Multi-Modal Paired Exchange*

Anand Siththaranjan†

August 11, 2025

Preliminary and Incomplete

This paper is frequently updated. Contact me for the latest version.

Abstract

This paper develops a model of paired exchange that integrates multiple donation technologies with observable risks. Integration provides more opportunities for patient-donor pairs to match, while reducing the risks placed on donors by providing safer procedures to be taken when possible. We construct pairwise exchange mechanisms that satisfy efficiency, stability, and strategy-proofness, and study applications including exchanges with different organ pools, kidney exchange with ABO-desensitization, and exchanges with multiple donors. We analyze the welfare improvements of integrating a kidney and liver exchange over non-integrated exchanges, where simulations find a 10 to 20% relative increase in transplants under reasonable conditions.

1 Introduction

For patients suffering from organ failure, such as renal or liver failure, transplantation from a willing donor is an effective means of improving the patient's life expectancy. However, a key concern is the donor's health, which can be negatively affected by undergoing the transplant procedure. As such, it is of utmost importance for doctors to ensure that risks and benefits are appropriately balanced and agreed upon by both parties. As opposed to direct donation, which relies on willing and compatible donors, when patients have willing but incompatible donors, paired donor exchange has emerged as a means to take advantage of a *coincident of wants*: when there are two pairs in such a situation, but the patient of

*We thank Haluk Ergin and Chris Shannon for their support throughout this project. Furthermore, we are grateful for helpful comments from Oğuzhan Çelebi, Tak-Huen Chau, Richard Chen, Yuichiro Kamada, Scott Kominers, Jianmeng Lyu, and Quitzé Valenzuela-Stookey, as well as assistance from Jason Choi.

†University of California, Berkeley, Department of Electrical Engineering and Computer Science, Berkeley, 2594 Hearst Ave, Berkeley, CA 94720, USA. E-mail: anandsranjan@berkeley.edu

each pair is compatible with the donor of the other pair, then they can *exchange* donors and undergo their respective operations.

This setting has been well studied for exchanges in many organ markets, such as for kidneys (Roth et al., 2005), livers (Ergin et al., 2020) and lungs (Ergin et al., 2017). Yet it might be possible to allow for exchanges across these markets. Such an idea was discussed by Dickerson and Sandholm (2017), who proposed a computational approach to showcase the benefits of such an integration. For example, consider allowing a pair that requires a kidney donation and pair that requires a liver donation, to exchange donors who then donate different organs. By allowing donors to donate organs different than what their patient requires, or potentially donate in different ways such as through ABO-desensitization[1] of the patient (Andersson and Kratz, 2020) or different lobes of the liver (Ergin et al., 2020), this can increase the number of compatible transplants a patient has. Such a swap has informally occurred before, where a patient-donor pair saw the work of Dickerson and Sandholm (2017) and were motivated by this new possibility (CMU, 2019). However, there is still no centralized mechanism that allows for such an option. Rather than force pairs to instigate such possibilities, which would require finding another willing pair to participate in the swap, we propose developing an exchange that integrates multiple modes of donations into a single paired exchange.

In addition to increasing the number of matching opportunities, donors whose only options were to undergo riskier transplants in order for their patient to receive a new organ now have more options available to them. In the case of kidneys and livers, the latter is a riskier donation for living donors in both mortality and morbidity[2]. Consider a kidney patient Alice whose donor Bob is willing to donate either a kidney or liver, but there is no other kidney patient they can be matched with. As well, there is a liver patient Claire whose donor David can donate either organ, and they have the option of being matched in an exchange with a kidney or liver patient. In this case, we can reduce unnecessary risks to David by allowing them to donate a kidney to Alice. Though Bob must now take on a greater risk by donating a liver to Claire as opposed to a kidney, it is the only available option to them. By combining exchanges, we have the option of limiting donor risks through such trades. We believe this to be an important consideration for increasing donor participation in paired exchange.

The literature has largely studied how to design organ exchanges for specific organ pools,

[1]This refers to a procedure that a patient undergoes in order to reduce the potential of rejection if they were to receive a blood-type incompatible kidney. In doing so, they can increase the set of feasible transplants available to them.

[2]Gaston et al. (2015) find that the mortality rate for living kidney donors to be 0.03%, and mortality rates to be less than 1%. As noted in Ergin et al. (2020), Lee (2010) find a mortality rate for left and right lobe liver donation to be 0.1% and 0.4-0.5% respectively, and Mishra et al. (2018) find morbidity rates to be 7.5% and 28% respectively.

yet different organs can introduce unique challenges. For example, liver exchanges can allow for two modes of transplants: left and right lobe transplant (Ergin et al., 2020). Both modes of transplants require blood type and size compatibility, where the latter requires a patient receives a liver greater than some minimum size specific to the patient. The right lobe is preferred in one sense - as they are bigger than left lobes, donating via them increases the number of compatible recipients compared to donating via the left lobe. This is helpful for a pair in the exchange as they have more opportunities to be match in a pairwise exchange with another pair. However, right lobe transplants are riskier than left lobe transplants for donors, and thus pairs may have different levels of risk they are willing to undertake. This can lead to non-trivial incentive issues that relate to truthful reporting of a pair's willingness to undergo different donation modes. Similarly, due to improvements in medical technology, the standard model of kidney exchange that typically has a single mode of donation can be expanded. Andersson and Kratz (2020) studies the use of immunosuppressant technology to allow patients to be blood type compatible with any donor. This can be seen as a new donation mode, and as emphasized in their work, this mode is less desirable to receive than being matched with an already compatible donor.

This paper takes a market design approach to implementing paired exchange across different modes of donation. We identify a *coincidence of wants* through the integration of new types of donation, such as different modes of donating a certain organ (e.g. left or right lobe of a liver) as well as donating different organs (e.g. kidney or liver). Given this, we ask how can we design *desirable* mechanisms. Our criteria for desirability include classic objectives in the field - efficiency, strategyproofness, and individual rationality - as well as a novel condition not usually studied in the literature[3] - weak-core stability. We provide a motivation for this as a means of mitigating strategic participation incentives by hospitals (Ashlagi and Roth, 2014). As we observed with comparing the risks between kidney and liver donation, as well as with ABO-incompatible donation, many risks are objective. We leverage these assumptions on the commonality of preferences when designing a mechanism. The main piece of private information is willingness level, that is, at what point is a modality by which to donate or receive via too risky.

