a portion of the donor's liver is transplanted to the patient, and hence a detailed analysis of patient and donor anatomies is required. Hence the logistics of liver exchange gets more complicated as the number of pairs increase in an exchange. Indeed, in a recent paper proposing an organized liver exchange to the members of the transplantation community, Mishra et. al. (2018) suggest:

> As with the initial experience with KPE, it is anticipated that LPE would begin with 2-way swaps, the simplest form of exchange.

Consistent with their suggestion, we assume that only two-way exchanges are feasible.[8]

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**.[9] The observable characteristics of a pair are summarized by an ordered pair of individual types $X - Y \in \mathbf{T} \times \mathbf{T^D}$, where X denotes the type of the patient and Y denotes the type of the donor; $X - Y$ is called the **pair type**.[10]

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

1. $\mathcal{I} = \{1, 2, \dots, K\}$ is a finite set of patient-donor pairs, and
2. $\tau : \mathcal{I} \to \mathbf{T} \times \mathbf{T^D}$ is a function, such that, for every pair $i \in \mathcal{I}$, $\tau(i)$ is its pair type.

For every pair $i \in \mathcal{I}$, we denote its type as $\tau(i) = \tau_P(i) - \tau_D(i)$, where $\tau_P(i) \in \mathbf{T}$ is the type of the patient of the pair, and $\tau_D(i) \in \mathbf{T^D}$ is the type of the donor of the pair.

Moreover, given a pair $i \in \mathcal{I}$, let $\tau_D^\ell(i) \in \mathbf{T}$ denote the type of its donor's left lobe, and $\tau_D^r(i) \in \mathbf{T}$ denote the type of its donor's right lobe.

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

[8] All liver exchanges reported in the literature as of March 2018 are between two patients and their donors.
[9] We use pronouns "she" for a patient, "he" for a donor, and "it" for a pair.
[10] We refer to a pair type as $X - Y$ instead of (X, Y) as a convention.

3.2 Feasible Grafts & Assignments

Since we rule out the possibility of right-lobe donation when a donor can more safely donate his left lobe, there is a unique donation "mode" between any donor and any patient: For any two pairs $(i, j) \in \mathcal{I} \times \mathcal{I}$, the donor of pair j can either feasibly donate his left lobe, or his right lobe, or neither of them to the patient of pair i. The following function keeps track of which lobe is to be donated (if any) in any potential assignment. Define the **transplant type** function $t : \mathcal{I} \times \mathcal{I} \to \{\ell, r, \emptyset\}$ as follows: For any $(j, i) \in \mathcal{I} \times \mathcal{I}$,

$$t(j,i) = \begin{cases} \ell & \text{if } \tau_P(i) \leq \tau_D^{\ell}(j) \\ r & \text{if } \tau_P(i) \not\leq \tau_D^{\ell}(j) \ \&\ \tau_P(i) \leq \tau_D^{r}(j) \\ \emptyset & \text{otherwise} \end{cases}$$

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

For any pair $i \in \mathcal{I}$, define

$$\mathcal{C}(i) \equiv \Big\{ j \in \mathcal{I} \ : \ t(j,i) \neq \emptyset \Big\}$$

be the set of pairs from whom the patient of i can receive a transplant from. Since the transplant type function uniquely determines which lobe of a compatible donor shall be transplanted to a patient, set $\mathcal{C}(i)$ also uniquely defines the set of feasible grafts for the patient of pair i. That is, for the patient of pair i

- the left lobe of the donor of pair j is **feasible** if and only if $j \in \mathcal{C}(i)$ and $t(j, i) = \ell$, and
- the right lobe of the donor of pair j is **feasible** if and only if $j \in \mathcal{C}(i)$ and $t(j, i) = r$.

With a slight abuse of terminology, we will also refer to set $\mathcal{C}(i)$ as **the set of feasible grafts** for the patient of pair i.

3.3 Preferences

There are three types of possible outcomes for a pair i:

1. The patient of pair i receives her own donor's feasible graft in a **direct transplant**.
2. Pair i exchanges donors with another pair to form a **(two-way) exchange**, so that the patient of each pair receives the only feasible graft from the other pair's donor.
3. Pair i remains **unmatched** and its patient does not receive a transplant. This outcome is denoted as $\emptyset$.

In any potential exchange, patients of both pairs have to each receive a transplant. Hence, define the set of **(feasible) assignments** for pair i as

$$\mathcal{E}(i) \equiv \Big\{ j \in \mathcal{C}(i) \ : \ i \in \mathcal{C}(j) \Big\}.$$

Observe that $i \in \mathcal{E}(i)$, whenever pair i is a compatible pair.

Since the function $t(i,.)$ uniquely identifies which lobe the donor of pair i donates in any assignment in $\mathcal{E}(i)$, we can further partition the set $\mathcal{E}(i)$ based on the donated lobe:

$$\mathcal{E}^{\ell}(i) \equiv \Big\{ j \in \mathcal{C}(i) \ : \ i \in \mathcal{C}(j) \text{ and } t(i,j) = \ell \Big\}, \text{ and}$$
$$\mathcal{E}^{r}(i) \equiv \Big\{ j \in \mathcal{C}(i) \ : \ i \in \mathcal{C}(j) \text{ and } t(i,j) = r \Big\}.$$

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 of a pair depend on the following three factors:

1. Observable characteristics of the received compatible graft.
2. Whether the donor donates his left lobe or his right lobe.
3. Whether the patient receives a graft through direct transplant or through an exchange.

For any pair $i \in \mathcal{I}$, the preference relation R_i is defined over the set $\mathcal{E}(i) \cup \{\emptyset\}$; that is, $R_i \subseteq \big(\mathcal{E}(i) \cup \{\emptyset\}\big) \times \big(\mathcal{E}(i) \cup \{\emptyset\}\big)$. Let P_i denote the asymmetric part and I_i denote the symmetric part of R_i.

In order to motivate the restrictions we make on the preferences, we next discuss the role of each of these factors on liver transplantation.

3.3.1 Preferences on Observable Characteristics of a Graft

While blood-type compatibility and size compatibility are the primary considerations for liver transplantation, other factors, such as the age of the donor, can also influence the outcome to a lesser extent. As a result, other things being equal, a pair may prefer one compatible graft to another, even if they are of the same type. For a given pair i, let the weak order $\succsim_i$ represent the **received-graft preference relation** over the set of feasible grafts $\mathcal{C}(i)$, with its asymmetric part indicated by $\succ_i$ and symmetric part indicated by $\sim_i$.

Since it purely depends on observable donor characteristics, we assume that the received-graft preference relation $\succsim_i$ is public information.

Given the high risk to donors, their screening is very strict for living-donor liver transplantation, and donation is ruled out unless the donor is in perfect health and the benefit to the patient is sufficiently high. Grafts must be sufficiently large (usually 40% of the patient's liver volume) to minimize the risk of *small-for-size syndrome*, a condition where a patient develops liver dysfunction and ascites when the transplanted graft is too small. As a result, the expected benefit to the patient is "similar" between any two grafts deemed compatible for the patient. It is also possible that the central planner may choose to disclose information on patient-donor compatibility only, and may not make any additional information on pairs available to third parties. Indeed, this is often the case for kidney exchange, where patient-donor pairs do not meet until after the exchange has taken place, and they only meet if each person agrees. In this likely scenario, pairs will be assumed to be indifferent between all compatible grafts under the received-graft preference relation $\succsim_i$.

Based on these observations, any asymmetric part of the received-graft preference relation on compatible grafts will play a secondary role of a "tie-breaker" on the preferences of a pair.

3.3.2 Preferences on Left-Lobe Donation vs Right-Lobe donation

While the expected benefit to a patient is similar for all compatible grafts, donor mortality and morbidity risks differ considerably between left-lobe donation and right-lobe donation. Based on the 2006 Vancouver Forum report, mortality rate exceeds 0.4% for right-lobe donation in contrast to approximately 0.1% for left-lobe donation.[11] Morbidity rates to the donor, significant complications other than mortality, are also much higher under right-lobe donation. Mishra et al. (2018) report that the morbidity rates are 28% for right-lobe donation and 7.5% for left-lobe donation.

As a result, pairs have much stronger preferences for their donors to donate their right lobes than their left lobes. Indeed, pairs may not be willing to have their donors to donate their right lobes at all. We refer to such pairs as **unwilling ($\boldsymbol{u}$)**. Pairs that are open to the possibility of right-lobe donation from their donors, on the other hand, are referred to as **willing ($\boldsymbol{w}$)**.

Formally,

- for an unwilling pair i, for all $j \in \mathcal{E}^\ell(i)$ and $j' \in \mathcal{E}^r(i)$,

$$j \; P_i \; \emptyset \; P_i \; j', \qquad \text{and}$$

- for a willing pair i, for all $j \in \mathcal{E}(i)$,

$$j \; P_i \; \emptyset.$$

Whether a pair i is willing or unwilling is its private information.

Importantly, the willingness of a pair is assumed to be independent of the graft received by its patient and whether it is received through a direct transplant or an exchange. Therefore which lobe is donated by the pair is the primary consideration in their preferences. We show in Example 4 in Appendix C.1 that, in the absence of this assumption, a mechanism

[11] More recent reports give similar mortality rates.

that is Pareto efficient, individually rational, and incentive compatible may fail to exist. Moreover, given that the risk-benefit ratio, one of the key considerations in living-donor organ transplantation, is highly responsive to the donated lobe but relatively irresponsive to the received graft, we believe this is a fairly realistic assumption. Indeed, this assumption is also supported by the findings of Molinari et al. (2014), where the authors analyze living liver donors' risk thresholds using decision analysis techniques based on the probability trade-off method. Individuals who were responsible for liver patients' daily care or who were emotionally or biologically related to the potential recipients were identified and screened for this study. The most important factor that influence participants' decisions was identified as the patient's expected life gain from living donor liver transplantation. The authors report, 88% of the participants would donate for a gain of 1 year, 95% would donate for a gain of 3 years, and 98% would donate for a gain of 5 years. Therefore, given the 3-year survival rate of 83% and the 5-year survival rate of 78% reported by Goldberg et al. (2014), in the absence of other factors more than 95% of the donors can be expected to donate regardless of which compatible liver lobe their loved one receives. Indeed, of the eight factors that influence the decision to donate, six of them are about the risks and inconveniences faced by the donor. Following the patient's expected life gain that emerges as the leading factor influencing the decision to donate, the remaining factors, ranked in order of importance, are reported in Molinari et al. (2014) as follows:

1. Donor's mortality risk
2. Donor's risk for decreased physical capacity
3. Donor's morbidity risk
4. Causes of recipient's liver disease
5. Donor's financial burden
6. Donor's hospital stay
7. Donor's time off work

Remarkably, more than 90% of donors are willing to donate even when the mortality risk is as high as 4%. Therefore the difference between the donor mortality rates of left-lobe donation vs. right-lobe donation, while very significant, may not have much of an influence in donor willingness to donate a right lobe. However, while more than 95% of the donors are willing to accept a morbidity rate of 10%, less than 70% of them are willing to accept a morbidity rate of 30%. This not only shows that pairs may have very strong preferences between left-lobe donation and right-lobe donation, but also suggests that a considerable rate of donors may not be willing to donate their right lobes.

3.3.3 Direct Donation Bias

In real-life applications of kidney exchange, it is well established that most pairs have *direct-transplant bias*, which means a compatible pair opts for a direct transplant even when its patient is committed to receive a more-favorable kidney through exchange. This preference can have a time-preference component: exchange option involves more waiting than direct transplant, which can be realized without finding a suitable exchange partner and expectation of a more-favorible match than a direct transplant may not outweigh this waiting cost. Often it also has an emotional component: direct donation from a loved one may have induce a higher utility for the pair than a donation from a third-party.

In an organized liver exchange, we expect the direct donation bias to be also prevalent, unless the risk to the donor can be reduced through a much safer left-lobe donation. While we do not expect many compatible pairs to participate in exchange in the absence of this tangible benefit, we allow for it in our model. We will assume, however that, a possible direct-transplant bias never dominates the safety concerns for the donor, and as such a pair always prefers an exchange with a left-lobe donation to a direct-transplant of the right lobe. We also rule out the possibility of a "mild" direct-transplant bias in the sense that, subject to donating the same lobe, a pair that has a direct-transplant bias strictly prefers direct-transplant to any other graft through exchange. We show in Example 5 in Appendix C.1 that, in the absence of this assumption, a mechanism that is Pareto efficient, individually rational, and incentive compatible may fail to exist.

To capture this private-information component of the preference relation, we introduce two *participation types* for left-lobe or right-lobe-only compatible pairs.

A left-lobe compatible or a right-lobe-only compatible pair i is an **exchange type** ($\boldsymbol{e}$) if,

$$\begin{array}{rc} \text{for all } j, j' \in \mathcal{E}^{\ell}(i) & j \; R_i \; j' \iff j \; \succsim_i \; j' \\ \text{for all } j, j' \in \mathcal{E}^{r}(i) & j \; R_i \; j' \iff j \; \succsim_i \; j' \\ \text{for all } j \in \mathcal{E}^{\ell}(i) \text{ and } j' \in \mathcal{E}^{r}(i) & j \; P_i \; j' \end{array}$$

A right-lobe-only compatible pair is a **direct-transplant type** ($\boldsymbol{d}$) if,

$$\begin{array}{rc} \text{for all } j \in \mathcal{E}^{\ell}(i) & j \; P_i \; i \\ \text{for all } j \in \mathcal{E}^{r}(i) & i \; P_i \; j \end{array}$$

A left-lobe compatible pair i is a **direct-transplant type** ($\boldsymbol{d}$) if,

$$\text{for all } j \in \mathcal{E}(i) \qquad i \; P_i \; j$$

Observe that, a left-lobe compatible pair of direct-transplant type has no reason to participate in exchange.

3.3.4 Preference Domain

For each pair $i \in \mathcal{I}$, fix the public information received-graft preference relation $\succsim_i$. Given the restrictions described in Sections 3.3.1-3.3.3, each pair has one of the four possible preference relations $R_i^{e/w}$, $R_i^{e/u}$, $R_i^{d/w}$, $R_i^{d/u}$, depending on whether it is willing or unwilling, and whether it is direct-transplant type or exchange type:

- For a willing exchange-type pair i , the possible outcomes are ranked as follows under its **exchange/willing preference relation** $R_i^{e/w}$:
    1. Feasible direct transplant and exchanges in which the pair donates a left lobe are ranked in order of its received-graft preferences $\succsim_i$.
    2. Feasible direct transplant and exchanges in which the pair donates a right lobe are ranked in order of its received-graft preferences $\succsim_i$.
    3. The pair is unmatched.
- For an unwilling exchange-type pair i, the possible outcomes are ranked in the following order under its **exchange/unwilling preference relation** $R_i^{e/u}$**:**
    1. Feasible direct transplant and exchanges in which the pair donates a left lobe are ranked in order of its received-graft preferences $\succsim_i$.
    2. The pair is unmatched.
    3. Feasible direct transplant and exchanges in which the pair donates a right lobe are ranked in order of its received-graft preferences $\succsim_i$.
- For a willing direct-transplant-type pair i, the possible outcomes are ranked in the following order under its **direct/willing preference relation** $R_i^{d/w}$**:**
    1. Feasible left-lobe direct transplant if i is left-lobe compatible.
    2. Feasible exchanges in which the pair donates a left lobe are ranked in order of its received-graft preferences $\succsim_i$.
    3. Feasible right-lobe direct transplant if i is right-lobe-only compatible.
    4. Feasible exchanges in which the pair donates a right lobe are ranked in order of its received-graft preferences $\succsim_i$.
    5. The pair is unmatched.
- For an unwilling direct-transplant-type pair i, the possible outcomes are ranked in the following order under its **direct/unwilling preference relation** $R_i^{d/u}$**:**
    1. Feasible left-lobe direct transplant if i is left-lobe compatible.
    2. Feasible exchanges in which the pair donates a left lobe are ranked in order of its received-graft preferences $\succsim_i$.
    3. The pair is unmatched.
    4. Feasible right-lobe direct transplant if i is right-lobe-only compatible.
    5. Feasible exchanges in which the pair donates a right lobe are ranked in order of its received-graft preferences $\succsim_i$.

Let $\mathbf{R}_i = \{R_i^{e/w}, R_i^{e/u}, R_i^{d/w}, R_i^{d/u}\}$ denote the set of possible preference relations for pair i.

An assignment $j \in \mathcal{E}(i)$ is **individually rational** for a pair i under a preference relation R_i if

$$\begin{array}{lll} j \; R_i \; \emptyset & \text{whenever} & i \notin \mathcal{E}(i), \text{ and} \\ j \; R_i \; \emptyset \text{ and } j \; R_i \; i & \text{whenever} & i \in \mathcal{E}(i). \end{array}$$

That is, an assignment is individually rational if it is not only weakly preferred to remaining unmatched, but also to a direct transplant whenever the pair is a compatible. Individual rationality is important because, no donor can be enforced for a donation he does not wish to. At any time during the donation process a living donor may change his or her mind. Throughout our analysis, we focus on individually rational outcomes.

Observe that

- if pair i is *incompatible*, i.e., $i \notin \mathcal{E}(i)$, then the two willing preferences coincide, i.e., $R_i^{e/w} = R_i^{d/w}$, and the two unwilling preferences also coincide, i.e., $R_i^{e/u} = R_i^{d/u}$; and
- if pair i is *left-lobe compatible*, i.e., $i \in \mathcal{E}^\ell(i)$, then the individually rational assignments are ranked the same way under the two exchange-type preferences, $R_i^{e/w}$ and $R_i^{e/u}$, and also the same way under the two direct-transplant-type preferences, $R_i^{d/w}$ and $R_i^{d/u}$.

Let $\mathbf{R} = \mathbf{R}_1 \times \ldots \times \mathbf{R}_K$ denote the set of preference profiles.

3.4 Outcome of the Problem: A Matching

We are ready to define the outcome of a liver-exchange problem. The set of all mutually compatible **matches** E_c is given as follows: For all $i, j \in \mathcal{I}$,[12]

$$\{i,j\} \in E_c \iff i \in \mathcal{E}(j) \iff j \in \mathcal{E}(i).$$

A match $\{i,j\} \in E_c$ is a (two-way) exchange if $i \neq j$ and a direct transplant if $i = j$.

The **compatibility graph** is defined as the undirected graph, with the pairs as its nodes and the mutually compatible matches as its edges, $G_c = (\mathcal{I}, E_c)$.[13]

Given a compatibility graph $G_c = (\mathcal{I}, E_c)$, a **matching** $M \subseteq E_c$ is a collection of compatible matches such that

$$\text{for all } \varepsilon, \varepsilon' \in M, \qquad \varepsilon \cap \varepsilon' \neq \emptyset \implies \varepsilon = \varepsilon'.$$

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

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

That is, no pair participates in two distinct exchanges or both in a direct transplant and in an exchange. Let $\mathbf{M}_c$ be the set of matchings supported by the compatibility graph G_c.

We denote the assignment of pair $i \in \mathcal{I}$ in matching $M \in \mathbf{M}_c$ 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 an exchange. If $M(i) = \emptyset$ (i.e., there is no $\varepsilon \in M$ such that $i \in \varepsilon$), then pair i remains unmatched.

Consider a match $\{i, j\}$ (possibly with $i = j$) in a matching M. Since a donor only donates a right lobe when his left lobe is too small for the intended receiver, 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.

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 3.3 can be directly extended to the set of matchings. We slightly abuse the notation, and let R_i also denote the resulting preference relation over all matchings $\mathbf{M}_c$ defined as:

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

3.5 Mechanisms and Axioms

Although the types of the participating pairs are observable, their preferences (or equivalently their willingness for a right-lobe donation and whether they have direct-transplant bias) are not. A (direct) mechanism determines a matching as a function of the reported preference profile.

Since we fix an exchange pool $(\mathcal{I}, \tau)$ and a received-graft preference profile $\succsim = (\succsim_i)_{i \in \mathcal{I}}$ throughout, we define a **mechanism** as a function $f : \mathbf{R} \to \mathbf{M}_c$.

A matching $M \in \mathbf{M}_c$ is **individually rational (IR)** at a preference profile $R \in \mathbf{R}$, if for every pair $i \in \mathcal{I}$, either $M(i)$ is an individually rational assignment at R or $M(i) = \emptyset$. A mechanism f is **individually rational (IR)** if $f(R)$ is individually rational at R for any $R \in \mathbf{R}$.

A matching $M \in \mathbf{M}_c$ is **Pareto efficient (PE) at** a preference profile $R \in \mathbf{R}$ if there does not exist a matching $M' \in \mathbf{M}_c$ 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 $R \in \mathbf{R}$.

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

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

4 An Efficient & Incentive-Compatible Mechanism

In this section we introduce a mechanism that is individually rational, Pareto-efficient, and incentive-compatible. In order to do so, we will make a number of key observations about the structure of the problem, and develop the tools required to equip a priority matching mechanism with the "safeguards" to avoid two potential complications due to availability of dual transplantation technologies for liver exchange.

4.1 Challenges to Overcome

While the starting point of our design is a priority mechanism, it has to be considerably modified to maintain Pareto efficiency and incentive compatibility. Under the most basic form of a *priority mechanism*, agents are committed (to the extent it is feasible) to be matched with one of their best possible matches –one at a time– following a fixed priority ordering. When there is a single transplantation technology, as in kidney exchange, this simple mechanism is both Pareto efficient and incentive compatible (Roth, Sönmez, and Ünver (2005)). As we present in our next example, the priority mechanism in this basic form is no longer incentive compatible in our problem due to the presence of dual transplantation technologies.

Example 1 *There is a set of three incompatible pairs $\mathcal{I} = \{i_1, i_2, i_3\}$ with the following types:*

$$\begin{aligned} \tau_P(i_1) &= (0,1,1) & \tau_D(i_1) &= (1,0,1,2) \\ \tau_P(i_2) &= (1,0,1) & \tau_D(i_2) &= (0,1,0,1) \\ \tau_P(i_3) &= (0,1,0) & \tau_D(i_3) &= (1,0,0,1) \end{aligned}$$

This pool of pairs result in the following set of feasible matches

$$E_c = \big\{\{i_1, i_2\}, \{i_2, i_3\}\big\},$$

and the following transplant type function

$$t(i_1,i_2) = \ell, \;\; t(i_1,i_3) = \emptyset, \;\; t(i_2,i_1) = r, \;\; t(i_2,i_3) = \ell, \;\; t(i_3,i_1) = \emptyset, \;\; t(i_3,i_2) = r.$$

Suppose all pairs are willing and none of them have a direct-transplant bias. That is, each pair $i \in \mathcal{I}$ has the preference relation $R_i^{e/w}$. The compatibility graph of this problem is depicted *in Figure 2.*

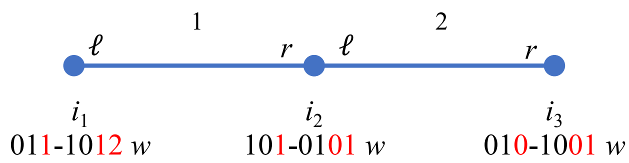

Figure 2: The compatibility graph for Example 1. All pairs are willing. The left- and right-lobe donations are denoted by letters ℓ and r, respectively. There are two individually rational exchanges.

Let us find the outcome of the priority mechanism for the priority order

$$\Pi = i_1 - i_2 - i_3.$$

First, we process pair i_1. Since exchange $\{i_1, i_2\}$ is the only feasible match for pair i_1, the priority mechanism commits for the exchange $\{i_1, i_2\}$. Next, we process pair i_2. Pair i_2 is already committed for a specific match, i.e. with pair i_1. Finally, we process pair i_3. There is no exchange that can match i_3 in addition to i_1 and i_2. So the outcome of the priority mechanism is the matching

$$M = \big\{\{i_1, i_2\}\big\},$$

where the donor of pair i_1 donates a left lobe, whereas the donor of pair i_2 donates a right lobe.

On the other hand, if pair i_2 declares itself as unwilling reporting its preferences as $R_i^{e/u}$, the exchange $\{i_1, i_2\}$ ceases to be individually rational, and thus pair i_2 enforces the priority mechanism to pick the unique individually rational exchange resulting with the matching

$$M' = \big\{\{i_2, i_3\}\big\}.$$

The donor of pair i_2 donates a left lobe under this alternative exchange, and thus the pair benefits from this manipulation.

Given the priority order Π, the priority mechanism enforces pair i_2 in Example 1 for a right-lobe donation, even though this pair is in a position to enforce an outcome where not only its patient receives a transplant, but also its donor donates the much safer left lobe. This example motivates our first departure from a basic priority mechanism: Under the modified mechanism, while we order pairs in a priority list and sequentially commit to matching them –one at a time– whenever it is feasible, we consider pairs for right-lobe donation only after their left-lobe donation possibilities are exhausted. Importantly, a pair's left-lobe donation possibilities include those exchanges where other pairs donate right lobes.

Under our modification, right-lobe donation from a pair is to be considered only when these possibilities are exhausted as well. It is easy to see that, this modification restores incentive compatibility of the mechanism. However, restoring incentive compatibility with this simple fix may come at the expense of losing Pareto efficiency. We illustrate this possibility with the following example:

Example 2 *There is a set of three incompatible pairs $\mathcal{I} = \{i_1, i_2, i_3\}$ with the following set of feasible matches*

$$E_c = \big\{\{i_1, i_2\}, \{i_2, i_3\}, \{i_3, i_1\}\big\},$$

and the transplant type function

$$t(i_1, i_2) = r,\ t(i_1, i_3) = \ell,\ t(i_2, i_1) = \ell,\ t(i_2, i_3) = r,\ t(i_3, i_1) = r,\ t(i_3, i_2) = \ell.$$

Suppose all pairs are willing and none of them have a direct-transplant bias. That is, each pair $i \in \mathcal{I}$ has the preference relation $R_i^{e/w}$. The compatibility graph of this problem is depicted in Figure 3.

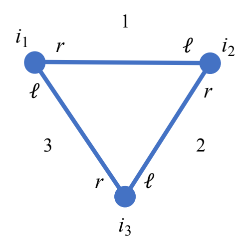

Figure 3: The compatibility graph for Example 2. All pairs are willing. The left- and right-lobe donations are denoted by letters ℓ and r. There are three individually rational exchanges.

The two critical observations in this example are:

1. *each of the three feasible exchanges involves one left-lobe donation and one right-lobe donation, and*
2. *each of the three pairs is part of two feasible exchanges, one with a left-lobe donation, and another with a right-lobe donation.*

What this means is, the only left-lobe donation possibility for each pair depends on a right-lobe donation from another pair. As such, our proposed modification of "right-lobe donation from a pair is to be considered only when its left-lobe donation possibilities are exhausted" simply means, no exchange can be picked by the modified priority mechanism, no matter how pairs are priority ordered.

For example, suppose the pairs are processed following the priority order

$$\Pi = i_1 - i_2 - i_3.$$

Pair i_1 is processed first. There is no left-lobe-only exchange in the problem, and at this initial stage of the modified process no pair is to be considered for right-lobe donation. Hence pair i cannot be committed for an assignment just yet. On the other hand, since $\{i_1, i_3\}$ is a feasible match with $t(i_1, i_3) = \ell$ and $t(i_3, i_1) = r$, pair i_1 can donate a left lobe if pair i_3 donates a right lobe. Thus, pair i_1 cannot be made available –just yet– for right-lobe donation either. Hence, under our proposed modification, the process has to bypass pair i_1, neither committing it an assignment, nor making it available for a right-lobe donation. Moreover, exactly analogous situation occurs for pairs i_2 and i_3 in steps 2 and 3, respectively. Therefore, no exchange becomes permissible under our proposed modification, and thus no exchange is conducted.

This is clearly a Pareto inefficient outcome since picking any of the individually rational exchanges, each involving one left-lobe donation and one right-lobe donation, results in a Pareto improvement.

The failure of the priority matching approach in Example 2 is directly linked to the existence of a cycle of exchanges, in which

- each pair donates a right lobe clockwise in the cycle, and
- a left lobe counter-clockwise in the cycle.

As it turns out, this difficulty is not unique to our priority matching approach. In Proposition 2 of Appendix C.1, we show that the existence of a *left-lobe–right-lobe exchange cycle* of any number of pairs in the underlying compatibility graph, rules out the existence of a mechanism that is individually rational, Pareto-efficient, and incentive compatible.

This negative result, however, does not imply that the priority matching approach has to be fruitless in our model. Observe that, unlike in Example 1, we have not given the types of pairs in Example 2. We have only given a set of (allegedly) feasible exchanges, along with a transplant type function that gives rise to a compatibility graph with a left-lobe–right-lobe exchange cycle. As we show in Lemma 1 in Section 4.2, one of the key observations in our paper, the existence of such a cycle is ruled out in our model. But while left-lobe–right-lobe exchange cycles do not exist, *left-lobe–right-lobe exchange chains* can exist, and these chains has to be taken into consideration in our design. Hence, this observation enables us to formulate a further refinement of the priority matching approach to construct an individually rational, Pareto-efficient, and incentive-compatible mechanism for liver exchange.

We next formalize the structure that rules out a left-lobe–right-lobe exchange cycle in our model.

4.2 The Precedence Digraph

Consider two pairs of types $X - Y,\ U - V \in \mathbf{T} \times \mathbf{T}^{\mathbf{D}}$. 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 donates his left lobe to the patient of type X. Observe that, the two pairs cannot form an exchange where the donor of type V donates his right lobe to the patient of type X, and the donor of type Y donates his left lobe to the patient of type U, for otherwise they could have formed a left-lobe-only exchange as well. Therefore, just focusing on these two pairs for the moment, it would be plausible to avail the pair of type $X - Y$ for right-lobe donation prior to the pair of type $U - V$, because the left-lobe-exchange possibilities of the pair of type $U - V$ expands with the availability of type $X - Y$ pair for right-lobe donation.

We can extend this line of reasoning to problems with more than two pairs as well, provided that there are no left-lobe–right-lobe exchange cycles (of the sort we have seen in Example 2) in the following directed graph.

Definition 1 *The* **precedence digraph** $(\mathbf{T} \times \mathbf{T^D}, D^r)$ *is a directed graph where,*

1. *each pair type in* $\mathbf{T} \times \mathbf{T^D}$ *is a node,*
2. *there is a directed edge from type* $X - Y$ *to type* $U - V$*, denoted as* $X - Y \longrightarrow U - V$*, if and only if*

$$X \le V^\ell,\ U \not\le Y^\ell \ \&\ U \le Y^r, \quad \textit{and}$$

3. D^r *is the resulting set of directed edges.*

We say that $X - Y$ **precedes** $U - V$, whenever $X - Y \longrightarrow U - V$. In a precedence digraph, 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, the two pairs fail to form a left-lobe-only exchange, but they can form an exchange once the pair of type $X - Y$ becomes available for right-lobe donation.

Figures 4 depicts the precedence digraph for the case of two sizes $\mathbf{S} = \{0, 1\}$ such that

$$\text{for all } Y \in \mathbf{T^D}, \quad Y_{3\ell} = 0 \implies Y_{3r} = 1$$

when potentially all pair types are available for exchange. Figure 5 depicts the precedence digraph for the same case, but when only left-lobe incompatible pair types are available for exchange. Figure 6 depicts the precedence digraphs for the case of three sizes $\mathbf{S} = \{0, 1, 2\}$ such that

$$\text{for all } Y \in \mathbf{T^D}, \quad [Y_{3\ell} = 0 \implies Y_{3r} = 1] \text{ and } [Y_{3\ell} = 1 \implies Y_{3r} = 2]$$

when only left-lobe incompatible pair types are available for exchange.

Note that the precedence digraphs in Figure 4, 5, and 6 are all acyclic. This observation

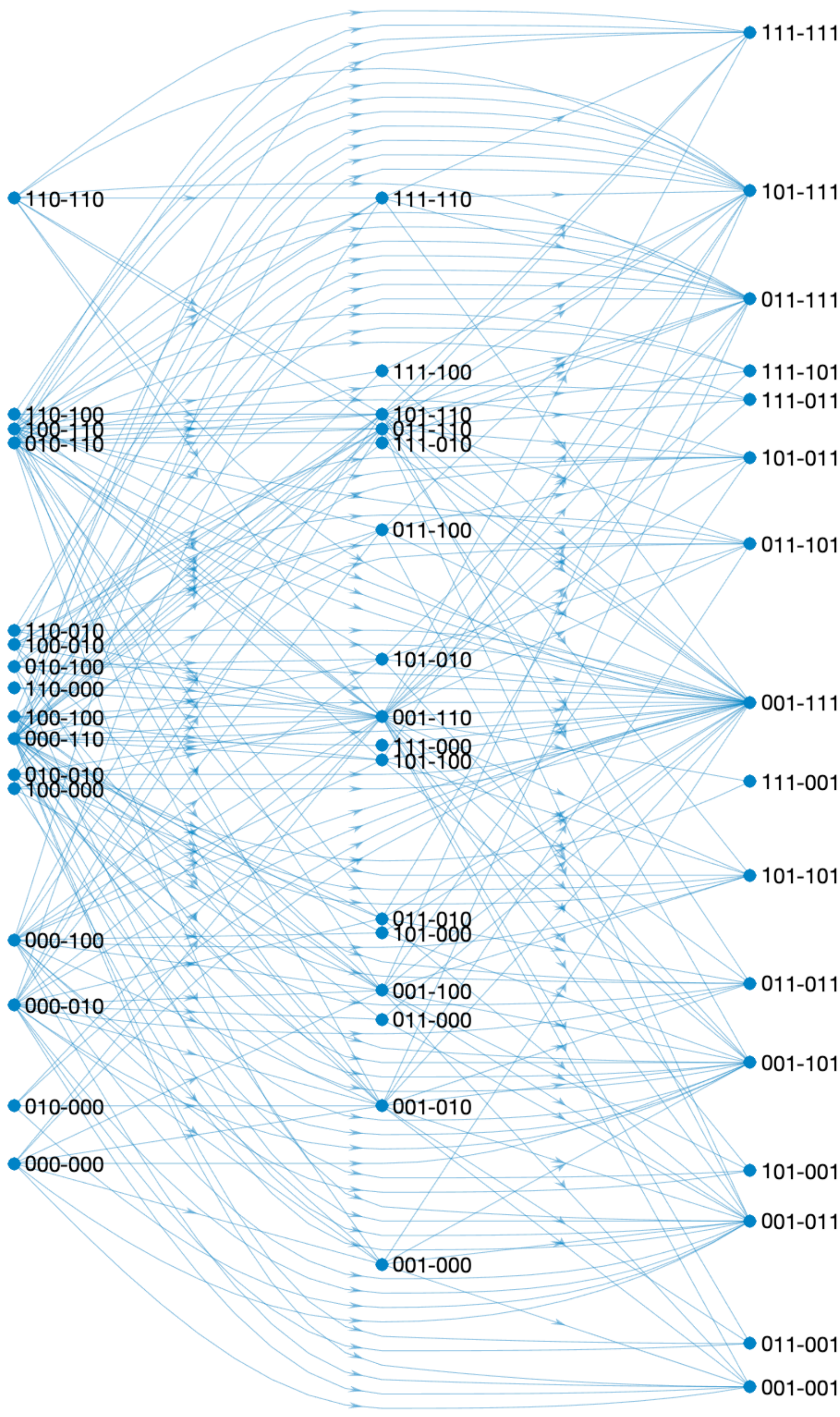

Figure 4: The precedence digraph with two sizes ($S = 2$). We only denote left-lobe size of the donor types in this depiction, as their right-lobe size is uniquely determined by their left-lobe size. 16 pair types have no adjacent edges in the digraph, so those are not shown.