We begin our exploration by studying *dual-mode* exchanges, that is exchanges where agents can participate by donating, or receiving, via two different means. We motivate this environment by considering two working examples. We first study kidney donation with ABO-incompatible donations, where patients can receive a blood type compatible or incompatible kidney from a donor, however the latter is less preferred due to the financial burden

[3]Though common in the general market design and matching literature, it is not usually studied in application to paired exchange.

it places on the patient (Andersson and Kratz, 2020). We also consider organ exchange when two donors are available to be used and preferences are "risk-ordered" (i.e. based in objective medical risk). These working examples are nested in an abstract model of dual-mode exchange, and in doing so we generally characterize what exchanges admit desirable mechanisms. In particular, we propose a condition called *weak acyclicity* and find it to be necessary and sufficient. We apply our sufficiency result to kidney exchange with ABO-incompatible donation, and with two risk-ordered donors, to construct a desirable mechanism using previous work by Ergin et al. (2020). On the other hand, we use our necessity result to find that liver exchange with two risk-ordered donors does not have a desirable mechanism.

We then study how to develop mechanisms for more general *multi-modal* exchanges. Our working example of interest is the integration of kidney and liver exchanges, where donors can donate via their kidney, left liver lobe, or right liver lobe. This is a specific instance of a general model that integrates arbitrarily many exchanges for different organs, where preferences over participating in an exchange reflect the objective medical risk of donating. We identify an ideal property of a multi-modal exchange that builds on the notion of acyclicity in Ergin et al. (2020), which we term *separability*. We find that this creates a structural decomposition within the space of modalities, and allows us to develop a simple desirable mechanism for multi-modal exchanges. When a multi-modal environment is not completely separable, but satisfies such properties in a certain partitional sense, we show that if there are desirable mechanisms for each element of a partition of the modality space, then we can create a modular "meta-mechanism" that retains the desirable properties if preferences on modalities are common. For example, if we have a desirable mechanism for kidney and liver exchange separately, we are able to "stitch" them together in a way that preserves their properties. The promise of this approach lies in modeling the integration of risk-ordered organ exchanges, which we find to naturally satisfy partition separability. We find such a modular approach to be important given that modes of organ donation are routinely being developed as the medical field progresses, both in existing organ exchanges and potentially new ones. Thus having a mechanism that can seamlessly integrate desirable mechanisms specific to each exchange into a single desirable mechanism for the whole exchange allows our approach to be robust to future technological advances. Though it is plausible that market integration can have positive welfare effects in this context[4], the idea that mechanisms specific to each market can be combined to form a new desirable mechanism for the integrated market is novel.

An application of key interest is the simple but practical model of integrating kidney and

[4]Market integration in general is not always welfare improving. See Kumar et al. (2022) for an analysis of the integration of Shapley-Scarf markets.

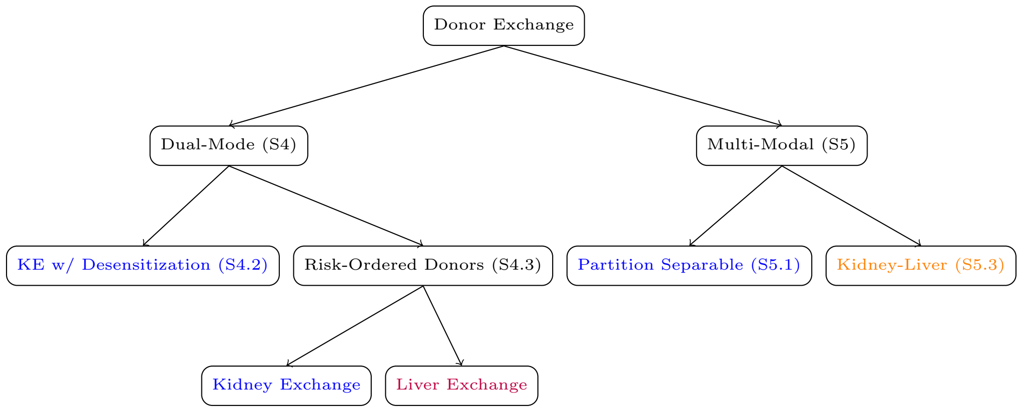


Figure 1: Diagram of Results and Applications. Result type: a desirable mechanism exists, a desirable mechanism does **not** exist, and ssimulation of a desirable mechanism.

liver exchanges. To characterize any welfare improvements over the status quo baseline, that is pairwise matchings in each organ exchange independently, we show that our mechanism always provides a weak increase in the number of matches. We complement this theoretical result with simulations that quantify, using Korean patient data, anticipated gains from integrating the exchanges across various assumptions on risk-tolerance of patients. We find, across such tolerances, anywhere from a 10 to 20 percent relative increase in the number of transplants over the baseline.

Outline. This paper is organized as follows. Section 2 describes the general model contained within this paper, as well as our desiderata and domain-specific background. Section 3 provides some background on medical and institutional factors pertaining to organ exchange, which informs the design of our solution. Section 4 studies exchanges with two donation modes, providing a theoretical characterization of which such exchanges admit desirable mechanisms (Section 4.1) and analyzing specific applications: kidney donation with ABO-desensitization (Section 4.2) and two risk-ordered donors (Section 4.3). Section 5 studies a specific environment with more than two modes, which we refer to as a risk-ordered integration of exchanges. Section 5.3 provides a simulation analysis of integrating kidney and liver markets. Our theoretical results are summarized in Figure 1.

1.1 Related Literature

Kidney Exchange. The kidney exchange literature in market design begins with Roth et al. (2004), which provides a modification of the top-trading cycles algorithm (TTC) to allow chains of kidney donation instigated by unpaired donors to be performed. Practical considerations such as pairwise cycle restrictions with dichotomous preferences[5] were explored in Roth et al. (2005), where tools from matching theory were leveraged. More general cycle restrictions were subsequently studied in Roth et al. (2007), where the goal was to identify transplant maximizing exchanges under different size restrictions. Such restrictions are important in practice, and appear as a necessary constraint in the literature on organ exchange (Ergin et al., 2020, 2017; Andersson and Kratz, 2020). When preferences can be richer, such as those over kidneys (Roth et al., 2004; Nicolo and Rodriguez-Alvarez, 2017), then the problem resembles a house exchange problem with cycle constraints. It is known that in general it is diffuclt to satisfy efficiency and strategyproofness as a result (Balbuzanov, 2020; Kamada and Yasuda, 2025).

Other Exchanges. Other types of organ donation and technologies have been studied, including liver exchanges (Ergin et al., 2020), dual-donor exchanges like lung and kidney-liver (Ergin et al., 2017), ABO-incompatible kidney donation via desensitization (Andersson and Kratz, 2020; Heo et al., 2022), and multi-donor kidney exchange (Gilon et al., 2019). We generalize the results of Ergin et al. (2020) to allow for wider application to dual-mode environments - that is environments with two means of donating or receiving - such as desensitization and two-donor kidney exchange under risk-preference assumptions. We also consider strategy-proofness and weak-core stability for pairwise exchanges as part of our key criteria, and previous literature has not studied at least one of these properties. Andersson and Kratz (2020) primarily characterize the structure of efficient matchings in their setting, and note that there can be incentive issues. Our work shows that this can be dealt with while maintaining efficiency. Gilon et al. (2019) study organ exchange where patients can bring multiple of their own donors to the exchange. Though they are motivated by kidney exchange, they consider an abstract compatibility. We impose structure on the set of feasible trades using the compatibility graph, as well as assumptions on preferences, to achieve positive results.

Exchange Integration. Similar to our work is that of Watanabe (2022), who take an approach like that of Ergin et al. (2020) and apply this to the integration of kidney and liver exchanges. Our algorithms are similar when we restrict attention to their environment, however our approach exposes a simpler, intuitive structure due to its modularity and abstraction. In particular, our algorithm can be agnostic to the underlying mechanism used in each of the kidney and liver exchange, while preserving their properties. Furthermore, our

[5] This refers to preferences where only compatible matches and being unmatched are individually rational, and all compatible matches are indifferent to one another but are strictly preferred to being unmatched. These are also know an 0-1 preferences.

model is able to generalize to new environments, and we provide a stronger theoretical characterization of our solution by showing weak-core stability and welfare improvements over the baseline. Dickerson and Sandholm (2017) also study a similar problem of integrating kidney and liver markets, but primarily focus on a computational approach to transplant maximization and largely ignore preferences and incentives.

2 Model

Let $\mathcal{I}$ be the set of agents, where $N = |\mathcal{I}|$, and $t_i \in \mathcal{T}_i$ be the type of agent i. Let $\mathcal{M} = \{m_1, m_2, ...\}$ denote the set of modalities, and $\emptyset$ denote the outside option (i.e. not being matched). We assume agents have strict preferences over $\mathcal{M} \cup \emptyset$ such that $m_k \succ_i m_{k+1}$ and $m_1 \succ_i \emptyset$. Note that this implies that agents have a common preference ranking over modalities, reflecting our common risk ordering assumption[6]. Let the space of preferences be given by $\mathcal{R}$. However, the preferences of agents may differ in how they rank the $\emptyset$. For a subset of agents $A \subseteq \mathcal{I}$, $m \in \mathcal{M}$ and a preference profile $\succ \in \mathcal{R}^N$, we let $A_\emptyset(m | \succ)$ be the set of agents that do not find m_i individually rational, i.e. $A_\emptyset(m | \succ) = \{i \in A | \emptyset \succ_i m_i\}$.

Let $\tau_m : \mathcal{T} \times \mathcal{T} \to \{0, 1\}$ be the compatibility function for modality $m \in \mathcal{M}$. That is, an agent i can donate via mode m to agent j if and only if $\tau_m(t_i, t_j) = 1$. Otherwise $\tau_m(t_i, t_j) = 0$. We use the notation $i \to_m j$ to mean that i can donate to j via modality m, i.e. $\tau_m(t_i, t_j) = 1$. Correspondingly, if i cannot donate to j via modality m, then we use the notation $i \not\to_m j$.

Define an **exchange problem** to be the tuple $\mathcal{E} = (\mathcal{I}, \mathcal{T}, \mathcal{M}, \tau, \mathcal{R})$, and the analogous family of exchange problems to be given by

$$\mathbf{E} = \{\mathcal{E} = (\{1, \ldots, n\}, \mathcal{T}^n, \mathcal{M}, \tau, \mathcal{R}) | n \in \mathbb{N}, \tau_m : \mathcal{T} \times \mathcal{T} \to \{0, 1\}\}$$

The compatibility graph with respect to an exchange is an edge-labeled directed graph $G_\mathcal{E} = (V, E)$, where $V = \mathcal{I}$ is the set of vertices and $E \subseteq V \times V \times \mathcal{M}$ the set of labeled edges such that $(i, j, m) \in E$ if and only if $i \to_m j$. Let G_C^m denote the directed graph induced by only considering edges corresponding to modality m, and $\bar{G}_C^m$ denote the undirected graph where an edge between (i, j) exists if $(i, j, m), (j, i, m) \in E$. We say a cycle is an n-cycle if it is of length n. A **matching** is a set of disjoint 2-cycles in the compatibility graph. Denote the set of matchings by $\mathbf{M}$, which is implicitly determined by the exchange problem $\mathcal{E}$ being considered. Let $\mathcal{I}(F)$ be the set of agents involved in a matching F. We will use the notation

[6]This is a necessary assumption in our work. We motivate this with the claim that medical risk are objective, and are often a key part of agents' preferences. However, there are reasonably cases where this may not always hold. We note such a case in Footnote 10.