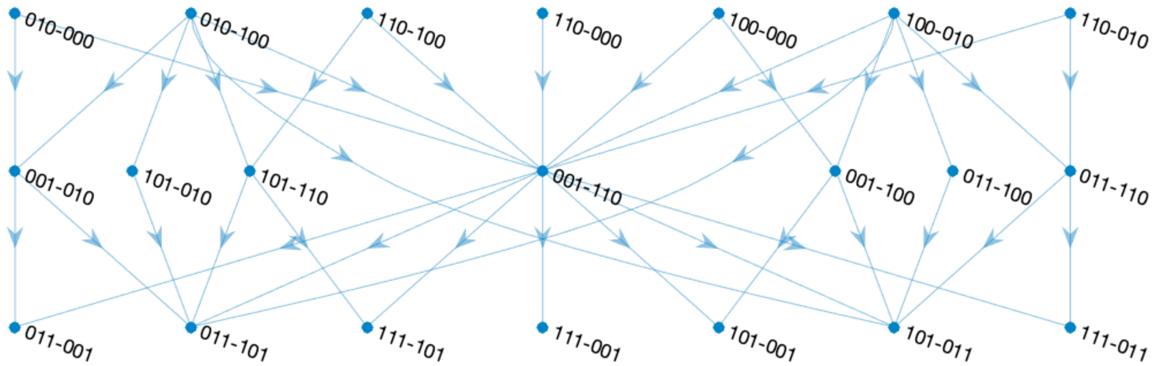

Figure 5: The precedence digraph with two sizes ($S = 2$) when left-lobe compatible pairs do not participate in exchange. We only denote left-lobe size of the donor types in this depiction, as their right-lobe size is uniquely determined by their left-lobe size. 16 pair types have no adjacent edges in the digraph, so those are not shown.

is not specific to these examples. As stated by the next lemma, the precedence digraph for any liver-exchange pool is acyclic.

Lemma 1 *The precedence digraph* $(\mathbf{T} \times \mathbf{T^D}, D^\tau)$ *is acyclic.*

Together with Lemma 4 in Appendix A, Lemma 1 imply that the precedence digraph is consistent with a linear order over the set of pair types, which can then be used to construct a linear order, called **topological order**, over the set of all pairs.

In order to do this, we extend the precedence digraph on the set of types of pairs $\mathbf{T} \times \mathbf{T^D}$ to the set of pairs $\mathcal{I}$. Let $(\mathbf{T} \times \mathbf{T^D}, D^\tau)$ be the precedence digraph of types of pairs. Construct a digraph $(\mathcal{I}, D)$ such that for any two types $X - Y,\ U - V \in \mathbf{T} \times \mathbf{T^D}$ and for any two pairs $i, j \in \mathcal{I}$ such that $\tau(i) = X - Y$ and $\tau(j) = U - V$:

$$(i, j) \in D \iff (X - Y, U - V) \in D^\tau.$$

We refer to $(\mathcal{I}, D)$ as the **precedence digraph** on $\mathcal{I}$. By Lemma 1, the precedence digraph on $\mathcal{I}$ is also acylic. By Lemma 4 (in Appendix A), we can fix a topological order of the precedence digraph on $\mathcal{I}$.

In order to maintain Pareto efficiency and incentive compatibility, our proposed mechanism relies on a topological order to determine both the processing order, and, if necessary, the activation order of the right-lobe donation possibilities for willing pairs.

4.3 Transformations & Deletions

In Section 4.2 we have established existence of a priority ordering, the topological order, that can potentially help us to overcome the challenges presented in Section 4.1. We next

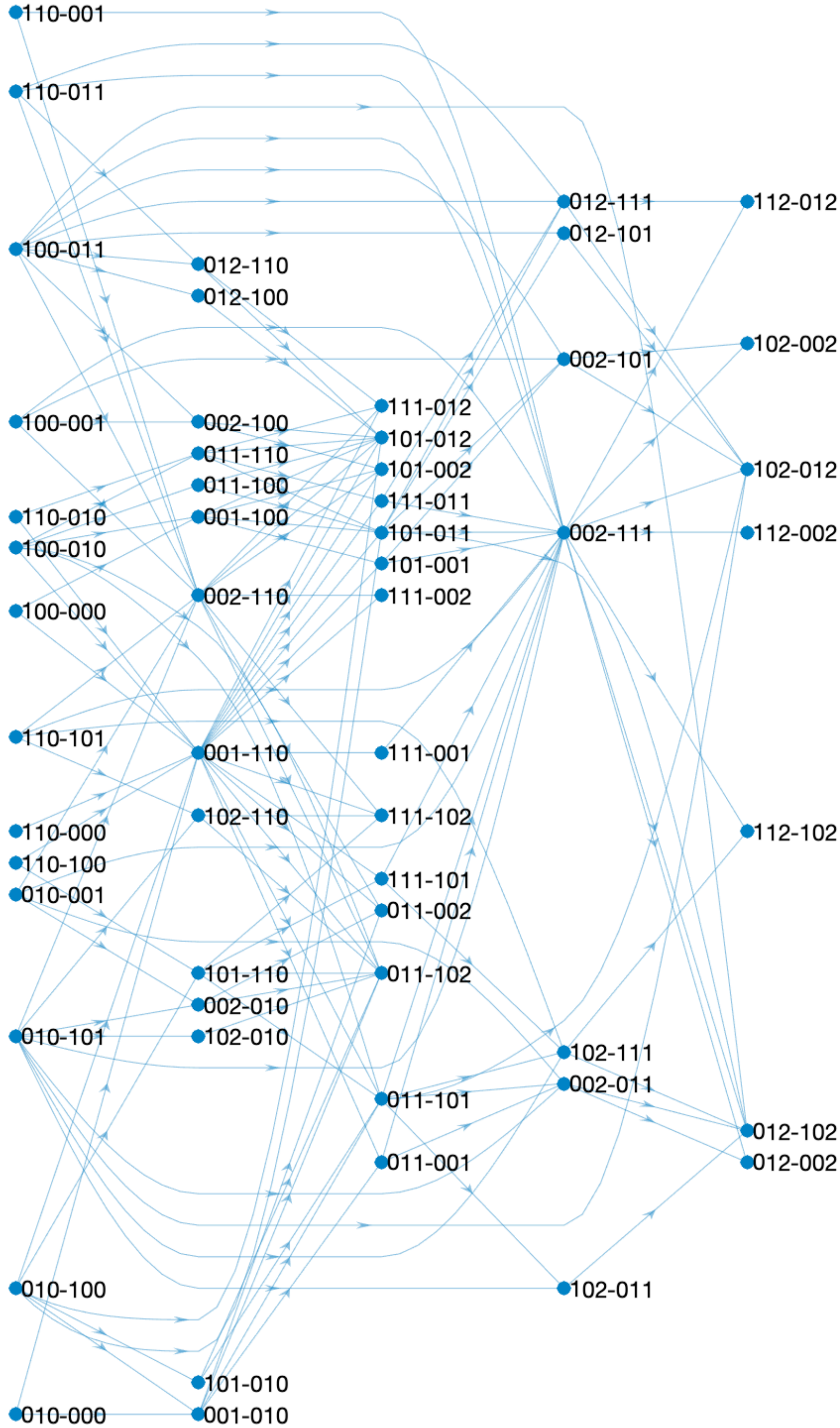

Figure 6: The precedence digraph with three sizes ($S = 3$) when left-lobe compatible pairs do not participate in exchange. We only denote left-lobe size of the donor types in this depiction, as their right-lobe size is uniquely determined by their left-lobe size. 34 pair types have no adjacent edges in the digraph, so those are not shown.

present some additional tools to facilitate the introduction of our proposed mechanism.

Fix a problem with the compatibility graph $G_c = (\mathcal{I}, E_c)$. For any set of matches $E' \subseteq E_c$, the graph $G' = (\mathcal{I}, E')$ is referred as a **reduced compatibility graph**.

For any set of matches $E' \subseteq E_c$ with the reduced compatibility graph G', let $\mathbf{M}[G'] \subseteq \mathbf{M}_c$ denote the resulting set of matchings. That is,

$$M \in \mathbf{M}[G'] \iff M \in \mathbf{M}_c \text{ and } M \subseteq E'.$$

Also denote the **set of matches involving i in E'** as

$$E'(i) = \Big\{\{i, j\} \in E'\Big\}.$$

For two reduced compatibility graphs $G' = (\mathcal{I}, E')$ and $G'' = (\mathcal{I}, E'')$ with $E'' \subseteq E'$, we refer to G'' as a **subgraph** of G'.

Given a preference profile R, let E_{IR} denote the set of **individually rational** matches, where, for any $i, j \in \mathcal{I}$,

$$\{i, j\} \in E_{IR} \iff \begin{cases} jR_i\emptyset \text{ and } jR_i i \\ \qquad \text{and} \\ iR_j\emptyset \text{ and } iR_j j \end{cases}$$

The reduced compatibility graph $G_{IR} = (\mathcal{I}, E_{IR})$ is referred as the **individually rational (IR) compatibility graph**.

In a general graph (i.e. not necessarily the compatibility graph of a liver-exchange pool) in which all pairs are indifferent between all exchanges, 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). While we allow for more general preferences in our model, we still rely on similar tools to design our mechanism.

A set of pairs $\mathcal{J} \subseteq \mathcal{I}$ is **matchable** in a reduced compatibility graph $G' = (\mathcal{I}, E')$, if there exists a matching $M \in \mathbf{M}[G']$ such that $M(j) \neq \emptyset$ for all $j \in \mathcal{J}$.[14]

Fix a pair $i \in \mathcal{I}$, a preference relation $R_i \in \mathbf{R}_i$, a reduced compatibility graph G', and a set of pairs $\mathcal{J} \subset \mathcal{I} \setminus \{i\}$ such that $\mathcal{J} \cup \{i\}$ is matchable in G'. For a given reduced compatibility graph G' and set of pairs $\mathcal{J}$ whose members are all committed to be matched, define the set of **achievable assignments** of pair i (while members of $\mathcal{J}$ are all matched) as

$$\mathcal{A}(i|\mathcal{J}, G') = \Big\{j \in \mathcal{I} \ : \ \exists M \in \mathbf{M}[G'] \text{ such that } M(h) \neq \emptyset \ \ \forall h \in \mathcal{J} \text{ and } M(i) = j\Big\},$$

[14]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).

and the set of **best achievable assignments** of i (while members of $\mathcal{J}$ are all matched) as

$$\mathcal{B}(i|\mathcal{J}, G') \equiv \max_{R_i} \mathcal{A}(i|\mathcal{J}, G').$$

Whenever a pair j in an achievable assignment for a pair i, the corresponding match i, j is referred as an **achievable match**. Similarly, whenever a pair j in a best achievable assignment for a pair i, the corresponding match i, j is referred as a **best achievable match**.

Whether a set of pairs is matchable or not can be checked in polynomial time.[15] Moreover, sets $\mathcal{A}(\cdot)$ and $\mathcal{B}(\cdot)$ can also be constructed in polynomial time.

We obtain the outcome of our proposed mechanism through an iterative algorithm introduced in Section 4.4. At the initiation of our algorithm, we start with the reduced compatibility graph where exclusively left-lobe-only matches are considered, which is modified through the following two operators as we run the algorithm:

1. **Deletions:** At the initiation of the algorithm, the reduced compatibility graph $G_0 = (\mathcal{I}, E_0)$ is activated, where

$$E_0 \equiv \Big\{ \{i, j\} \in E_{IR} \ : \ t(i, j) = \ell \text{ and } t(j, i) = \ell \Big\}.$$

   Left-lobe donation possibilities of pairs are fully utilized, by sequentially processing them following a fixed topological order, and committing to each processed pair one of its best achievable matches while donating a left lobe. Whenever a pair is committed to receive one of its best achievable matches in our algorithm, all less-preferred matches are *deleted* from the active reduced compatibility graph, and subsequently activating the resulting subgraph.
   Let $G' = (\mathcal{I}, E')$ be the active reduced compatibility graph prior to processing of pair i, and let $\mathcal{J}$ be the set of pairs whose members are committed for an assignment up to this point. If pair i is deemed matchable in G' in addition to pairs in $\mathcal{J}$, i.e., $\mathcal{A}(i|\mathcal{J}, G') \neq \emptyset$, the **deletion** operator activates the reduced compatibility graph $G'' = (\mathcal{I}, E'')$ in the algorithm, where

$$E'' \equiv \Big(E' \setminus E'(i) \Big) \cup \Big\{ \{i, j\} \in E'(i) \ : \ j \in \mathcal{B}(i|\mathcal{J}, G') \Big\}.$$

2. **Transformations:** As each pair is processed sequentially, the options for pairs further ahead in the topological order potentially shrink. If a pair i cannot be matched through a left-lobe donation, that means, not only its existing left-lobe donation possibilities are exhausted, but also any potential left-lobe donation possibilities are also exhausted by Lemma 1.[16] At this point we **transform** pair i deeming it available for right-lobe

[15] We provide a polynomial-time method for checking matchability in Appendix C.2.

[16] Observe that, processing the pairs following a topological order is key for this last argument to hold.

donation, for the first time and until the termination of the algorithm, provided that it is willing.

Given a preference profile R, let $G' = (\mathcal{I}, E')$ be the reduced compatibility graph that is active under our algorithm, prior to pair i's transformation. After we transform pair i, a new reduced compatibility graph $G'' = (\mathcal{I}, E'')$ is activated, where

$$E'' = E' \cup \Big\{ \{i,j\} \in E_{IR}(i) \ : \ t(i,j) = r \Big\}.$$

Since $E'' \supseteq E'$, a transformation potentially enlarges the set of matches in the compatibility graph.

4.4 Precedence-Induced Adaptive-Priority Mechanism

We are ready to present an iterative algorithm, which can be used to find the outcome of our proposed **Precedence-Induced Adaptive-Priority Mechanism**.

For the rest of this section, fix a liver-exchange pool $(\mathcal{I}, \tau)$, a preference profile $R \in \mathbf{R}$, a topological order over pairs Π_ℓ, and an arbitrary priority order over pairs Π_r. While the priority order Π_r can be the same as the topological order Π_ℓ, it does not have to be. We refer to Π_ℓ as the **left-lobe matching topological order** and to Π_r as the **right-lobe matching priority order**.

Precedence-Induced Adaptive-Priority-Matching Algorithm:

Step 1: Let $\mathcal{I} = \{i_1, \ldots, i_K\}$ be the enumeration of pairs with respect to the left-lobe matching topological order Π_ℓ. We inductively construct

1. a sequence of reduced compatibility graphs

$$G_0 = (\mathcal{I}, E_0), \ldots, G_K = (\mathcal{I}, E_K)$$

   such that $E_k \subseteq E_{IR}$ for all k, and
2. two sequences of enlarging pair sets

$$\mathcal{J}_0 \subseteq \mathcal{J}_1 \subseteq \ldots \subseteq \mathcal{J}_K \subseteq \mathcal{I} \quad \text{and} \quad \widetilde{\mathcal{J}}_0 \subseteq \widetilde{\mathcal{J}}_1 \subseteq \ldots \subseteq \widetilde{\mathcal{J}}_K \subseteq \mathcal{I}$$

through substeps of Step 1.

At the initiation, activate the reduced compatibility graph $G_0 = (\mathcal{I}, E_0)$ where

$$E_0 \equiv \Big\{ \{i,j\} \in E_{IR} \ : \ t(i,j) = \ell \text{ and } t(j,i) = \ell \Big\}.$$

That is, E_0 is restricted to IR matches in which each donor donates a left lobe.

Also define

$$\mathcal{J}_0 \equiv \emptyset \quad \text{and} \quad \widetilde{\mathcal{J}}_0 \equiv \emptyset.$$

Step 1's substeps proceed inductively:

Step 1.(k): Consider pair i_k, the k'th highest-priority pair under Π_ℓ. Sets $\mathcal{J}_{k-1}$, $\widetilde{\mathcal{J}}_{k-1}$, and the reduced compatibility graph $G_{k-1} = (\mathcal{I}, E_{k-1})$ are defined at Step 1.($k-1$).

- If $\mathcal{J}_{k-1} \cup \{i_k\}$ is matchable in G_{k-1}, then let

$$\mathcal{J}_k \equiv \mathcal{J}_{k-1} \cup \{i_k\}, \quad \widetilde{\mathcal{J}}_k \equiv \widetilde{\mathcal{J}}_{k-1}, \text{ and } G_k = (\mathcal{I}, E_k) \text{ where}$$

$$E_k \equiv \Big[E_{k-1} \setminus E_{k-1}(i_k)\Big] \cup \Big\{\{i_k, j\} \ : \ j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_{k-1})\Big\}.$$

  That is, E_k is obtained from E_{k-1} by *deleting* all matches involving pair i_k except its best achievable matches (when all pairs in $\mathcal{J}_{k-1}$ can also be matched).

- Otherwise, let

$$\mathcal{J}_k \equiv \mathcal{J}_{k-1} \quad \text{and}$$

  * if i_k is not willing, then let

$$\widetilde{\mathcal{J}}_k \equiv \widetilde{\mathcal{J}}_{k-1} \text{ and } G_k = (\mathcal{I}, E_k) \equiv G_{k-1};$$

  * if i_k is willing, then let

$$\widetilde{\mathcal{J}}_k \equiv \widetilde{\mathcal{J}}_{k-1} \cup \{i_k\} \text{ and } G_k = (\mathcal{I}, E_k) \text{ where}$$

$$E_k \equiv E_{k-1} \cup \Big\{\{i_k, j\} \in E_{IR}(i_k) \ : \ t(i_k, j) = r\Big\}.$$

    That is, E_k is obtained by *transforming* pair i_k, i.e., by adding to E_{k-1} the individually rational matches in which pair i_k donates a right lobe.

Proceed with Step 1.($k+1$).