$F_{\mathcal{I}}(i) = j$ for $j \in \mathcal{I}$ such that there exists $m \in \mathcal{M}$ where $(i, j, m) \in F$. Similarly, $F_{\mathcal{M}}(i) = m$ for $m \in \mathcal{M}$ such that there exists $j \in \mathcal{I}$ where $(i, j, m) \in F$. That is, $F_{\mathcal{I}}(i)$ is the agent i donates to in F, and $F_{\mathcal{M}}(i)$ is the mode that i donates via in F.

With abuse of notation, we say that $M \succeq_i M'$ for $M, M' \in \mathbf{M}$ if and only if $M_{\mathcal{M}}(i) \succeq_i M'_{\mathcal{M}}(i)$. In other words, an agent i (weakly) prefers the matching M over M' if and only if they are donate via a weakly preferred modality in M than in M'.

Given a graph $G = (V, E)$, the induced subgraph with respect $A \subseteq V$ is $G(A) = (A, E_A)$ where $E_A = E \cap (A \times A)$. Define a priority order Π to be a strict total order on $\mathcal{I}$. We let $\texttt{MaxMatch}(G|\Pi)$ denote a maximum cardinality matching of the graph G with priority determined by Π. For a bipartite graph G with partitions A and B, define similarly $\texttt{BipartiteMatch}(G|\Pi, A, B)$. Though both functions maximize the cardinality of a matching in G, because of priorities the underlying algorithms are different as `BipartiteMatch` exploits the structure of a bipartite graph, whereas `MaxMatch` applies to general graphs[7].

For $(i, m) \in \mathcal{I} \times \mathcal{M}$ and $A \subseteq \mathcal{I} \times \mathcal{M}$, let $\texttt{Matchable}(i|G, A)$ output `True` if $i \in \mathcal{I}$ can be matched via m in G while ensuring that for all $(j, m') \in A$, j can be matched via m'. Otherwise it will return `False`[8].

Given a family of exchange problems $\mathbf{E}$, a family of matching mechanisms is $\phi : \mathbf{E} \times \mathcal{R}^N \to \mathbf{M}$. A mechanism with respect to a specific $\mathcal{E} \in \mathbf{E}$ will be denoted $\phi_{\mathcal{E}}$, and if $\mathcal{E}$ or $\mathbf{E}$ is clear from context, we will simply use ϕ and refer to it as a mechanism.

2.1 Desiderata

In this section, we describe various desiderata. Fix $\mathcal{E} \in \mathbf{E}$ and $\succeq \in \mathcal{R}^N$. A matching M is

1. **individually rational** if for all $i \in \mathcal{I}$, $M \succeq_i \emptyset$.

2. **Pareto efficient** if there does not exist $M' \in \mathbf{M}$ such that $M' \succeq_i M$ for all $i \in \mathcal{I}$ and there is some $j \in \mathcal{I}$ such that $M' \succ_j M$.

3. **weak-core stable** if there does not exist non-empty $C \subseteq \mathcal{I}$ and a matching $M' \in \mathbf{M}$ such that for all $i \in C$, $M' \succ_i M$, and for all $i \in C$ we have that $j = M'_{\mathcal{I}}(i) \in C$.

A generally weaker condition than weak-core stability is **pairwise stability**, which requires that there does not exist $i, j \in \mathcal{I}$ and m, m' such that $i \to_m j$, $j \to_{m'} i$, $m \succ_i M_{\mathcal{M}}(i)$,

[7] This affects the time complexity of the algorithm, unlike in maximum matching without priorities where the Micali-Vazirani algorithm (Micali and Vazirani, 1980) for general graphs matches the worst-case time complexity of the Hopcroft-Karp algorithm (Gabow, 2017) for bipartite matching. When there are priorities, bipartite graph algorithms (Turner, 2015) generally have a lower worst-case time complexity than for general graphs (Okumura, 2014).

[8] Checking matchability for pairwise exchanges can be done in polynomial time. Ergin et al. (2020) describe such a function.

and $m' \succ_j M_{\mathcal{M}}(j)$. Note that when indifferences exist in a model, only the weak-core and not the strong-core[9] is guaranteed to exist.

Finally, a mechanism ρ is **strategy-proof** if for all $i \in \mathcal{I}$, $\succ \in \mathcal{R}^N$ and $\succ'_i \in \mathcal{R}$, then $\rho(\succ) \succeq_i \rho(\succ'_i, \succ_{-i})$. That is, it is a weakly dominant strategy for agents to report preferences truthfully. For $\mathcal{E} \in \mathbf{E}$, we say a mechanism $\phi_{\mathcal{E}}$ is **satisfactory** if $\phi_{\mathcal{E}}$ is strategy-proof, and all elements of its range are Pareto-efficient, individually rational, and weak-core stable. For a family of mechanisms ϕ, if for every $\mathcal{E} \in \mathbf{E}$ we have that $\phi_{\mathcal{E}}$ is satisfactory, then we say that ϕ is **desirable**.

The following simple result highlights that, in the setting of matchings, pairwise stability is sufficient for weak-core stability:

Proposition 1. *An individually rational pairwise stable matching is weak-core stable (among matchings).*

Proof. See Appendix A.1. □

3 Background

We provide some institutional background to motivate common assumptions, considerations and restrictions in the literature that will be leveraged in our work.

3.1 Simultaneous Operations.

An important point is when an exchange between multiple pairs is done, it tends to be the case that the operations are done simultaneously. The reason for this is that if the patient of one pair has received an organ from another donor, and their donor has not yet donated, there cannot be any punishment to the donor or the patient should the former choose not to donate. That is, we cannot force the donor to donate, nor take back the organ transplanted to the patient. Thus doing sequential or asynchronous operations can lead to a holdup problem. Although this problem is alleviated when considering chains of donation that are initiated by a deceased donation, in our setting we only consider paired donation.

3.2 Small Exchanges.

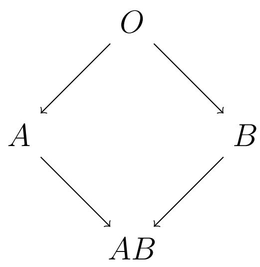
The literature often considers pairwise exchanges, that is the restriction that any pair i whose donor donates to the patient of pair j also has the donor of pair j donating to the

[9]The strong core requires that in a blocking coalition, all agents weakly improve and at least one agent strictly improves over their original allocation.

patient of pair i. Though it is possible to do larger size exchanges, and this has been done in reality, there are a number of considerations that make pairwise exchange a good starting point in theory and practice. Due to the requirement of having simultaneous operations, smaller size exchanges are preferred as large exchanges can impose a prohibitively great logistical and medical burden on the hospital performing the multiple operations. However it is also common to have three-way exchanges in kidney exchange (Ashlagi and Roth, 2021). Our work introduced different methods of donation in a single model, hence may require different expertise and considerations of potentially complications due to the need of simultaneous operations. As such, we consider pairwise exchanges primarily as this is practically most feasible in an initial implementation. Beginning with such an approach in novel environments is common in the literature, as in the initial papers on kidney exchange with ABO-desensitization (Andersson and Kratz, 2020) and liver exchange (Ergin et al., 2020).

3.3 Tissue-Type Compatibility.

From a medical point of view, an important factor in the ability of a donor to be able to donate to a patient is their mutual biological compatibility (Roth et al., 2007). This is a product of different types of compatibility, such as blood-type, size and tissue-type compatibility. For all donors and patients, their individual blood-type and size are known, and their mutual compatibility can be inferred from this. As such, we can incorporate this information into a mechanism by treating this as a restriction on the set of feasible exchanges. However, tissue-type compatibility is more difficult to ascertain. In particular, the patient must not have antibodies to a donor's human leukocyte antigens (HLA), thus requiring testing specific to the patient and donor (Ashlagi and Roth, 2021). Thus to incorporate this as a restriction on the feasible set requires that many patient-donor pairs be tested in this fashion. This can be difficult in practice, and a simplifying assumption is to assume that donors and patients from different pairs are tissue-type compatible (Roth et al., 2005). On the other hand, patients and donors from the same pair are often already tested for tissue-type compatibility in the first place. The reason we usually have this information is that donors brought by a patient in a pair are usually those who want to donate specifically to that patient. As such, they must have already checked compatibility prior to entering the paired exchange mechanism. We assume for simplicity that the donor in all pairs is tissue-type incompatible with their patient.



(a) Compatibility Graph

$X \leq_B Y$	O	A	B	AB
O	✓	✓	✓	✓
A		✓		✓
B			✓	✓
AB				✓

(b) Table description

Figure 2: Blood type compatibility.

3.4 Kidney Exchange

The standard models of kidney exchange primarily consider blood-type compatibility as the main determinant of biological compatibility. A patient's blood type is given by the presence of two antigens, either A or B. If they are missing both antigens, we say they have blood type O. Hence the set of possible blood types considered are A, B, AB, or O. A person i can donate to another person j if whenever j is missing some antigen $k \in \{A, B\}$, then so is i. For example, O donors are compatible with all patients as they are missing every antigen, however O patients are only compatible with O donors. On the other hand, AB donors can only donate to AB patients as they are missing no antigen, but as patients they can receive from any donor. We depict the blood type compatibiltiy relation, where $x \leq_{\mathcal{B}} y$ for $x, y \in \mathcal{B} = \{O, A, B, AB\}$ if x can donate to y, in Figure 2.

Though an ideal kidney exchange is one where the patients and donors are blood-type compatible with one another, there exists technology by which blood-type incompatible donations can be performed. As Andersson and Kratz (2020) note, though the graft survival rates for such a transplant are identical to compatible donations, the main issues with such an approach are the financial cost of the immunosuppressant, longer waiting time prior to transplantation, and the need for additional medical treatment. As such, it is reasonable to assume that this is a less preferred mode of receiving a kidney than a compatible donation.

3.5 Liver Exchange

In liver exchange, a donor donates a portion of their liver, referred to as a lobe, rather than their whole organ as in kidney exchange. Furthermore, liver exchange differs from kidney exchange in two key ways: biological compatibility and modes of donation. The main features of compatibility for liver donation is blood type and size compatibility. By the latter we mean that a potential donor is compatible with a patient if the lobe they donate is larger than what the patient requires. This is important as donors are also able to choose which lobe - the left or right - they donate, given that they differ in size. However, it is known that the right lobe is more dangerous to donate in terms of mortality and morbidity than the left lobe (see Footnote 2). Given this, we assume that agents will prefer donating their left lobe over their right lobe. Similarly, kidney donation is also known to be safer on both metrics than liver donation, whether left or right lobe. Thus we maintain the same assumption on preferences when comparing kidney to liver donation[10].

Though desensitization is possible, it is mainly done with donations from brain dead patients (Egawa et al., 2023). We discount this as a possibility due to its limited current implementation for living donation, but note that it might have scope for exploration in future work.

4 Dual-Mode Exchanges

In this section, we study the setting where our exchange only has two modes, i.e. $\mathcal{M} = \{m_1, m_2\}$. This section provides a generalization of Ergin et al. (2020), which studies a model that characterizes the existence of desirable mechanisms for liver exchange. In their work, the two modes of interest corresponded to donating a left lobe as opposed to a right lobe. It is assumed that objective medical risks associated with donation determines an individuals preference ordering between the two modes. We consider a similar assumption as their environment, and find a weaker condition that guarantee the existence of a desirable mechanism. We terms this condition *weak acyclicity*, which generalizes the notion of *acyclicity* (Ergin et al., 2020). We further find that not only is it sufficient, but it is also a necessary condition.

It is natural to ask whether our generalization was necessary in so far as practical applications are concerned. For example, is acyclitiy naturally satisfied in other dual-mode environments? To showcase the importance of our generalization, we find practical environments - kidney exchange with ABO-desensitization, and with two risk-ordered donors - that satisfy weak acyclicity but not acyclicity. We highlight our necessity result by showing the the non-existence of a desirable mechanism in the setting of liver exchange with multiple donors.