Step 1 terminates at substep K, where $K = |\mathcal{I}|$ is the number of pairs. This step determines:

1. the reduced compatibility graph $G_K = (\mathcal{I}, E_K)$, where one of its matchings is to be selected as the eventual outcome of the algorithm,
2. the finalized set of pairs $\mathcal{J}_K$ whose members are each matched by donating a left lobe, and their eventual utility level,
3. the set of pairs $\mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{J}_K]$, whose members –unwilling to donate a right lobe – definitely remain unmatched, and

4. the set of pairs $\widetilde{J}_K$, whose members may be matched by donating right lobe or remain unmatched.

We proceed with Step 2, which determines the utility level of pairs in $\widetilde{J}_K$, along with the final outcome of the algorithm.

Step 2: Inductively, we continue with the pairs in $\widetilde{J}_K$, whose members can be potentially matched –albeit through right-lobe donation– in addition to members of $\mathcal{J}_K$ (who are already committed to be matched). Let $\widetilde{\mathcal{J}}_K = \{i_1^*, \ldots, i_N^*\}$ be the enumeration of those pairs with respect to the right-lobe matching priority order Π_r. We construct

1. a sequence of shrinking reduced compatibility graphs

$$G_0^* = (\mathcal{I}, E_0^*), \ldots, G_N^* = (\mathcal{I}, E_N^*)$$

such that $E_N^* \subseteq \ldots \subseteq E_0^*$, and

2. a sequence of enlarging pair sets

$$\mathcal{J}_0^* \subseteq \ldots \subseteq \mathcal{J}_N^*.$$

At the initiation of Step 2, set

$$\mathcal{J}_0^* \equiv \emptyset \quad \text{and} \quad G_0^* = (\mathcal{I}, E_0^*) \equiv G_K.$$

Step 2.(n): Consider i_n^*, the n'th highest-priority pair in $\widetilde{\mathcal{J}}_K$ according to the right-lobe matching priority order Π_r. Pair set $\mathcal{J}_{n-1}^*$ and reduced compatibility graph $G_{n-1}^* = (\mathcal{I}, E_{n-1}^*)$ are constructed at Step 2.$(n-1)$.

- If $\mathcal{J}_K \cup \mathcal{J}_{n-1}^* \cup \{i_n^*\}$ is matchable in G_{n-1}^*, then let

$$\mathcal{J}_n^* \equiv \mathcal{J}_{n-1}^* \cup \{i_n^*\} \text{ and } G_n^* = (\mathcal{I}, E_n^*) \text{ where}$$

$$E_n^* \equiv \Big[E_{n-1}^* \setminus E_{n-1}^*(i_n^*)\Big] \cup \Big\{\{i_n^*, j\} \ : \ j \in \mathcal{B}(i_n^* | \mathcal{J}_K \cup \mathcal{J}_{n-1}^*, G_{n-1}^*)\Big\}.$$

  That is, E_n^* is obtained from E_{n-1}^* by *deleting* the matches involving pair i_n^*, except its best achievable matches (when all pairs in $\mathcal{J}_K \cup \mathcal{J}_{n-1}^*$ can also be matched).
- Otherwise, let

$$\mathcal{J}_n^* \equiv \mathcal{J}_{n-1}^* \quad \text{and} \quad G_n^* = (\mathcal{I}, E_n^*) \equiv G_{n-1}^*.$$

Proceed with Step 2.$(n+1)$.

When Step 2 terminates at substep $N = |\widetilde{\mathcal{J}}_K|$, the mechanism picks, as its outcome, a matching of G^*_N that matches all pairs in $\mathcal{J}_K$ and $\mathcal{J}^*_N$.

As we prove in Lemma 3 in Section 5, the utility level of each pair is uniquely determined at each such matching, thus, the outcome of the mechanism is utility-wise well defined.[17]

5 Results

We first state the properties of the reduced compatibility graphs, the resulting matchings, and the sets of pairs constructed through the steps of our algorithm. We define the following sets of matchings, which can be interpreted as the outcomes of Step 1.(k) for any $k \in \{1, ..., K\}$ and Step 2.(n) for any $n \in \{1, ..., N\}$ of the mechanism. We analyze the properties of these sets of matchings in Lemmas 2 and 3 below.

Define
$$\mathbf{M}_k \equiv \left\{M \in \mathbf{M}[G_k] \ : \ M(j) \neq \emptyset \ \ \forall j \in \mathcal{J}_k\right\}$$
as the subset of matchings of G_k, the compatibility graph activated at the end of Step 1.(k) of the algorithm, such that each matching in this set matches each pair in $\mathcal{J}_k$, which is the set of committed pairs up to the end of Step 1.(k).

Define
$$\mathbf{M}^*_n \equiv \left\{M \in \mathbf{M}[G^*_n] \ : \ M(j) \neq \emptyset \ \ \forall j \in \mathcal{J}_K \cup \mathcal{J}^*_n\right\}$$
as the subset of matchings of G^*_n, the compatibility graph activated at the end of Step 2.(n) of the algorithm, such that each matching in this set matches each pair in $\mathcal{J}_K \cup \mathcal{J}^*_n$, which is the set of committed pairs up to the end of Step 2.(n). Thus, any matching in $\mathbf{M}^*_N$ can be the outcome of the precedence-induced adaptive-priority mechanism.

We state the properties of matchings in $\mathbf{M}_K$ below in Lemma 2. This lemma shows, among other results, that

(i) set $\mathcal{J}_K$ is matchable in G_K such that each pair in $\mathcal{J}_K$ donates a left lobe,
(ii) when every pair in $\mathcal{J}_K$ is matched in a matching of G_K, pairs in $\mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K]$ remain unmatched, and
(iii) each pair in $\mathcal{J}_K$ is indifferent between all matchings of G_K that match all pairs in $\mathcal{J}_K$.

Lemma 2 (Properties of Constructs in Step 1) *Consider the sequences of reduced compatibility graphs $\{G_k\}_{k=1}^K$ and pair sets $\{\mathcal{J}_k, \widetilde{\mathcal{J}}_k\}_{k=1}^K$ constructed through the substeps of Step 1 of the precedence-induced adaptive-priority-matching algorithm.*

1. *Set of pairs $\mathcal{J}_K$ is matchable in G_K.*
2. *$\mathbf{M}_K \neq \emptyset$, and, for any matching $M \in \mathbf{M}_K$,*
   (a) for all $j \in \mathcal{I} \setminus \widetilde{\mathcal{J}}_K$ and all $M' \in \mathbf{M}_K$, $M(j) \ I_j \ M'(j)$,

[17] We provide a polynomial-time method for how to find such a matching in Appendix **??**.

(b) *for all* $j \in \mathcal{J}_K$, $M(j) \in \mathcal{E}^\ell(j)$,
(c) *for all* $j \in \widetilde{\mathcal{J}}_K$, $M(j) \notin \mathcal{E}^\ell(j)$,
(d) *for all* $j \in \mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K]$, $M(j) = \emptyset$, *and*
(e) *for all* $i_k \in \mathcal{J}_K$, $M(i_k) \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_K)$; *moreover,* $M(i_k)\ I_{i_k}\ j$ *for all* $j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_k)$.

We state the properties of matchings in $\mathbf{M}^*_N$ below in Lemma 3. This lemma shows, among other results, that

(i) the set of pairs $\mathcal{J}_K \cup \mathcal{J}^*_N$ is the unique maximal matchable set in G^*_N such that each pair in $\mathcal{J}_K$ donates a left lobe and each pair in $\mathcal{J}^*_N$ donates a right lobe in every maximal matching, and
(ii) each pair is indifferent among all matchings of G^*_N that match pairs in $\mathcal{J}_K \cup \mathcal{J}^*_N$, and thus, the mechanism outcome is utility-wise uniquely defined.

Lemma 3 (Properties of Constructs in Step 2) *Consider the sequences of reduced compatibility graphs* $\{G^*_n\}_{n=1}^N$ *and pair sets* $\{\mathcal{J}^*_n\}_{n=1}^N$ *constructed through the substeps of Step 2 of the precedence-induced adaptive-priority-matching algorithm.*

1. $\mathbf{M}^*_N \subseteq \mathbf{M}_K$ *and* $\mathcal{J}^*_N \subseteq \widetilde{\mathcal{J}}_K$.
2. *Set of pairs* $\mathcal{J}_K \cup \mathcal{J}^*_N$ *is matchable in* G^*_N.
3. $\mathbf{M}^*_N \neq \emptyset$, *and, for any matching* $M \in \mathbf{M}^*_N$,
   (a) *for all* $j \in \mathcal{I}$ *and all* $M' \in \mathbf{M}^*_N$, $M(j)\ I_j\ M'(j)$,
   (b) *for all* $j \in \mathcal{J}_K$, $M(j) \in \mathcal{E}^\ell(j)$,
   (c) *for all* $j \in \mathcal{J}^*_N$, $M(j) \in \mathcal{E}^r(j)$,
   (d) *for all* $j \in \mathcal{I} \setminus [\mathcal{J}_K \cup \mathcal{J}^*_N]$, $M(j) = \emptyset$,
   (e) *for all* $i_k \in \mathcal{J}_K$, $M(i_k) \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G^*_N)$, *and*
   (f) *for all* $i^*_n \in \mathcal{J}^*_N$, $M(i^*_n) \in \mathcal{B}(i^*_n|\mathcal{J}_K \cup \mathcal{J}^*_{n-1}, G^*_N)$; *moreover,* $M(i^*_n) \in \mathcal{B}(i^*_n|\mathcal{J}_K \cup \mathcal{J}^*_{n-1}, G^*_n)$.

Parts of these two lemmas will also be used in proving our main result, Theorem 1 below.

We are ready to present our main result.

Theorem 1 *The precedence-induced adaptive-priority mechanism is individually rational, Pareto efficient, and incentive compatible.*

5.1 Transplant Maximization and Incentive Compatibility

It is well known that when all patients are indifferent among compatible grafts and right-lobe donation is not allowed, every Pareto efficient matching maximizes the number of transplants (see Korte and Vygen, 2011, Roth, Sönmez, and Ünver, 2005, and Sönmez and Ünver, 2014).

When right-lobe donation becomes feasible, this equivalence no longer holds. Moreover, although our proposed precedence-induced adaptive-priority mechanism is Pareto efficient, it may not maximize the number of transplants even when there are only two sizes and all patients are indifferent among all compatible grafts.

The example below shows that one has to sacrifice incentive compatibility in order to maximize the number of transplants, even when patients are indifferent among all compatible grafts. Indeed the same example also shows that, one has to sacrifice incentive compatibility in order to maximize the number of the safer left-lobe transplants as well.

Example 3 *Suppose there are two sizes, small and large, denoted as* $S = \{0, 1\}$*. Consider a liver-exchange pool with* $\mathcal{I} = \{i_1, i_2, i_3, i_4\}$*. The pair types are given as follows:*

$$\begin{aligned} \tau_P(i_1) &= (1,0,1) \qquad \tau_D(i_1) = (0,1,1,1), \\ \tau_P(i_2) = \tau_P(i_4) &= (0,1,1) \qquad \tau_D(i_2) = \tau_D(i_4) = (1,0,0,1) \\ \tau_P(i_3) &= (1,0,0) \qquad \tau_D(i_3) = (0,1,1,1). \end{aligned}$$

Suppose all pairs are indifferent among all compatible grafts under the received-graft preference profile $\succsim$*. Suppose also that, pairs* i_2 *and* i_4 *are both willing. The individually rational compatibility graph of this problem is given in Figure 7.*

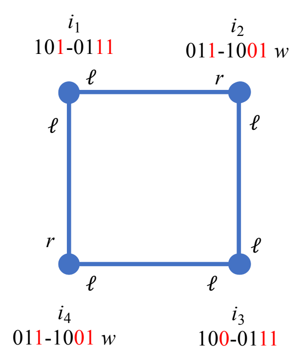

Figure 7: The individually rational compatibility graph for Example 3. The willing types are denoted by letter w following their types. The left- and right-lobe donations are denoted by letters ℓ and r. There are four individually rational exchanges.

Any left-lobe-donation-maximizing or total-transplant-maximizing matching (two of which can be obtained by swapping i_2 *and* i_4 *with each other) generates two exchanges. Consider these two matchings:*

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

Observe that $t(i_2, i_1) = t(i_4, i_1) = r$ while $t(i_2, i_3) = t(i_4, i_3) = \ell$. 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_2 has an incentive to announce its right-lobe donation willingness type as unwilling by revealing $R'_{i_2} = R^{e/u}_{i_2}$, 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 3 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 even when all patients are indifferent among compatible grafts.*

Establishing such an impossibility is straightforward when received-graft preferences admit strict preferences, thus we skip it.

6 Simulations

In this section, we report the results of computer simulations to determine the potential welfare gains from liver exchange. We use South Korean aggregate statistics in our simulations, since this country leads the world both in living-donor liver transplants and in liver exchange.

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.[18]

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 blood-type compatible, the donor's right-lobe volume is at least 40% of the total liver volume of the patient, although the donor's left lobe volume is less than 40% of the total liver volume of the patient.

We generate $K = 50$, 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 willingness rate, we randomly determine each pair's willingness.

We make the following two assumptions for preferences of pairs, as we do not have a better measure of more nuanced preferences over liver exchanges:

1. All pairs are *direct-transplant type*. This implies all left-lobe compatible pairs prefer a direct transplant to any type of exchange and all right-lobe compatible pairs prefer a direct transplant to any exchange in which they donate right lobe. Incompatible pairs are not affected by this assumption.
2. All pairs are indifferent among all compatible received-grafts in their received-graph preferences $\succsim$.

These two assumptions are standard in kidney exchange literature, as well as most of its real-life applications. In most kidney exchange applications, pairs have direct-transplant bias as they may not want to wait for an exchange. Moreover, transplant doctors mostly care about the compatibility of the received graft as the first-order coarse received-graph preference relation. Thus, we expect these also to be the case in liver exchange, especially in the short run until more data becomes available regarding exchanges.

We consider the following four treatments:

1. *No exchange.* Left-lobe-compatible pairs and right-lobe-only-compatible willing pairs participate in direct transplants.
2. *RSÜ priority mechanism for left-lobe exchanges.* Left-lobe-compatible pairs participate exclusively in direct transplants. Restricting the compatibility graph to left-lobe-only exchanges in the remaining problem, the outcome of the Roth, Sönmez, and Ünver (2005)

[18] We use the following weight determination formula as a function of height (also see Ergin et al. 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).

(RSÜ) priority mechanism, mainly introduced for kidney exchange, is determined for an arbitrary priority order. Right-lobe-only-compatible willing pairs participate in direct transplants only if they cannot be matched through left-lobe only exchanges.

3. *Proposed Pareto-efficient, individually rational, and incentive-compatible mechanism.* An outcome of our precedence-induced adaptive-priority mechanism is determined for arbitrary topological and priority orders.
4. *A maximum individually rational matching under full information.* Assuming that the willingness profile is known, we find a maximum individually rational matching as follows: We first deem each left-lobe compatible pair only compatible with themselves. For willing right-lobe-only compatible pairs, in addition to their left-lobe transplant options, we only make direct (right-lobe) transplant a feasible option. We transform all willing incompatible pairs at the initiation, deeming them available for right-lobe transplantation right away. Then, we find a maximum matching of the induced compatibility graph.[19]

The second treatment depicts a baseline scenario for measuring the benefits from liver exchange using off-the-shelf methods introduced for kidney exchange. Thus, it utilizes exchanges only for left-lobe transplants. Although this procedure is individually rational and incentive compatible as a mechanism, it is not Pareto efficient.

The fourth treatment depicts a hypothetical situation assuming willingness profiles of the pairs are known. We use this as a benchmark when the goal is to maximize number of transplants. This procedure is not incentive compatible as a mechanism.

The results of the simulations are given in Table 2 and Figure 8.[20] About 12.5% of all pairs are left-lobe compatible and their patients receive a direct left-lobe transplant. About 45.5% of all pairs are right-lobe-only compatible, and up to this percentage of the patients receive a direct right-lobe transplant as a linear, increasing function of the willingness rate (see the no exchange treatment in the figure). 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 78% of all pairs, in a seemingly concave, increasing function of the willingness rate for

[19]We implement Sönmez and Ünver (2014) priority mechanism in this case to find a maximum matching from an arbitrary priority order. This mechanism is maximum when the compatibility graph includes two-way exchanges and direct transplants.

[20]We caution the readers that we do not consider the possible genetic relationship between a paired donor and patient and assume that their blood types and sizes are independently distributed. Also pairs consisting of spouses may have positive correlation for their sizes although their blood types are unrelated. This independence assumption works in favor of our simulated gains from exchange. Also we only consider adult patients. Living-donor transplants from parents to their children are in non-negligible numbers in countries such as the US. The left-lobe compatibility instances within such pairs should be more frequent than within baseline pairs. The exclusion of such pairs works in favor of our simulated gains from exchange as well. On the other hand, we used the recipient percentages to determine the gender of donors and patients. Females are in general smaller than males. This data has selection bias as probably we observe more size-compatible pairs than the underlying entry population. This effect works against our exchange simulations.