[10] Note this is it not necessarily obviously true. Since the liver regrows whereas the kidney does not, it is plausible to imagine that some agents might have the opposite preference based on the preference to "feel whole". We focus on preferences that reflect medical risk, though such alternative preferences may have a foundation in practice. Understanding whether such preferences exist is an interesting empirical question beyond the scope of this paper.

4.1 Characterization

In characterizing the existence of desirable mechanisms, we begin with the same graph construction as Ergin et al. (2020). Given G_C and a set of agents $A \subseteq \mathcal{I}$, construct the following digraph $G(A) = (A, E')$: let $(i, j) \in E'$ if and only if

1. i can donate to j via m_1, i.e. $i \rightarrow_{m_1} j$, and
2. j can donate to i only via m_2, i.e. $j \not\rightarrow_{m_1} i$ and $j \rightarrow_{m_2} i$.

We refer to $G(A)$ as the **precedence digraph**, as in Ergin et al. (2020). We say a pair of agents $i, j \in \mathcal{I}$ is **mutually compatible** via $m \in \mathcal{M}$ if $i \leftrightarrow_m j$. We say that $G(A)$ is **weakly acyclic** if every cycle contains a pair of agents mutually compatible via m_1. Furthermore, it is **acyclic** if there are no cycles. This definition of acyclicity follows from Ergin et al. (2020). Let `TopOrder`(G) return a topological order on G if it exists[11].

Observation 1. *$G(\mathcal{I})$ is weakly acyclic if it is acyclic.*

This follows from the fact that if there are no cycles, then it is vacuously true that all cycles contain a pair mutually compatible via m_1 donation. We now informally describe a basic version of the "Preference Adaptive" (PA) algorithm studied in Ergin et al. (2020), whose properties we aim to generalize: let $\Pi =$ `TopOrder`$(G(\mathcal{I}))$.

1. Compute a maximum match M_0 using some priority order via m_1 and promise matched agents a match via m_1
2. For agents not matched but willing to donate via m_2, add them to the set A.
3. Process through A in order given by Π:
   (a) If k is `Matchable` (while preserving promises), then promise a match via m_1
   (b) Transform to m_2 otherwise (if m_2 is feasible)
4. Compute a maximum match via m_2 while maintaining promises.

Note that for the algorithm as-is to be well-defined, we require that $G(\mathcal{I})$ is acyclic. Ergin et al. (2020) ensure this by showing that, in their setting of left and right lobe liver exchange, the directed graph is always acyclic due to restrictions imposed by the compatibility graph.

[11] A topological order of a directed graph G with vertices V and edges E, if it exists, is a strict total order $\rhd$ over V such that for all $i, j \in V$ such that $(i, j) \in E$, then $j \rhd i$. It is known that such an order always exists if and only if G is acyclic, however it may not be unique. We can break ties using an exogenous priority order.

However, we can observe that their proof leverages acyclicity of $G(\mathcal{I})$, whereas, in general, A is a subset of $\mathcal{I}$. Thus the computation of the order is only required with respect to $G(A)$, not $G(\mathcal{I})$. Hence instead consider a version of the PA algorithm where Π is computed as $\texttt{TopOrder}(G(A))$ after A is constructed, which we refer to by the same name[12]. We leverage this observation by showing that this change ensures the algorithm is well-defined whenever $G(\mathcal{I})$ is weakly-acyclic, in the sense that whenever the topological order is attempted to be computed for any set A that could be generated, an order will necessarily exist.

Proposition 2. *If $G(\mathcal{I})$ is weakly acyclic, then the PA algorithm is well-defined.*

Proof. See Appendix A.2. □

With the key insight that the construction of the order is only necessary **after** agents are matched mutually via m_1, we can note that these are exactly the agents that will make up any cycle in $G(\mathcal{I})$ when it is weakly acyclic. Hence, if they are precluded, $G(A)$ is acyclic and thus admits a topological order. We can see the process of the initial mutual m_1 matching step as "breaking" any cycles that might exist. Given this, we subsequently find weak acyclicity to be a sufficient condition for this mechanism to be desirable:

Theorem 1 (Weak Acyclicity is Sufficient)**.** *If $G(\mathcal{I})$ is weakly acyclic then the Preference Adaptive algorithm of Ergin et al. (2020) is a desirable mechanism.*

Proof. See Appendix A.4. □

Since the mechanism is the same as that of Ergin et al. (2020), our proof primarily builds on their analysis by extending their results to weak-core stability, and showing the other properties hold under the weaker conditions. That is, the acyclicity of $G(\mathcal{I})$ is not required for their results, only the acyclicity of $G(A)$.

We have shown that the condition of weak acyclicity is sufficient, but how about necessity? The following results shows a converse to our sufficiency theorem.

Theorem 2 (Weak Acyclicity is Necessary)**.** *If $G(\mathcal{I})$ is not weakly acyclic, then there is no desirable mechanism.*

Proof. See Appendix A.5. □

For a graph to not be weakly acyclic, that means we must be able to find a cycle where no pair of agents is m_1 mutually compatible. The proof follows by contradiction, where

[12] Alternatively, one can observe that if instead of computing the topological order we used the condensation of $G(\mathcal{I})$, which always exists, then the partial order induced by this would be a strict total order when restricted to A whenever $G(\mathcal{I})$ is weakly acyclic.