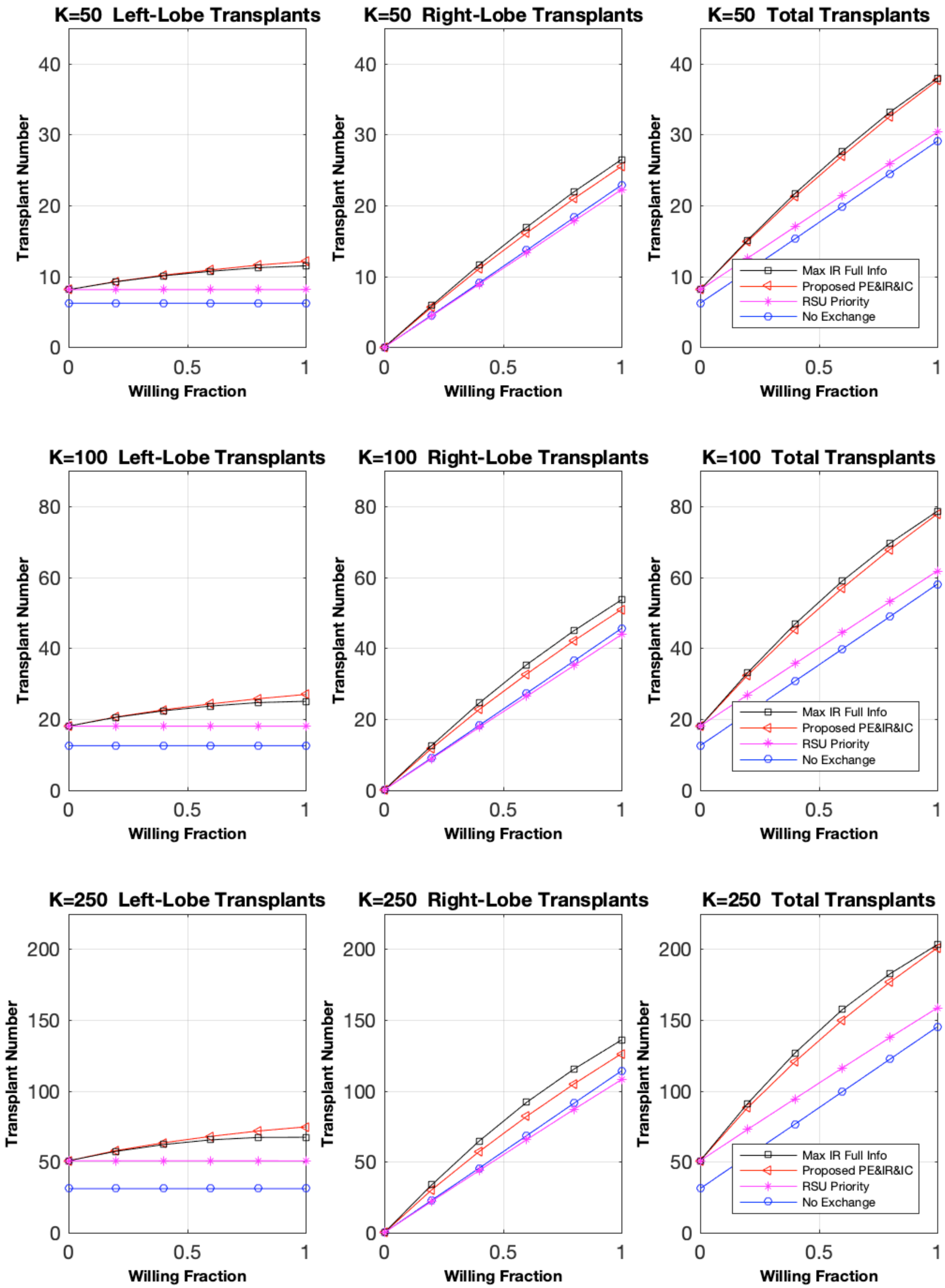

Figure 8: Simulation averages

		No Exchange			RSÜ Priority Mechanism			Proposed PE&IR&IC Mechanism			A Maximum IR Matching under Full Information		
Pop. Size K	w Fraction	Transplants Left L.	Right L.	Total	Transplants Left L.	Right L.	Total	Transplants Left L.	Right L.	Total	Transplants Left L.	Right L.	Total
50	0	6.204	0	6.204	8.122	0	8.122	8.122	0	8.122	8.122	0	8.122
		(2.280)	(0.000)	(2.280)	(2.831)	(0.000)	(2.831)	(2.831)	(0.000)	(2.831)	(2.831)	(0.000)	(2.831)
	0.2	6.204	4.529	10.733	8.12	4.439	12.559	9.272	5.649	14.921	9.232	5.885	15.117
		(2.280)	(2.064)	(2.912)	(2.826)	(2.050)	(3.305)	(2.869)	(2.270)	(3.609)	(2.855)	(2.313)	(3.694)
	0.4	6.204	9.081	15.285	8.12	8.881	17.001	10.164	11.059	21.223	10.074	11.616	21.69
		(2.280)	(2.810)	(3.330)	(2.826)	(2.775)	(3.644)	(2.920)	(2.899)	(3.966)	(2.895)	(2.993)	(4.139)
	0.6	6.204	13.665	19.869	8.122	13.318	21.44	10.904	16.103	27.007	10.715	16.922	27.637
		(2.280)	(3.179)	(3.491)	(2.831)	(3.150)	(3.756)	(2.923)	(3.211)	(3.945)	(2.887)	(3.296)	(4.096)
	0.8	6.204	18.301	24.505	8.122	17.808	25.93	11.589	20.963	32.552	11.224	21.932	33.156
		(2.280)	(3.496)	(3.581)	(2.831)	(3.489)	(3.766)	(2.931)	(3.476)	(3.825)	(2.868)	(3.542)	(3.924)
	1	6.204	22.87	29.074	8.12	22.239	30.359	12.101	25.507	37.608	11.481	26.465	37.946
		(2.280)	(3.624)	(3.507)	(2.832)	(3.671)	(3.607)	(2.940)	(3.614)	(3.558)	(2.833)	(3.695)	(3.630)
100	0	12.497	0	12.497	17.995	0	17.995	17.995	0	17.995	17.995	0	17.995
		(3.367)	(0.000)	(3.367)	(4.526)	(0.000)	(4.526)	(4.526)	(0.000)	(4.526)	(4.526)	(0.000)	(4.526)
	0.2	12.497	9.097	21.594	17.995	8.839	26.834	20.593	11.754	32.347	20.484	12.56	33.044
		(3.367)	(2.931)	(4.255)	(4.526)	(2.909)	(5.107)	(4.586)	(3.263)	(5.544)	(4.582)	(3.401)	(5.752)
	0.4	12.497	18.196	30.693	17.991	17.641	35.632	22.594	22.671	45.265	22.329	24.5	46.829
		(3.367)	(3.928)	(4.605)	(4.520)	(3.865)	(5.298)	(4.548)	(4.079)	(5.725)	(4.514)	(4.310)	(6.082)
	0.6	12.497	27.277	39.774	17.989	26.378	44.367	24.292	32.679	56.971	23.667	35.318	58.985
		(3.367)	(4.442)	(4.814)	(4.524)	(4.412)	(5.336)	(4.545)	(4.479)	(5.595)	(4.380)	(4.761)	(5.944)
	0.8	12.497	36.402	48.899	17.989	35.106	53.095	25.738	42.111	67.849	24.658	44.937	69.595
		(3.367)	(4.884)	(5.032)	(4.524)	(4.935)	(5.323)	(4.526)	(4.895)	(5.389)	(4.318)	(4.999)	(5.572)
	1	12.497	45.561	58.058	17.971	43.806	61.777	26.945	50.897	77.842	25.006	53.629	78.635
		(3.367)	(5.186)	(5.062)	(4.513)	(5.306)	(5.234)	(4.514)	(5.138)	(5.243)	(4.229)	(5.196)	(5.255)
250	0	31.031	0	31.031	50.683	0	50.683	50.683	0	50.683	50.683	0	50.683
		(5.236)	(0.000)	(5.236)	(7.681)	(0.000)	(7.681)	(7.681)	(0.000)	(7.681)	(7.681)	(0.000)	(7.681)
	0.2	31.031	22.895	53.926	50.683	22.109	72.792	57.889	30.228	88.117	57.354	33.686	91.04
		(5.236)	(4.746)	(6.572)	(7.681)	(4.692)	(8.329)	(7.820)	(5.175)	(9.060)	(7.661)	(5.597)	(9.579)
	0.4	31.031	45.5	76.531	50.679	43.81	94.489	63.368	57.189	120.557	62.138	64.615	126.753
		(5.236)	(6.355)	(7.263)	(7.677)	(6.280)	(8.592)	(7.819)	(6.488)	(9.052)	(7.501)	(6.963)	(9.827)
	0.6	31.031	68.387	99.418	50.659	65.528	116.187	67.925	81.993	149.918	65.457	91.913	157.37
		(5.236)	(7.287)	(7.639)	(7.673)	(7.329)	(8.502)	(7.797)	(7.272)	(8.508)	(7.411)	(7.730)	(9.107)
	0.8	31.031	91.294	122.325	50.643	86.973	137.616	71.718	104.914	176.632	67.178	115.387	182.565
		(5.236)	(7.870)	(7.777)	(7.668)	(8.097)	(8.275)	(7.765)	(7.838)	(8.262)	(7.139)	(8.105)	(8.449)
	1	31.031	114.084	145.115	50.613	107.917	158.53	74.598	126.228	200.826	67.291	135.859	203.15
		(5.236)	(8.290)	(7.744)	(7.654)	(8.660)	(7.879)	(7.677)	(8.480)	(7.759)	(6.888)	(8.391)	(7.726)

Table 2: Simulation results for population sizes $K = 50$, 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. For the standard errors of the averages all these standard errors need to be divided by $\sqrt{1000} \approx 31.62$.

$K = 100$ (see proposed PE&IR&IC treatment in the figure).[21] Thus, for a population size of $K = 100$, the percentage-wise 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.[22] 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,

[21] This concavity is caused by the initial fast increase in the scope of exchange when right-lobe donation becomes feasible for lower willingness fractions. For example, for w-fraction equal to 0, the exchange's contribution is rather low only an additional 10% of pairs are matched over no exchange.

[22] The increase in the number of transplants (rather than the percentage-wise increase compared to no exchange scenario) is higher for higher willingness rates.

the rate of increase in left-lobe transplants is higher than the rate of increase in right-lobe transplants.[23]

The baseline off-the-shelf RSÜ priority treatment matches more patients than no exchange, and the difference slightly decreases as willingness rate increases. However, when compared to our mechanism, it results in substantially fewer transplants whenever right-lobe transplantation is a viable option, i.e., w-fraction is positive. The percentage-wise increase in the number of transplants due to the availability of our mechanism instead of the RSÜ priority mechanism is in the range of 20.5% to 28.4% for $K = 100$ when right-lobe transplant option is viable, and more than 26% when w-fraction is 0.4 or more.

When compared to the maximum IR matching treatment, our mechanism does fairly well in terms of the numbers of transplants, despite the favorable treatment received by the former mechanism due to the full information assumption on willingness to donate a right lobe. Unlike our proposed mechanism, this mechanism is not incentive compatible, and hence it's outcome is best interpreted as a hypothetical maximum. Indeed, the number of left-lobe transplants are higher under our proposed mechanism than this hypothetical maximum.[24] Since the total number of transplants has to be weakly higher under the hypothetical maximum, our proposed mechanism yields fewer right-lobe transplants. As the worst case, when $K = 250$, the total number of transplants change between 100% (when w-fraction is 0) to 95.1% (when w-fraction is 0.4) and then back to 98.9% (when w-fraction is 1) of those of the hypothetical maximum IR matching. These ratios, while always less than 100% by definition, they are more favorable for our proposed mechanism when the population size is smaller with $K = 50$ and $K = 100$.

While the total number of transplants is higher under the maximum IR matching treatment than our proposed mechanism, the total number of transplants does not necessarily represent the best metric for social welfare. The *double equipoise* is a widely accepted concept in evaluating the balance between donor risk and recipient benefit in living donor liver transplantation, and according to this theory it should be performed only if the donor risk is justified by the acceptable outcome for the recipient. Based on this theory, Roll et al. (2013) propose the metric of *recipient lives saved at 5 years per donor death* to evaluate various liver transplantation policies. For a population size of $K = 100$ and willingness rate of 100%, the expected number of left lobe transplants under our proposed mechanism is more

[23] Even under the most conservative predictions, our simulations also show that potential gains from liver exchange has not been fully realized in South Korea, especially those due to size incompatibility. ASAM Medical Center, the leading living-donor liver transplantation center in the world, reports that between 2003 and 2011 only 26 patients were transplanted through exchange, which is only 1.2% of all living-donor transplants conducted in the center (see Jung et al., 2014). Moreover, they note that only 4 of these patients participated in exchange for size incompatibility reasons, while the rest participated in exchange because of blood-type incompatibility.

[24] We choose one arbitrary full-information priority matching in this graph, and we do not aim to minimize the number of right-lobe transplants among all possible priority matchings of the graph, as there may be multiple priority matchings that match the same set of pairs in the induced compatibility graph.

than 2 units higher than under the maximum IR matching treatment, whereas the expected number of right lobe transplants is less than 3 units lower. Therefore, given the five-fold donor-mortality risk under the right-lobe transplantation, our proposed mechanism performs significantly better than the maximum IR matching treatment based on this performance metric proposed in liver transplantation literature.

7 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 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.

References

Abdulkadiroğlu, A. and T. Sönmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Agarwal, N., E. Azevedo, I. Ashlagi, C. Featherstone, and Ömer Karaduman (2018): "Market failure in kidney exchange," Working paper.

Andersson, T. and J. Kratz (2019): "Pairwise Kidney Exchange over the Blood Group Barrier," *Review of Economic Studies*, forthcoming.

Barr, M. L., J. Belghiti, F. G. Villamil, E. A. Pomfret, D. S. Sutherland, R. W. Gruessner, A. N. Langnas, and F. L. Delmonico (2006): "A Report of the Vancouver Forum on the Care of the Live Organ Donor: Lung, Liver, Pancreas, and Intestine Data and Medical Guidelines," *Transplantation*, 81.

Budish, E. and E. Cantillon (2012): "The Multi-unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard," *American Economic Review*, 102, 2237–71.

Chan, S. C., C. M. Lo, B. H. Yong, W. J. C. Tsui, K. K. C. Ng, and S. T. Fan (2010): "Paired donor interchange to avoid ABO-incompatible living donor liver transplantation," *Liver Transplantation*, 16, 478–481.

Cronin, D. C., J. M. Millis, and M. Siegler (2001): "Transplantation of Liver Grafts from Living Donors into Adults - Too Much, Too Soon," *New England Journal of Medicine*, 344, 1633–1637.

Delacrétaz, D., S. D. Kominers, and A. Teytelboym (2017): "Refugee Resettlement," Working paper.

Dickerson, J. P. and T. Sandholm (2014): "Multi-Organ Exchange: The Whole is Greater than the Sum of its Parts," in *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, 220–229.

Echenique, F. and M. B. Yenmez (2015): "How to Control Controlled School Choice," *American Economic Review*, 105, 2679–2694.

Edelman, B., M. Ostrovsky, and M. Schwarz (2007): "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *American Economic Review*, 97, 242–259.

Edmonds, J. (1965): "Paths, Trees, and Flowers," *Canadian Journal of Mathematics*, 17, 449–467.

Ergin, H., T. Sönmez, and M. U. Ünver (2017): "Dual-donor organ exchange," *Econometrica*, 85, 1645–1671.

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

Goldberg, D. S., B. French, P. L. Abt, K. Olthoff, and A. Shaked (2014): "Superior survival using living donors and donor-recipient matching using a novel living donor risk index," *Hepatology*, 60, 1717–1726.

Grady, D. (2002): "Donor's Death at Hospital Halts Some Liver Surgeries," *New York Times*, January, 16.

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2013): "Effective affirmative action in school choice," *Theoretical Economics*, 8, 325–363.

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

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

Hwang, S., S.-G. Lee, D.-B. Moon, G.-W. Song, C.-S. Ahn, K.-H. Kim, T.-Y. Ha, D.-H. Jung, K.-W. Kim, N.-K. Choi, G.-C. Park, Y.-D. Yu, Y.-I. Choi, P.-J. Park, and H.-S. Ha (2010): "Exchange living donor liver transplantation to overcome ABO incompatibility in adult patients," *Liver Transplantation*, 16, 482–490.

Jones, W. and A. Teytelboym (2017): *Journal of Refugee Studies*, fex022.

Jung, D.-H., S. Hwang, C.-S. Ahn, K.-H. Kim, D.-B. Moon, T.-Y. Ha, G.-W. Song, G.-C. Park, and S.-G. Lee (2014): "Section 16. Update on Experience in Paired-Exchange Donors in Living Donor Liver Transplantation For Adult Patients at ASAN Medical Center," *Transplantation*, 97.

Kahn, A. B. (1962): "Topological sorting of large networks," *Communications of the ACM*, 5, 558–562.

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

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

Lee, S.-G. (2010): "Living-donor liver transplantation in adults," *British Medical Bulletin*, 94, 33–48.

Milgrom, P. (2000): "Putting Auction Theory to Work: The Simultaneous Ascending Auction," *Journal of Political Economy*, 108, 245–272.

Mishra, A., A. Lo, G. S. Lee, B. Samstein, P. S. Yoo, M. H. Levine, D. S. Goldberg, A. Shaked, K. M. Olthoff, and P. L. Abt (2018): "Liver paired exchange: Can the liver emulate the kidney?" *Liver Transplantation*, 24, 677–686.

Molinari, M., J. Matz, S. DeCoutere, K. El-Tawil, B. Abu-Wasel, and V. Keough (2014): "Live liver donors' risk thresholds: Risking a life to save a life," *HPB: The Official Journal of the International Hepato Pancreato Biliary Association*, 16, 560–574.

Moraga, J. F.-H. and H. Rapoport (2014): "Tradable immigration quotas," *Journal of Public Economics*, 115, 94–108.

Nicolò, A. and C. Rodriguez-Álvarez (2017): "Age-based preferences in paired kidney exchange," .

Okumura, Y. (2014): "Priority matchings revisited," *Games and Economic Behavior*, 88, 242–249.

Pomfret, E. A., J. P. A. Lodge, F. G. Villamil, and M. Siegler (2011): "Should We Use Living Donor Grafts for Patients With Hepatocellular Carcinoma? Ethical Considerations," *Liver Transplantation*, 17, S128–S132.