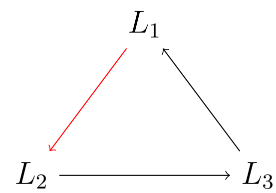


Figure 3: Precedence digraph with the matching for the preference profile WWW in Example 1 highlighted.

we consider the environment that induces this cycle and assume that there does exists an efficient matching from a strategy-proof mechanism. It is without loss to assume that the cycle is simple: if there is a $\mathcal{E} \in \mathbf{E}$ with cycle that is not simple, then we can find $\mathcal{E}' \in \mathbf{E}$ where there is a simple cycle by eliminating parallel paths in the original cycle[13]. If we can show that there is no satisfactory mechanism in this case, then there is no desirable mechanism for the considered class of problems. To show this, we leverage the fact that pairs can match one of two ways: one agent matches via m_1 and the other by m_2, or both via m_2. The former case means that the pair of agents are adjacent in the cycle. Assuming for contradiction that a desirable mechanism exists and thus can be applied to this setting, we can use the constraints of efficiency and strategy-proofness to determine how agents could be matched. We can observe that in any match, there is at least one agent that must donate via the least preferred modality m_2. Thus strategyproofness requires that misreporting, i.e. saying that m_2 is not individually rational, must mean that agent is not matched at all. Efficiency requires that the agent who succeeds them in the cycle, who they could donate to via m_1 and receive via m_2, must be matched. Repeatedly considering misreports by an agent that donates via m_2 allows us to identify a contradiction when we arrive at the final agent who needs to be matched to ensure strategyproofness for the report of the penultimate agent, but has no matching opportunities as every other agent finds m_2 infeasible. The following illustrates the impossibility result through an example with three agents.

Example 1 (An environment without a desirable mechanism)**.** Consider the precedence digraph in Figure 3 with the cycle $L_1 \to L_2 \to L_3 \to L_1$, and assume there is a desirable mechanism. Since this is a cycle of length 3, and no two agents that are adjacent can be m_1-mutually compatible, there are no m_1-mutually compatible pairs in this cycle. Let XYZ for $X, Y, X \in \{U, W\}$ be a preference profile, where W means that m_2 is IR (i.e. willing) and U it is not (i.e. unwilling). Note that even if there were mutually compatible m_2 pairs, such a matching would never be efficient. Given this and the fact that at most two agents

[13]Recall that the definition of $\mathbf{E}$ ensures that we can eliminate agents from a $\mathcal{E}$ and it is still a member of $\mathbf{E}$.

can be matched in a matching, we can denote a matching by (L_i, L_j), where L_i donates via m_1 and L_j donates via m_1.

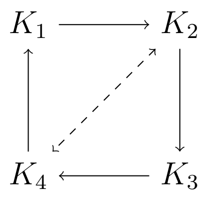
First consider the profile WWW. Without loss of generality, assume that (L_1, L_2) is the matching given by the mechanism, where L_1 donates by m_1 and L_2 donates by m_2. Now consider UUW, and observe that (L_2, L_3) is the only individually rational and efficient match. For UWW, the possible individually rational, efficient matches are (L_2, L_3) and (L_1, L_2). The former would contradict strategyproof-ness of the mechanism, as L_1 can misreport from U to W and be strictly better off. However the latter would also contradict strategyproof-ness, as a misreport by L_2 from W to U would give them a strictly better outcome. As no match is feasible, this mechanism cannot be desirable. △

We have provided a general characterization of dual-mode settings that admit desirable mechanisms. This approach builds on previous work, and an open question is how practically necessary was our generalization. For example, is it the case that other applications of interest already induce an acylic preference digraph? We illustrate the usefulness of our results through different practical applications, two of which admit non-acyclic but weakly acyclic environments, and one of which leverages our converse result to show the general non-existence of a desirable mechanism for the given setting.

4.2 Incompatible Donation via ABO-Desensitization

The prototypical model of kidney donation considers agents that consist of patient-donor pairs who are incompatible either due to tissue-type incompatibility or blood-type incompatibility. With the assumption that all patients are tissue-type compatible with all other donors, the problem of maximizing the number of pairwise exchanges is the same as identifying the maximum match in a compatibility graph. Because certain blood types are more rare than others and thus have heterogeneous demand in the exchange market, certain agents fare better than others due to their blood type. To overcome this barrier, the development of ABO-desensitization has allowed patients to receive transplants from ABO-incompatible donors. In the context of studying the properties of efficient matchings, this problem has been studied by Andersson and Kratz (2020). We extend this line of work by showing how this model satisfies weak-acyclicity strictly, and thus there exists a desirable mechanism.

The model is formally described as follows. Consider agents $i \in \mathcal{I}$ with type $X_i - Y_i \in \mathcal{B} \times \mathcal{B}$, where X_i is the blood type of the patient and Y_i is the blood type of the donor. Their ability to participate in an exchange with $j \in \mathcal{I}$ is by one of two modes. The first mode, m_1 is the standard exchange mode, where $i \to_{m_1} j$ if j's donor is able to donate to i's patient. That is $X_i \leq_{\mathcal{B}} Y_j$, where the blood type order $\leq_B$ is shown in Figure 2. The second mode is



(a)

$\mathcal{I}$	Patient	Donor 1
K_1	A	A
K_2	B	A
K_3	B	B
K_4	A	B

(b) Observable characteristics

Figure 4: Isolated component, and example observable characteristics. Assume all agents in the component have preferences $K \succ L \succ \emptyset$, that is they find all donations acceptable.

m_2, which allows i's patient to undergo desensitization in order to become compatible with j's donor. Note that by doing so, a patient is compatible with all donors. However, this is less desirable than compatiblity via the non-desensitization mode, as noted in Andersson and Kratz (2020).

The authors also observe a connection between the model of kidney donation with desensitization and that of liver exchange as studied in Ergin et al. (2020). However the following example highlights how the application of the same mechanism will not obviously preserve the same properties. The approach of Ergin et al. (2020) utilizes the existence of a specific acyclic directed graph, which is to draw an edge from i to j if i can interact via m_1 but not m_2[14]. Again we can consider the induced graph as in Ergin et al. (2020), where a directed edge from i to j means that j's donor is compatible with i, but j can only receive a donation from i if they undergo desensitization. Alternatively, we can frame this as j is compatible with i but i is not compatible with j, since by being desensitized they are able to receive a donation from any blood type. The example in Figure 4 shows that the graph is not always acyclic.

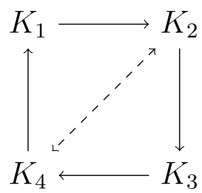
However, we can observe that once all donation possibilities via regular, that is without desensitization, pairwise exchanges have been exhausted and we've removed those agents, we would have removed K_2 and K_4. Thus the remaining graph is acyclic. In the following result, we show that in general the preference digraph is weakly acyclic but not acyclic by showing that only cycles similar to that in Figure 4 can exist.