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

Roll, G. R., J. R. Parekh, W. F. Parker, M. Siegler, E. A. Pomfret, N. L. Ascher, and J. P. Roberts (2013): "Left Hepatectomy Versus Right Hepatectomy for Living Donor Liver Transplantation: Shifting the Risk From the Donor to the Recipient," *Liver Transplantation*, 19, 472–481.

Roth, A. E. and E. Peranson (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 748–780.

Roth, A. E., T. Sönmez, and M. U. Ünver (2004): "Kidney Exchange," *Quarterly Journal of Economics*, 119, 457–488.

——— (2005): "Pairwise Kidney Exchange," *Journal of Economic Theory*, 125, 151–188.

——— (2007): "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review*, 97, 828–851.

Samstein, B., I. de Melo-Martin, S. Kapur, L. Ratner, and J. Emond (2018): "A liver for a kidney: Ethics of trans-organ paired exchange," *American Journal of Transplantation*, 18, 1077–1082.

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

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

Schummer, J. and R. V. Vohra (2013): "Assignment of Arrival Slots," *AEJ: Microeconomics*, 5, 164–185.

Sönmez, T. and M. U. Ünver (2010): "Course Bidding at Business Schools," *International Economic Review*, 51, 99–123.

——— (2014): "Altruistically unbalanced kidney exchange," *Journal of Economic Theory*, 152, 105–129.

Sönmez, T., M. U. Ünver, and M. B. Yenmez (2018): "Incentivized Kidney Exchange," Working paper.

Um, E.-H., S. Hwang, G.-W. Song, D.-H. Jung, C.-S. Ahn, K.-H. Kim, D.-B. Moon, G.-C. Park, and S.-G. Lee (2015): "Calculation of standard liver volume in Korean adults with analysis of confounding variables," *Korean Journal of Hepatobiliary Pancreatic Surgery*, 19, 133–138.

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

Appendix A 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 $x = y$.[25]

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

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

[25]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).

for all $x \in \mathcal{V}$. We call $M(x)$ the **assignment 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 $\varepsilon = \{x, y\}$ 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 an undirected graph, in the definition of a digraph, we are ruling out loops, i.e., directed edges (x,y) such that $x = y$.[26]

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

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: $v_1 \to v_2 \to \ldots \to v_n \to v_1$.

The following lemma is a standard result in graph theory.[27]

Lemma 4 *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 1 in Subection 4.2.

Proof of Lemma 1: 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 $k \in \{0, 1, \ldots, n-1\}$ where all indexes are written in modulo n (i.e., $n \equiv 0$)

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

implies that $X_3^{(k)} \leq Y_{3\ell}^{(k+1)}$. It also implies that $Y_{3\ell}^{(k+1)} < X_3^{(k+2)}$ since $Y^{(k+1)\,\ell} \not\geq X^{(k+2)}$ and $Y^{(k+1)\,r} \geq X^{(k+2)}$. Therefore, $X_3^{(k)} < X_3^{(k+2)}$. 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:

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

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

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

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

Appendix B Proofs

Proof of Lemma 2:

Parts (1) *and* (2.b): For each i_k, we check whether $\mathcal{J}_{k-1} \cup \{i_k\}$ is matchable in G_{k-1} (recall that by construction of G_{k-1}, $\{i_k, j\} \in E_{k-1} \implies j \in \mathcal{E}^\ell(i_k)$). When the answer is affirmative, we include i_k in $\mathcal{J}_k$. Moreover, for all $m > k$, no right-lobe donating match of i_k is ever included in G_m and G_m is constructed from G_{m-1} making sure that $\mathcal{J}_k$ is still matchable. These imply Parts (1) and (2.b) when $k = K$.

If $\mathcal{J}_K = \emptyset$ then $\mathbf{M}_K = \mathbf{M}[G_K] \supseteq \{\varnothing\} \neq \emptyset$. If $\mathcal{J}_K \neq \emptyset$, then we showed that $\mathcal{J}_K$ is matchable in G_K by some matching $M' \in \mathbf{M}[G_K]$ by Part 1. Thus, $M' \in \mathbf{M}_K$. This shows in either, case $\mathbf{M}_K \neq \emptyset$. Suppose $M \in \mathbf{M}_K$ for the remaining parts.

Part (2.c): Suppose that there exists some $i_k \in \widetilde{\mathcal{J}}_K$ such that $M(i_k) \in \mathcal{E}^\ell(i_k)$. This and Part (2.b) imply that all of the pairs in $\mathcal{J}_{k-1} \cup \{i_k\}$ are matched in M by donating their left lobes. By construction, $\mathcal{J}_{k-1}$ is matchable in G_{k-1} and $\{i_k\} \cup \mathcal{J}_{k-1}$ is not matchable in G_{k-1}. Again by construction, for all $i \in \{i_{k+1}, \ldots, i_n\}$, $\{i, j\} \in E_{k-1} \implies j \in \mathcal{E}^\ell(i)$. Therefore, $M \notin \mathbf{M}_{k-1}$ and there is some $i \in \mathcal{J}_{k-1} \cup \{i_k\}$ such that $i \in \mathcal{E}^r(M(i))$ and $M(i) \in \{i_{k+1}, \ldots, i_n\}$. Hence, $i \; \Pi_\ell \; M(i)$. We also have $M(i) \in \mathcal{E}^\ell(i)$ as established above. Thus, by definition of the precedence digraph, $\tau(M(i)) \to \tau(i)$. By construction of the topological order, $M(i) \; \Pi_\ell \; i$, which is a contradiction to $i \; \Pi_\ell \; M(i)$.

Part (2.d): Let $i_k \in \mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K]$. Thus, $\{i_k, j\} \in E_K \implies j \in \mathcal{E}^\ell(i_k)$. Suppose $M(i_k) \neq \emptyset$. Then $M(i_k) \in \mathcal{E}^\ell(i_k)$. This and Part (2.b) imply that all of the pairs in $\mathcal{J}_{k-1} \cup \{i_k\}$ are matched in M by donating their left lobes. By construction, $\mathcal{J}_{k-1}$ is matchable in G_{k-1} and $\{i_k\} \cup \mathcal{J}_{k-1}$ is not matchable in G_{k-1}. Thus, $M \notin \mathbf{M}_{k-1}$. Again by construction, for all $i \in \{i_{k+1}, \ldots, i_n\}$, $\{i, j\} \in E_{k-1} \implies j \in \mathcal{E}^\ell(i)$. Therefore, there is some $i \in \mathcal{J}_{k-1} \cup \{i_k\}$ such that $i \in \mathcal{E}^r(M(i))$ and $M(i) \in \{i_{k+1}, \ldots, i_n\}$. Hence, $i \; \Pi_\ell \; M(i)$. We have $M(i) \in \mathcal{E}^\ell(i)$ as established above. Thus, by definition of the precedence digraph, $\tau(M(i)) \to \tau(i)$. By construction of the topological order, $M(i) \; \Pi_\ell \; i$, which is a contradiction to $i \; \Pi_\ell \; M(i)$.

Parts (2.a) *and* (2.e): We prove the following claim to prove these parts:

Claim: For all indices k, all indices $k' \geq k$, and all pairs $i \in \mathcal{J}_k$, the following hold for the induced match sets of pair i at Step 1.(k) and Step 1.(k') reduced compatibility graphs:

1. $E_{k'}(i) \subseteq E_k(i)$, and
2. $\{i, j\} \in E_k \implies$ for all $M' \in \mathbf{M}_{k'}$, $M'(i) \; I_i \; j$.

Proof of Claim:

1. Suppose to the contrary, there exists some $\{i,j\} \in E_{k'} \setminus E_k$. Therefore, j is processed and, in particular, transformed after Step 1.(k). Since $i \in \mathcal{J}_k$, i is not transformed and thus, $j \in \mathcal{E}^\ell(i)$ and $i \in \mathcal{E}^r(j)$ implying that $\tau(j) \to \tau(i)$ in the precedence digraph (by definition). This, in turn, implies $j \; \Pi_\ell \; i$. But this is a contradiction that j is processed after Step 1.(k) while i is processed before or at Step 1.(k).
2. Let $\{i,j\} \in E_k$. Since $i = i_m$ for some index $m \leq k$, by the first part of the claim $\{i,j\} \in E_k \implies \{i,j\} \in E_m$ since $m \leq k$. We delete from E_{m-1} all matches of i but its best achievable matches (while all pairs in $\mathcal{J}_{m-1}$ can simultaneously be matched), i.e., $j \in \mathcal{B}(i|\mathcal{J}_{m-1}, G_{m-1})$. Hence, i is indifferent among all matchings that match it in G_k. Since no new matches of i are added to $E_{k+1}, \ldots, E_{k'}$ by first part of the Claim, for all $M' \in \mathbf{M}_{k'} = \big\{M'' \in \mathbf{M}[G_{k'}] \;:\; M''(j'') \neq \emptyset \;\; \forall j'' \in \mathcal{J}_{k'}\big\}$, $M'(i) \; I_i \; j$. $\diamond$

Pick $i \in \mathcal{J}_K$ and $M' \in \mathbf{M}_K$. Then $M(i) \; I_i \; M'(i)$ by the Claim's second statement. Moreover, by Part (2.d) for all $j \in \mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K]$, $M(j) = M'(j) = \emptyset$. These prove Part (2.a).

Now $i = i_k$ for some k. Since $M \in \mathbf{M}_K$, $M(i_k) \in E_K(i_k) \subseteq E_k(i_k)$ by the Claim's first statement. Since $E_k(i_k) = \big\{\{i_k, j\} \;:\; j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_{k-1})\big\}$ (by definition of graph G_k) and since $E_K(i_k) \subseteq E_k(i_k)$, we have $E_K(i_k) = \big\{\{i_k, j\} \;:\; j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_K)\big\}$, this in turn implies $M(i_k) \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_K)$ and $M(i_k) \; I_{i_k} \; j$ for all $j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_k)$. This proves Part (2.e) and completes the proof of the lemma. ∎

Proof of Lemma 3:

Part (1): For all $n = 1, ..., N$, $\mathbf{M}^*_n \subseteq \mathbf{M}_K$ follows from the facts that $G^*_0 = G_K$ and the match sets satisfy $E^*_N \subseteq \ldots \; E^*_0 = E_K$; moreover, $\mathcal{J}^*_n \subseteq \widetilde{\mathcal{J}}_K$ follows from the definition of Step 2. Thus, Part (1) is proven when $n = N$.

Part (2): $\mathcal{J}_K \cup \mathcal{J}^*_n$ is matchable in G^*_n follows from the definition of Step 2. Thus, Part (2) follows for $n = N$.

If $\mathcal{J}_K \cup \mathcal{J}^*_N = \emptyset$, then $\mathbf{M}^*_N = \mathbf{M}[G^*_N] \supseteq \{\varnothing\} \neq \emptyset$. If $\mathcal{J}_K \cup \mathcal{J}^*_N \neq \emptyset$, there exists some $M' \in M^{[}G^*_N]$ such that M' matches all pairs in $\mathcal{J}_K \cup \mathcal{J}^*_N$ as we showed in Part 1. Thus, in either case, $M^*_N \neq \emptyset$.

Let $M \in \mathbf{M}^*_N$ for the rest of the proof.

Parts (3.*b*)*,* (3.*c*)*, and* (3.*e*): $\mathbf{M}^*_N \subseteq \mathbf{M}_K$ and Lemma 2 Parts (2.b), (2,c), and (2.e) imply Parts (3.b), (3.c), and (3.e), respectively.

Part (3.*d*): Suppose contrary to the claim there exists $j \in \mathcal{I} \setminus [\mathcal{J}_K \cup \mathcal{J}^*_N]$ such that $M(j) \neq \emptyset$. By Lemma 2 Part (2.d), $j \in \widetilde{\mathcal{J}}_N \setminus \mathcal{J}^*_N$. Thus, $j = i^*_n$ for some $n \leq N$. Since $i^*_n \notin \mathcal{J}^*_n \subseteq \mathcal{J}^*_N$, in substep 2.$(n)$, $\mathcal{J}_K \cup \mathcal{J}^*_{n-1} \cup \{i^*_n\}$ is not matchable in G^*_{n-1}. Thus, no matching in $\mathbf{M}^*_{n-1}$ matches i^*_n. This contradicts $M(i^*_n) \neq \emptyset$ because $M \in \mathbf{M}^*_N \subseteq \mathbf{M}^*_{n-1}$.

Part $(3.f)$: Let $i_n^* \in \mathcal{J}_N^*$. By construction in Step 2.(n),

$$E_n^* = \Big[E_{n-1}^* \setminus E_{n-1}^*(i_n^*)\Big] \cup \Big\{\{i_n^*, j\} \;:\; j \in \mathcal{B}(i_n^* | \mathcal{J}_K \cup \mathcal{J}_{n-1}^*, G_{n-1}^*)\Big\}.$$

That is, while we are obtaining G_n^*, we delete all edges involving i_n^* in G_{n-1}^* except those would match it to one of its best assignments in G_{n-1}^* given that all pairs in $\mathcal{J}_K \cup \mathcal{J}_{n-1}^*$ are simultaneously matched. Since $E_n^* \subseteq E_N^*$ and all pairs in $\mathcal{J}_K \cup \mathcal{J}_N^*$ are matchable in G_N^*, we have $M(i_n^*) \in \mathcal{B}(i_n^* | \mathcal{J}_K \cup \mathcal{J}_{n-1}^*, G_N^*)$. Since no new edges are added to the graph through the substeps of Step 2, we also have $M(i_n^*) \; I_{i_n^*} \; j$ for all $j \in \mathcal{B}(i_n^* | \mathcal{J}_K \cup \mathcal{J}_{n-1}^*, G_n^*)$.

Part $(3.a)$: For all $j \in \mathcal{J}_K$, the statement holds by Part (3.e). For all $j \in \mathcal{I} \setminus [\mathcal{J}_K \cup \mathcal{J}_N^*]$, the statement holds by Part (3.d). For all $j \in \mathcal{J}_N^*$, the statement holds by Part (3.f).

■

We prove Theorem 1 in three parts for each property in Lemmas 5, 6, and 7. Let $f^{\mathbf{P}}$ refer to the precedence-induced adaptive-priority mechanism.

Lemma 5 (IR) *Mechanism $f^{\mathbf{P}}$ is individually rational.*

Proof of Lemma 5 (IR): In every step of the algorithm, the active reduced compatibility graphs are subgraphs of the IR compatibility graph given the submitted preference profile R. Since $f^{\mathbf{P}}$ chooses a matching of the final graph of the algorithm G_N^*, it is individually rational. ■

Lemma 6 (PE) *Mechanism $f^{\mathbf{P}}$ is Pareto efficient.*

Proof of Lemma 6 (PE): Fix $R \in \mathbf{R}$. Recall that $G_{IR} = (\mathcal{I}, E_{IR})$ is the individually rational compatibility graph of the problem induced by R.

Let $M \equiv f^{\mathbf{P}}(R)$. Suppose $M' \in \mathbf{M}_c$ satisfies $M'(i) \; R_i \; M(i)$ for all $i \in \mathcal{I}$. We will show that $M'(i) \; I_i \; M(i)$ for all $i \in \mathcal{I}$ to prove Pareto efficiency of $f^{\mathbf{P}}(R)$. Since $M'(i) \; R_i \; M(i)$ for all $i \in \mathcal{I}$, and M is individually rational by Lemma 5, we obtain that M' is individually rational, as well.

We consider three separate cases for pairs in $\mathcal{J}_K$, $\mathcal{J}_N^*$, and $\mathcal{I} \setminus [\mathcal{J}_K \cup \mathcal{J}_N^*]$.

1. $\mathcal{J}_K$: By induction we prove that (i) for all Steps 1.(k) and all $i \in \mathcal{J}_k$, $M'(i) \; I_i \; M(i)$ and (ii) $M' \in \mathbf{M}_K$.
   Fix $k \leq K$. As the inductive assumption, assume that for $k' = k-1$, the following holds:

$$\text{for all } i_m \in \mathcal{J}_{k'}, \;\; M(i_m) \; I_{i_m} \; M'(i_m) \;\text{ and }\; M'(i_m) \in \mathcal{B}(i_m | \mathcal{J}_{m-1}, G_{k'}).$$

We will prove the same holds for $k' = k$. Two cases hold for i_k: Either $i_k \notin \mathcal{J}_k$ or $i_k \in \mathcal{J}_k$:

- First, assume $i_k \notin \mathcal{J}_k$. Thus, $\mathcal{J}_k = \mathcal{J}_{k-1}$. Hence, the inductive assumption for $k-1$ implies for all $i_m \in \mathcal{J}_k$, $M(i_m)\ I_{i_m}\ M'(i_m)$.
  - If i_k is not willing, then by definition of Step 1.(k), $G_k = G_{k-1}$, and hence, the inductive assumption for $k-1$ implies $M'(i_m) \in \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_k)$.
  - If i_k is willing, then by definition of Step 1.(k) graph $G_k = (\mathcal{I}, E_k)$ satisfies:

    $$E_k = E_{k-1} \cup \Big\{\{i_k, j\} \in E_{IR}(i_k)\ :\ t(i_k, j) = r\Big\}.$$

    Fix $i_m \in \mathcal{J}_{k-1}$. Since $i_m\ \Pi_\ell\ i_k$, by the definition of the precedence graph and topological order, $t(i_k, i_m) \neq r$. Thus,

    $$\{i_m, i_k\} \in E_{k-1} \iff \{i_m, i_k\} \in E_k.$$

    Thus, $\mathcal{B}(i_m|\mathcal{J}_{m-1}, G_{k-1}) = \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_k)$. Hence, by the inductive assumption, we still have $M'(i_m) \in \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_k)$.
- Next, assume $i_k \in \mathcal{J}_k$. Then $\mathcal{J}_k = \mathcal{J}_{k-1} \cup \{i_k\}$. By the definition of Step 1.(k), graph $G_k = (\mathcal{I}, E_k)$ is obtained from graph $G_{k-1} = (\mathcal{I}, E_{k-1})$ as follows:

  $$E_k = \Big[E_{k-1} \setminus E_{k-1}(i_k)\Big] \cup \Big\{\{i_k, j\}\ :\ j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_{k-1})\Big\}. \tag{1}$$

  We first prove the inductive statement for i_k and then for pairs in $\mathcal{J}_{k-1} = \mathcal{J}_k \setminus \{i_k\}$:
  - Since $i_k \in \mathcal{J}_K$, $M(i_k) \neq \emptyset$. Since by assumption $M'(i_k)\ R_{i_k}\ M(i_k)$, we have $M'(i_k) \neq \emptyset$. Moreover, $M'(i_k) \in \mathcal{E}^\ell(i_k)$, as $M(i_k) \in \mathcal{E}^\ell(i_k)$ and i_k prefers donating left lobe to donating right lobe under any match. Suppose $i_m \equiv M'(i_k)$ for some m. Two subcases exist: $m > k$ or $m \leq k$:
    - Suppose $m > k$, then $i_k \in \mathcal{E}^\ell(i_m)$ by the fact that $i_k\ \Pi_\ell\ i_m$ and the definition of the precedence graph and topological order (as otherwise, $i_k \in \mathcal{E}^r(i_m)$ and $i_m \in \mathcal{E}^\ell(i_k)$ implies that $\tau(i_m) \to \tau(i_k)$, and thus, $i_m\ \Pi_\ell\ i_k$, a contradiction). As $i_m \in \mathcal{E}^\ell(i_k)$ and by the inductive assumption $\{i, M'(i)\} \in E_{k-1}$ for all $i \in \mathcal{J}_{k-1}$, we also have $\{i_k, i_m\} \in E_{k-1}$.
    - Suppose $m \leq k$. Two further subcases: $i_m \in \mathcal{J}_k$ or $i_m \notin \mathcal{J}_k$.
      (a) Suppose $i_m \in \mathcal{J}_k$. Thus, $\{i_m, \underbrace{M'(i_m)}_{=i_k}\} \in E_{k-1}$ by the inductive assumption.
      (b) Suppose $i_m \notin \mathcal{J}_k$. Then $i_m \notin \mathcal{J}_m$, either. However, the inductive assumption for $k-1$ implies that $M' \in \mathbf{M}_{m-1}$ and M' is a matching that matches all pairs $\mathcal{J}_{m-1} \cup \{i_m\}$ in the graph G_{m-1}, contradicting $i_m \notin \mathcal{J}_m$. Thus, this subcase cannot hold.

    We showed that in either case $\{i_k, i_m\} \in E_{k-1}$. Since $M(i_k)\ I_{i_k}\ j$ for all $j \in$

$\mathcal{B}(i_k|\mathcal{J}_{k-1}, G_{k-1})$ by Lemma 2 Part (2.e) and $M'(i_k)\ R_{i_k}\ M(i_k)$ by assumption, we have $\{i_k, i_m\} \in E_k(i_k) = \big\{\{i_k, j\} : j \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_{k-1})\big\}$, where the equality follows from Equation 1. Thus, $i_m = M'(i_k) \in \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_k) = \mathcal{B}(i_k|\mathcal{J}_{k-1}, G_{k-1})$ and $M'(i_k)\ I_{i_k}\ M(i_k)$.

  – Finally, consider any $i_m \in \mathcal{J}_{k-1} = \mathcal{J}_k \setminus \{i_k\}$. The inductive assumption for $k-1$ implies $M'(i_m)\ I_{i_m}\ M(i_m)$.

    We finish this case by showing $M'(i_m) \in \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_k)$. The inductive assumption for $k-1$ implies $M'(i_m) \in \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_{k-1})$.

    Two subcases exist: $M'(i_m) \neq i_k$ or $M'(i_m) = i_k$.

    * In the first subcase (i.e., $M'(i_m) \neq i_k$): The definition of E_k, which is restated in Equation 1, implies $E_k \subseteq E_{k-1}$ and $\{i_m, M'(i_m)\} \in E_k$. Thus, $M'(i_m) \in \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_k)$.
    * In the second subcase (i.e., $M'(i_m) = i_k$): Through Subcase (a) of $m \leq k$ above, we showed that $\{i_m, i_k\} \in E_k$. The definition of E_k, which is restated in Equation 1, implies $E_k \subseteq E_{k-1}$. Thus, $M'(i_m) \in \mathcal{B}(i_m|\mathcal{J}_{m-1}, G_k)$.

2. $\mathcal{J}^*_N$: By induction, we prove that for all Steps 2.(n) and all $i \in \mathcal{J}^*_n$, $M'(i)\ I_i\ M(i)$. Fix $n \leq N$. As the inductive assumption, assume that for $n' = n-1$, the following holds:

$$\text{for all } i^*_m \in \mathcal{J}^*_{n'},\ \ M(i^*_m)\ I_{i^*_m}\ M'(i^*_m) \ \text{ and } \ M'(i^*_m) \in \mathcal{B}(i^*_m|\mathcal{J}_K \cup \mathcal{J}^*_{m-1}, G^*_{n'}),\ \text{and}$$

$$\text{for all } i \in \mathcal{J}_K,\ \ M(i)\ I_i\ M'(i) \ \text{ and } \ M'(i) \in \mathcal{B}(i|\mathcal{J}_K, G^*_{n'}).$$

We will prove the same holds for $n' = n$. Two cases hold for i^*_n: Either $i^*_n \notin \mathcal{J}^*_n$ or $i^*_n \in \mathcal{J}^*_n$:

- First, assume $i^*_n \notin \mathcal{J}^*_n$. Thus, $\mathcal{J}^*_n = \mathcal{J}^*_{n-1}$ and $G^*_n = G^*_{n-1}$. Hence, the inductive assumption for $n-1$ implies the same holds for n.
- Next, assume $i^*_n \in \mathcal{J}^*_n$. Thus, $\mathcal{J}^*_n = \mathcal{J}^*_{n-1} \cup \{i^*_n\}$. Recall that by the definition of Step 2.(n), subgraph $G^*_n = (\mathcal{I}, E^*_n)$ is obtained from graph $G^*_{n-1} = (\mathcal{I}, E^*_{n-1})$ as follows:

$$E^*_n = \Big[E^*_{n-1} \setminus E^*_{n-1}(i^*_n)\Big] \cup \Big\{\{i^*_n, j\}\ :\ j \in \mathcal{B}(i^*_n|\mathcal{J}_K \cup \mathcal{J}^*_{n-1}, G^*_{n-1})\Big\}. \tag{2}$$

  We first prove the inductive statement for i^*_n, then for pairs in $\mathcal{J}_K \cup \mathcal{J}^*_{n-1}$:

  – Since $i^*_n \in \mathcal{J}^*_N$, $M(i^*_n) \neq \emptyset$. Since by assumption $M'(i^*_n)\ R_{i^*_n}\ M(i^*_n)$, we have $M'(i^*_n) \neq \emptyset$. The proof of Part 1 for $\mathcal{J}_K$ above and the inductive assumption for $\mathcal{J}^*_{n-1}$ imply that for all $i \in \mathcal{J}_K \cup \mathcal{J}^*_{k-1}$, $M'(i)\ I_i\ M(i)$ and $\{i, M'(i)\} \in E^*_{n-1}$. Thus, $M'(i^*_n)\ R_{i^*_n}\ M(i^*_n)$ also implies that and $\{i^*_n, M'(i^*_n)\} \in E^*_n(i^*_n) = \big\{\{i^*_n, j\} : j \in \mathcal{B}(i^*_n|\mathcal{J}_K \cup \mathcal{J}^*_{n-1}, G^*_{n-1})\big\}$, where equality follows from Equation 2. Since $M(i) \in \mathcal{B}(i^*_n|\mathcal{J}_K \cup \mathcal{J}^*_{n-1}, G^*_{n-1})$ by Step 2.(n), we also have $M'(i^*_n)\ I_{i^*_n}\ M(i^*_n)$. Equation 2 also implies $M'(i^*_n) \in \mathcal{B}(i^*_n|\mathcal{J}_K \cup \mathcal{J}^*_{n-1}, G^*_n)$.
  – Next, consider $i \in \mathcal{J}_K \cup \mathcal{J}^*_{n-1}$. The inductive assumption (if $i \in \mathcal{J}^*_{n-1}$) and the

Part 1 for $\mathcal{J}_K$ above (if $i \in \mathcal{J}_K$) imply that $M'(i)\ I_i\ M(i)$ and $\{i, M'(i)\} \in E^*_{n-1}$. If $M'(i) = i^*_n$ then the part for i^*_n (the above paragraph) implies $\{i, i^*_n\} \in E^*_n$. If $M'(i) \neq i^*_n$, then Equation 2 implies $\{i, M'(i)\} \in E^*_n$. Thus, if $i = i^*_m \in \mathcal{J}^*_{n-1}$ for some m, then the inductive assumption implies that $M'(i^*_m) \in \mathcal{B}(i^*_m|\mathcal{J}_{m-1}, G^*_n)$, and if $i = i_k \in \mathcal{J}_K$ for some k, then the part for $\mathcal{J}_K$ above implies $M'(i_k) \in \mathcal{B}(i_k|\mathcal{J}_K, G^*_n)$.

3. $\mathcal{I} \setminus [\mathcal{J}_K \cup \mathcal{J}^*_N]$: Part 2 for $\mathcal{J}^*_N$ also establishes that $M' \in \mathbf{M}^*_N$. Lemma 3 Part (2.d) implies for both M and M',

$$M'(i) = M(i) = \emptyset \quad \text{for all } i \in \mathcal{I} \setminus [\mathcal{J}_K \cup \mathcal{J}^*_N]$$

finishing the induction and showing that $M'(i)\ I_i\ M(i)$ for all $i \in \mathcal{I}$, and hence, $M = f^{\mathbf{P}}(R)$ is Pareto efficient.

■

Lemma 7 (IC) *Mechanism $f^{\mathbf{P}}$ is incentive compatible.*

Proof of Lemma 7 (IC): Fix $R \in \mathbf{R}$. Let $M \equiv f^{\mathbf{P}}(R)$. Consider the algorithm executed to find M under R, and let $\mathcal{J}_K$ and $\widetilde{\mathcal{J}}_K$ be the corresponding sets of pairs determined in Step 1.

Consider pair $i \in \mathcal{I}$. Let its preference relation be denoted as $R_i^{a/v} \equiv R_i$ for some bias type $a \in \{d, e\}$ and for some willingness type $v \in \{u, w\}$. Three mutually exclusive cases are possible: $i \in \mathcal{J}_K$, $i \in \widetilde{\mathcal{J}}_K$, and $i \in \mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K]$:

1. If $i \in \mathcal{J}_K$: Then $M(i) \in \mathcal{E}^\ell(i)$. Since it is never transformed,

$$M(i)\ I_i\ f^{\mathbf{P}}(R_i^{a/x}, R_{-i})(i) \ \text{ for } x \ \in \{u, w\} \setminus \{v\}. \tag{3}$$

There are two subcases for its bias type a:

- If $a = d$, i.e., it is direct-transplant type: If it is also a left-lobe compatible pair, then $M(i) = i$ and this is its first choice. Thus, it cannot benefit by misreporting. On the other hand, if it is not left-lobe compatible then $R_i = R_i^{d/v} = R_i^{e/v}$. Thus, $M(i)\ I_i\ f^{\mathbf{P}}(R_i^{b/x}, R_{-i})(i)$ for any bias type $b \in \{d, e\}$ and willingness type $x \in \{u, w\}$ by previous statement and Equation 3.
- if $a = e$, i.e., it is exchange type: By individual rationality of $f^{\mathbf{P}}$, we have

$$M(i)\ R_i \left\{ \begin{array}{cc} i & \text{if } i \text{ is left-lobe compatible} \\ M(i) & \text{if } i \text{ is left-lobe incompatible} \end{array} \right\} I_i\ f(R_i^{d/x}, R_{-i})(i) \quad \text{for all } x \in \{u, w\}.$$

This together with Equation 3 establishes that i cannot benefit from misreporting.

2. If $i \in \widetilde{\mathcal{J}}_K$: Then $M(i) = \emptyset$ or $M(i) \in \mathcal{E}^r(i)$. Moreover, by individual rationality of M, i is not a left-lobe-compatible pair and $R_i = R_i^{a/v} \in \{R_i^{d/w}, R_i^{e/w}\}$.
   Let $i_k \equiv i$ and was transformed in Step 1.(k) for some k. It was not matchable by left-lobe donation in addition to pairs in $\mathcal{J}_{k-1}$ in G_{k-1}. Thus, reporting $R_i^{d/u}$ (or $R_i^{e/u}$, which has the same individually rational portion as $R_i^{d/u}$, because i is unwilling and left-lobe incompatible under both) instead of R_i will not change the fact that i is not matchable by left-lobe donation in addition to pairs in $\mathcal{J}_{k-1}$ in G_{k-1}, as the same graph will occur under both revelations of preferences (as it is not left-lobe compatible, the individually rational options of i in which it donates a left lobe are the same under all preferences). Thus, $M(i)\ R_i\ \emptyset = f^{\mathbf{P}}(R_i^{d/u}, R_{-i})(i)$. Finally consider the remaining manipulation possibility by revealing $R_i^{b/x} \in \{R_i^{d/w}, R_i^{e/w}\} \setminus \{R_i\}$:
   - If $R_i = R_i^{d/w}$, then the remaining manipulation is $R_i^{b/x} = R_i^{e/w}$. If i is not right-lobe-only compatible then $R_i = R_i^{b/x}$, so we are done. On the other hand, if i is right-lobe-only compatible, then $M(i) = i$ by individual rationality. Moreover, $M(i) = i\ R_i\ j$ for all $j \in \mathcal{E}^r(i)$ and $M(i) = i\ P_i\ \emptyset$ by individual rationality again. Since $f^{\mathbf{P}}(R_i^{b/x}, R_{-i})(i) \notin \mathcal{E}^\ell(i)$, we obtain $M(i)\ R_i\ f^{\mathbf{P}}(R_i^{b/x}, R_{-i})(i)$.
   - If $R_i = R_i^{e/w}$, then the remaining manipulation is $R_i^{b/x} = R_i^{d/w}$. If i is not right-lobe-only compatible then $R_i = R_i^{b/x}$, so we are done. On the other hand, if i is right-lobe-only compatible $M(i)\ R_i\ i = f^{\mathbf{P}}(R_i^{b/x}, R_{-i})(i)$ by individual rationality of $f^{\mathbf{P}}$.
3. If $i \in \mathcal{I} \setminus [\mathcal{J}_K \cup \widetilde{\mathcal{J}}_K]$: Then $M(i) = \emptyset$ and it is unwilling and left-lobe incompatible, i.e., $R_i = R_i^{d/u} = R_i^{e/u}$ and $\emptyset\ P_i\ j$ for all $j \in \mathcal{E}^r(i)$. Suppose $i_k \equiv i$ for some k and thus, i is not left-lobe matchable in addition to $\mathcal{J}_{k-1}$ in G_{k-1} (as otherwise $i \in \mathcal{J}_k$, a contradiction). When it announces $R_i^{b/x}$, the same graph G_{k-1} occurs at the end of Step 1.$(k-1)$. That is because, as it is not left-lobe compatible, its individually rational left-lobe donation options are the same under all preferences available to it. Thus, it is still not matchable in addition to $\mathcal{J}_{k-1}$ and $f^{\mathbf{P}}(R_i^{b/x}, R_{-i})(i) \notin \mathcal{E}^\ell(i)$, implying $M(i) = \emptyset\ R_i\ f^{\mathbf{P}}(R_i^{b/x}, R_{-i})(i)$ as R_i is an unwilling preference relation.

■

Appendix C Additional Results

C.1 Impossibilities

Proposition 2 *Consider an exchange pool $(\mathcal{I}, \tau)$ with $\mathcal{I} = \{i_1 \ldots, i_K\}$ in which the underlying precedence digraph $(\mathbf{T} \times \mathbf{T}^{\mathbf{D}}, D^\tau)$ is a cycle $\tau(i_1) \to \tau(i_2) \to \cdots \to \tau(i_K) \to \tau(i_1)$ for $|\mathcal{I}| = K \geq 3$ such that for all k and all $n \notin \{k-1, k+1\}$ in modulo K, $i_n \notin \mathcal{E}(i_k)$. There exists no individually rational, Pareto-efficient, and incentive-compatible mechanism for this* *exchange pool.*

Proof of Proposition 2: Let f be an individually rational, Pareto-efficient, and incentive-compatible mechanism for this pool. We will show that this will lead to a contradiction. In the proof, all indices are meant in modulo K (i.e., $i_K \equiv i_0$).

Let

$$R^{(K+1)} \equiv (R^{e/w}_{i_1}, R^{e/w}_{i_2}, \dots, R^{e/w}_{i_K})$$

be the preference profile in which all pairs are willing (and exchange type[28]). Since f is Pareto efficient and individually rational, there exists some $\{i_k, i_{k+1}\} \in f(R^{(K+1)})$. Without loss of generality, subject to reindexing of the pairs

- if K is odd, suppose $\{i_{K-1}, i_K\} \in f(R^{(K+1)})$, and
- if K is even, suppose $\{i_K, i_1\} \in f(R^{(K+1)})$.

Define for any $k \in \{1, 2, \dots, K\}$,

$$R^{(k)} \equiv (R^{e/w}_{\{i_1,i_2,\dots,i_{k-1}\}}, R^{e/u}_{\{i_k,i_{k+1},\dots,i_K\}}).$$

We prove the following claim:

Claim: For all $k = K, K-1, \dots, 3$,

- if k is odd, $\{i_{k-1}, i_k\} \in f(R^{(k)})$, and
- if k is even, $\{i_{k-2}, i_{k-1}\} \in f(R^{(k)})$

Proof of Claim: We prove the claim by induction on decreasing k. Fix $k \in \{1, \dots, K\}$. As the <u>inductive assumption</u>, suppose the Claim is true for $k+1$ if $k < K$. We will prove it also holds for k (the initial step will be handled for $k = K$ below).

Consider the preference profile $R^{(k)}$ as defined above. It satisfies

$$R^{(k)} = (R^{e/u}_{i_k}, R^{(k+1)}_{-i_k}).$$

Two cases for k:

<u>k is odd</u> If $k \neq K$, by the inductive assumption for $k+1$ (which is even), and if $k = K$, by the labeling and choice of i_K, we have $\{i_{k-1}, i_k\} \in R^{(k+1)}$. Observe that $t(i_k, i_{k-1}) = \ell$ by the fact that $\tau(i_{k-1}) \to \tau(i_k)$. Moreover, $t(i_k, i_{k+1}) = r$ as $\tau(i_k) \to \tau(i_{k+1})$. Thus, by incentive compatibility of f for i_k, we still have $\{i_{k-1}, i_k\} \in f(R^{(k)})$.

<u>k is even</u> If $k \neq K$, by the inductive assumption for $k+1$ (which is odd), and if $k = K$, by the labeling and choice of i_K, we have $\{i_k, i_{k+1}\} \in f(R^{(k+1)})$. Since $\tau(i_k) \to \tau(i_{k+1})$, we have $t(i_k, i_{k+1}) = r$. By reporting $R^{e/u}_{i_k}$ instead of $R^{e/w}_{i_k}$, the match $\{i_k, i_{k+1}\}$ becomes

[28] It does not matter whether they are exchange type or direct-transplant type as they have the same preferences as each pair is incompatible

individually irrational, and hence, $\{i_k, i_{k+1}\} \notin f(R^{(k)})$ by individual rationality of f. Moreover, by incentive compatibility of f for i_k, it should not be able to get a match by donating left lobe, i.e., $\{i_{k-1}, i_k\} \notin f(R^{(k)})$.
We claim that $\{i_{k-2}, i_{k-1}\} \in f(R^{(k)})$. Suppose not. Since

$$E_{IR}[R^{(k)}](i_{k-1}) = \big\{\{i_{k-2}, i_{k-1}\}, \{i_{k-1}, i_k\}\big\}$$

is the set of individually rational matches for pair i_{k-1}, then i_{k-1} is unmatched in $f(R^{(k)})$. Similarly i_k is unmatched in $f(R^{(k)})$ since

$$E_{IR}[R^{(k)}](i_k) = \big\{\{i_{k-1}, i_k\}\big\}$$

is the set of individual matches for pair i_k. Then the following is a matching,

$$f(R^{(k)}) \cup \{i_{k-1}, i_k\},$$

and it Pareto dominates $f(R^{(k)})$ under $R^{(k)}$ contradicting f's Pareto efficiency. Thus, $\{i_{k-2}, i_{k-1}\} \in f(R^{(k)})$, completing the induction. $\diamond$

By the Claim, we are left with the following preference profile and chosen match (as $k = 3$, the last step index of the induction, is odd):

$$R^{(3)} = (R^{e/w}_{\{i_1,i_2\}}, R^{e/u}_{\{i_3,\ldots,i_K\}}) \text{ and}$$

$$\{i_2, i_3\} \in f(R^{(3)}).$$

As $E_{IR}[R^{(3)}] = \big\{\{i_1, i_2\}, \{i_2, i_3\}\big\}$ is the set of individually rational matches and f is individually rational, we have $f(R^{(3)}) = \big\{\{i_2, i_3\}\big\}$.

Consider the preference profile $R^{(2)} = (R^{e/u}_{i_2}, R^{(3)}_{-i_2})$. We have $E_{IR}[R^{(2)}] = \big\{\{i_1, i_2\}\big\}$. Thus, by individual rationality and Pareto efficiency of f, $f(R^{(2)}) = \big\{\{i_1, i_2\}\big\}$. Since $\tau(i_1) \to \tau(i_2)$, $t(i_2, i_1) = \ell$. On the other hand, since $\tau(i_2) \to \tau(i_3)$, $t(i_2, i_3) = r$. Thus, pair i_2 benefits from reporting its type e/u instead of e/w, contradicting the incentive compatibility of f.

■

Example 4 *In this example we show that, if a pair's willingness to donate a right lobe is allowed to be contingent on the specific compatible liver lobe its patient receives, then a Pareto efficient, individually rational, and incentive compatible mechanism may not exist.*

Consider a liver-exchange pool with four incompatible pairs $\mathcal{I} = \{i_1, i_2, i_3, i_4\}$ *with the* *following types:*

$$\tau_P(i_1) = \tau_P(i_3) = (1,0,1) \qquad\qquad \tau_D(i_1) = \tau_D(i_3) = (0,1,0,1)$$
$$\tau_P(i_2) = \tau_P(i_4) = (0,1,1) \qquad\qquad \tau_D(i_2) = \tau_D(i_4) = (1,0,0,1)$$

The set of mutually compatible exchanges are given as

$$E_c = \big\{\{i_1, i_2\}, \{i_2, i_3\}, \{i_3, i_4\}, \{i_4, i_1\}\big\}.$$

Observe that, since the left lobe of each donor is too small for any patient, each donor donates his right lobe under each of these exchanges.

The public information received-graft preference relation over the set of compatible grafts is given as follows for each pair:

$$i_2 \succ_{i_1} i_4$$
$$i_3 \succ_{i_2} i_1$$
$$i_4 \succ_{i_3} i_2$$
$$i_1 \succ_{i_4} i_3$$

Suppose that each pair is willing to donate a right lobe regardless of which graft its patient receives, and thus the preference profile R is given as follows:

$$i_2 \; P_{i_1} \; i_4 \; P_{i_1} \; \emptyset$$
$$i_3 \; P_{i_2} \; i_1 \; P_{i_2} \; \emptyset$$
$$i_4 \; P_{i_3} \; i_2 \; P_{i_3} \; \emptyset$$
$$i_1 \; P_{i_4} \; i_3 \; P_{i_4} \; \emptyset$$

The mutual compatibility graph is depicted in Figure 9.

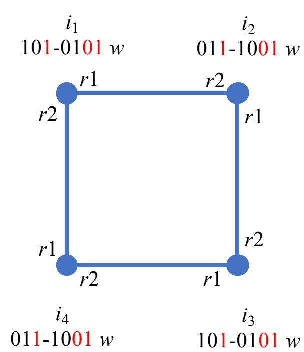

Figure 9: The mutual compatibility graph for Example 4. The right-lobe donations are denoted by letter r and preferences are denoted by numbers $1, 2$ next to the donated lobe for each exchange.

Suppose f is a Pareto efficient, individually rational, and incentive compatible mechanism. By Pareto efficiency of f, there exists some $\{i_k, i_{k+1}\} \in f(R)$ (all indices in modulo $n = 4$). Without loss of generality suppose $\{i_1, i_2\} \in f(R)$ (i.e., subject to relabeling of pairs).

Next consider the preference relations for pairs i_2, i_3, and i_4, where each of these pairs is willing to donate a right lobe only if their patient receives their first choice graft under the public information received-graft preference relation. In this case, the preferences R'_{i_2}, R'_{i_3}, and R'_{i_4}, are given as follows:

$$i_3 \ P'_{i_2} \ \emptyset \ P'_{i_2} \ i_1$$
$$i_4 \ P'_{i_3} \ \emptyset \ P'_{i_3} \ i_2$$
$$i_1 \ P'_{i_4} \ \emptyset \ P'_{i_4} \ i_3$$

We next show that, the mechanism f cannot satisfy all three of our axioms in the presence of preference relations R'_{i_2}, R'_{i_3}, and R'_{i_4}:

By assumption, $\{i_1, i_2\} \in f(R)$.

By incentive compatibility of f for i_2,

1. *$\{i_2, i_3\} \notin f(R_{i_1}, R'_{i_2}, R_{i_3}, R_{i_4})$, and thus*
2. *pair i_2 remains unmatched under $f(R_{i_1}, R'_{i_2}, R_{i_3}, R_{i_4})$ since only pair i_3 is acceptable under R'_{i_2}.*

Then $\{i_3, i_4\} \in f(R_{i_1}, R'_{i_2}, R_{i_3}, R_{i_4})$: Otherwise both i_2 and i_3 would be unmatched in $f(R_{i_1}, R'_{i_2}, R_{i_3}, R_{i_4})$, and $f(R_{i_1}, R'_{i_2}, R_{i_3}, R_{i_4}) \cup \big\{\{i_2, i_3\}\big\}$ would Pareto dominate $f(R_{i_1}, R'_{i_2}, R_{i_3}, R_{i_4})$ contradicting mechanism f's Pareto efficiency.

By incentive compatibility of f for i_3, $\{i_3, i_4\} \in f(R_{i_1}, R'_{i_2}, R'_{i_3}, R_{i_4})$.

By Pareto efficiency and individual rationality of f, $\{i_1, i_4\} \in f(R_{i_1}, R'_{i_2}, R'_{i_3}, R'_{i_4})$.

However, the last statement contradicts incentive compatibility of f for i_4: Pair i_4 reports R'_{i_4} instead of R_{i_4} and benefits, gets matched to pair i_1, which is more preferable than i_3 under its preference R_{i_4}. $\diamond$

Example 5 *In this example we show that, if a pair is allowed to prefer a direct transplant to some (but not all) of the strictly better-fit grafts based on its public information received-graft preferences,[29] then a Pareto efficient, individually rational, and incentive compatible mechanism may not exist.*

Consider a liver-exchange pool with three left-lobe compatible pairs $\mathcal{I} = \{i_1, i_2, i_3\}$ with for

[29] This can be interpreted as a "mild" direct transplant bias.

all i_k

$$\tau_P(i_k) = (0,1,0) \qquad\qquad \tau_D(i_k) = (0,1,0,1).$$

The set of mutually compatible exchanges are given as

$$E_c = \big\{\{i_1\},\{i_2\},\{i_3\},\{i_1,i_2\},\{i_1,i_3\},\{i_2,i_3\}\big\}.$$

Observe that, since the left lobe of each donor is sufficiently large for any patient, each donor donates his left lobe under each of these exchanges. Hence whether the pairs are willing to donate their right lobes or not is immaterial in this example.

The public information received-graft preference relation over the set of compatible grafts is given as follows for each pair:

$$\begin{array}{ccccc} i_2 & \succ_{i_1} & i_3 & \succ_{i_1} & i_1 \\ i_3 & \succ_{i_2} & i_1 & \succ_{i_2} & i_2 \\ i_1 & \succ_{i_3} & i_2 & \succ_{i_3} & i_3 \end{array}$$

Suppose no pair has a direct-transplant bias, and thus the preference profile R is given as follows:

$$\begin{array}{c} i_2 \; P_{i_1} \; i_3 \; P_{i_1} \; i_1 \; P_{i_1} \; \emptyset \\ i_3 \; P_{i_2} \; i_1 \; P_{i_2} \; i_2 \; P_{i_2} \; \emptyset \\ i_1 \; P_{i_3} \; i_2 \; P_{i_3} \; i_3 \; P_{i_3} \; \emptyset \end{array}$$

The mutual compatibility graph for this problem is depicted in Figure 10.

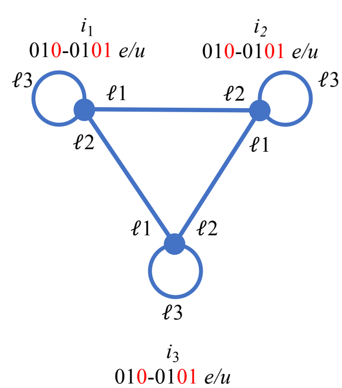

Figure 10: The mutual compatibility graph for Example 5. The left-lobe donations are denoted by letter ℓ and preferences are denoted by numbers $1, 2, 3$ next to the donated lobe for each match.

Suppose f is a Pareto efficient, individually rational, and incentive compatible mechanism. By Pareto efficiency of f, there exists some $\{i_k, i_{k+1}\} \in f(R)$ (all indices are in modulo $n = 3$). Without loss of generality suppose $\{i_1, i_2\} \in f(R)$ (i.e., subject to relabeling of pairs).

Consider the following preferences R'_{i_1}, R'_{i_2}, where pairs i_1 and i_2 have mild direct-transplant bias that allows them to improve the ranking of direct transplant above some of the public information better-fit grafts but not all of them:

$$i_2 \; P'_{i_1} \; i_1 \; P'_{i_1} \; i_3 \; P'_{i_1} \; \emptyset$$
$$i_3 \; P'_{i_2} \; i_2 \; P'_{i_1} \; i_1 \; P'_{i_2} \; \emptyset$$

We next show that, the mechanism f cannot satisfy all three of our axioms in the presence of preference relations R'_{i_1} and R'_{i_2}:

By assumption, $\{i_1, i_2\} \in f(R)$.

By incentive compatibility of f for i_1, $\{i_1, i_2\} \in f(R'_{i_1}, R_{i_2}, R_{i_3})$.

By Pareto efficiency and individual rationality of f, $\{i_2, i_3\} \in f(R'_{i_1}, R'_{i_2}, R_{i_3})$.

However, this contradicts incentive compatibility of f for i_2: Pair i_2 reports R'_{i_2} instead of R_{i_2} and benefits, gets matched to pair i_3, which is more preferable than i_1 under its preference R_{i_2}.

Observe that a similar example can be generated for right-lobe donation decision, by changing all patients' sizes to 1 instead of 0 and making all pairs willing. $\diamond$

C.2 Computation

We give a polynomial-time method in $K = |\mathcal{I}|$ to find our mechanism outcome.

The precedence digraph and a topological order can be constructed in polynomial time (for example see Kahn, 1962). There are at most $2K$ substeps for the algorithm, K in Step 1 and K in Step 2. We can check matchability, construct reduced compatibility graphs, and find an outcome matching in the final reduced compatibility graph in polynomial time. Thus, overall the algorithm runs in polynomial time.

Checking matchability: We can use the following method in each substep for checking matchability of a set $\mathcal{J}$ in the active reduced compatibility graph $G = (\mathcal{I}, E)$:

Define pair weights $\pi^{\mathcal{I}}(j)$ for all $j \in \mathcal{I}$ such that

- $\pi^{\mathcal{I}}(j) \neq \pi^{\mathcal{I}}(i)$ for any $i \neq j$, and
- $\pi^{\mathcal{I}}(j) > \pi^{\mathcal{I}}(i)$ for all $j \in \mathcal{J}$ and $i \in \mathcal{I} \setminus \mathcal{J}$.

Define match weights

$$\pi^E(\varepsilon) \equiv \sum_{j\in\varepsilon} \pi^{\mathcal{I}}(j) \quad \text{for all } \varepsilon \in E.$$

Find an outcome matching $\hat{M}$ of the (polynomial-time) edge-weighted matching algorithm of Edmonds (1965) for edge weights π^E on G. This solves the integer-programming problem

$$\max_{M\in\mathbf{M}[G]} \sum_{\varepsilon\in M} \pi^E(\varepsilon) = \max_{M\in\mathbf{M}[G]} \sum_{i:M(i)\neq\emptyset} \pi^{\mathcal{I}}(i).$$

All pairs in $\mathcal{J}$ are matched in $\hat{M}$ if and only if $\mathcal{J}$ is matchable in G.[30]

Finding the outcome matching: In the final substep of Step 2, Substep 2.(N), by setting $\mathcal{J} \equiv \mathcal{J}_K \cup \mathcal{J}_N^*$ and $G \equiv G_K^*$, we can use the outcome of this above procedure to find the outcome of our mechanism.

Construction of the set of best achievable assignments: In each subset of the algorithm, while pair i is being processed, $\mathcal{J}$ is the set of already committed pairs, and G is the active reduced compatibility graph, first we check using the above method whether $\mathcal{J} \cup \{i\}$ is matchable in G. If so, we can construct $\mathcal{B}(i|\mathcal{J}, G)$ as follows in polynomial time:

Let $\mathcal{I}_1$ be the set of pairs that are best individually rational assignments of i in $E(i)$ with respect to R_i:

$$\mathcal{I}_1 \equiv \max_{R_i} \big\{ j \in \mathcal{I} \ : \ \{j,i\} \in E(i) \big\}.$$

For each $j \in \mathcal{I}_1$, we form the reduced compatibility graph $G^j = (\mathcal{I}, E^j)$ such that

$$E^j \equiv \big[E \setminus E(i)\big] \cup \big\{\{i,j\}\big\},$$

in which the only match of i is with j and all other matches are as in E.

- If $\mathcal{J} \cup \{i\}$ is matchable in G^j then we include j in $\mathcal{B}(i|, \mathcal{J}, G)$, we continue with the next pair in $\mathcal{I}_1$.
- Otherwise, j is not included in $\mathcal{B}(i|\mathcal{J}, G)$, we continue with the next pair in $\mathcal{I}_1$.

[30]The equality follows from Okumura (2014) when there are no direct transplants. This determines a priority matching by Proposition 2 Roth, Sönmez, and Ünver (2005) because of the matroid property of matchings on a graph and this algorithm finds a priority matching with respect to priority induced by pair weights $\pi^{\mathcal{I}}$. Since the weights of the pairs in $\mathcal{J}$ are higher than any other pair in $\mathcal{I} \setminus \mathcal{J}$, it will match pairs in $\mathcal{J}$ whenever it can. Extension with direct transplants is straightforward after showing that matroid property extends with direct transplants (also see Sönmez and Ünver, 2014).

After we process all pairs in $\mathcal{I}_1$, if we placed at least one pair in $\mathcal{B}(i|\mathcal{J}, G)$, then $\mathcal{B}(i|\mathcal{J}, G)$ is constructed at the end of the above process. Otherwise, we consider the next indifference class of i among matches in $E(i)$, $\mathcal{I}_2$, with respect to R_i, similarly, and continue so on until $\mathcal{B}(i|\mathcal{J}, G)$ is constructed. Then, we activate the next reduced compatibility graph using $\mathcal{B}(i|\mathcal{J}, G)$.