Proposition 3. *The digraph is always weakly acyclic, but not always acyclic.*

Proof. See Appendix A.6. □

As a result, by Theorem 1 we are able to apply the mechanism of Ergin et al. (2020) to this setting, and thus achieve desirable outcomes.

[14]Note that since it is the patient that undergoes desensitization, we frame the modality as an interaction rather than what they donate.



$\mathcal{I}$	Patient	Donor 1	Donor 2
K_1	A	A	B
K_2	A	B	A
K_3	B	B	O
K_4	B	A	B

Figure 5: Isolated component, and example observable characteristics. Assume all agents in the component have preferences $K \succ L \succ \emptyset$, that is they find all donations acceptable.

4.3 Exchange with Two Risk-Ordered Donors

In this section, we interpret the modes of interaction with another agent as having multiple donors by which one can donate the corresponding organ. We consider the case where different donors face different, objective medical risks. Hence they can be ranked in order of the risk imposed on them by undergoing the surgery. We assume the agent's preference over which donor donates is reflected in this risk, and without loss we assume that the first (listed) donor is preferred over the second donor in terms of who should donate. This model is studied for kidney and liver exchange, and we show a possibility result with the former and an impossibility result for the latter.

Kidney Exchange. We now consider the standard kidney exchange model, but allow for an agent to list two donors rather than one. This occurs in practice. An agent is a triple composed of a patient and two donors. We let an agent i's type be given by $X_i - Y_i, Z_i$, where X_i is the patient's blood type, and Y_i and Z_i are the blood types of the first and second donor respectively.

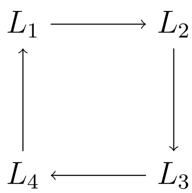
Our modes of donation are m_1 and m_2, where donation via m_l from agent i to j means that i's l-th donor can donate to j. Hence a directed edge in our digraph $i \to j$ can be interpreted as i's first donor being compatible with j's patient, but only j's second donor is compatible with i.

Figure 5 gives an example where there is a cycle, similar to the example in Figure 4. The following result shows that although this digraph can contain a cycle, the digraph is always weakly acyclic.

Proposition 4. *The digraph is weakly acyclic but not acyclic.*

Proof. See Appendix A.7. □

The setting of objective preferences over donors is not without contention. In particular,



$\mathcal{I}$	Patient	Donor 1	Donor 2
L_1	$A, 1$	$AB, 2$	$B, 1$
L_2	$AB, 2$	$A, 2$	$A, 1$
L_3	$A, 2$	$O, 1$	$O, 2$
L_4	$B, 1$	$O, 1$	$O, 2$

Figure 6: Counterexample for liver exchange with multiple donors.

there may be other considerations beyond medical risk, such as different donors having different abilities to take time off work due to familial obligations. In such a case, this is not reflected in observable medical risks.

Liver Exchange (Left-Lobe Only). Could we extend this to other organ exchanges, such as liver exchange? The model is as follows. Let $\mathcal{S} = \{1, ..., S\} \subseteq \mathbb{R}_+$ be a set of lobe sizes[15]. An agent i's type is $X_i - Y_i - Z_i$ where $X_i, Y_i, Z_i \in \mathcal{B} \times \mathcal{S}$ refer to the blood and size type of the patient, the first donor, and the second donor respectively. As before, we assume the donors are ordered by risk. Let $X^{\mathcal{B}}$ and $X^{\mathcal{S}}$ refer to the blood type and size of X. Y can donate to X if they are blood and size compatible:

1. $X^{\mathcal{B}} \leq_{\mathcal{B}} Y^{\mathcal{B}}$, and
2. $X^{\mathcal{S}} \leq Y^{\mathcal{S}}$

The model of liver exchange is similar to that of kidney exchange except with the addition of a size compatibility requirement. By choosing the size of all patients and donors in an example to be identical, we can replicate any cycles in some kidney exchange with two donors problem in this liver setting. However, the following result highlights that this setting is **not** weakly acyclic.

Proposition 5. *There is no desirable mechanism for liver exchange with two risk-ordered donors.*

Proof. Given the example in Figure 6, which provides a simple cycle, we can use Theorem 2 to conclude that there is no desirable mechanism. □

Since this is a negative result in a restricted setting, that is of a known common preference ordering that reflects objective risk, we then have an impossibility result in the general environment with potentially differing preference orderings as well.

[15]This is assumed to be discrete as in Ergin et al. (2020). This is without loss of generality as there are finitely many agents and their sizes are observable, hence $\mathcal{S}$ can be chosen given a set of agents.

5 Multi-Modal Exchanges

In this section, we study how to go beyond dual-mode exchanges. Our main application is in studying risk-ordered exchanges, where there is an ordered set of different organ exchanges, each potentially with multiple modes. As before, we assume agents share a common preference ordering consistent with the exchange ordering over these modes, which is motivated by objective medical risks being identical across agents and the determinant of their preferences. By placing restrictions on the compatibility structure through properties we call *separability* and *partition separability*, we are able to model risk-ordered exchanges and identify desirable mechanisms. We conclude this subsection by formally describing the application of integrating multiple organ exchanges.

5.1 Partition Separability

We say G_C is **partition separable** if there exists a partition $\{\mathcal{M}_k\}_{k=1}^K$ of $\mathcal{M}$ such that

1. $\mathcal{M}_k$ is **contiguous**, that is if $m_a, m_b \in \mathcal{M}_k$ then for all c such that $a \leq c \leq b$, $m_c \in \mathcal{M}_k$, and
2. there exists $\{A_k\}_{k=1}^K$ such that $\mathcal{I} = \uplus_{k=1}^K A_k$ and for all $k \in \{1, \ldots, K\}$ and $m \in \mathcal{M}_k$, $i \rightarrow_m j$ only if $j \in A_k$.

Furthermore, if $|\mathcal{M}_k| = 1$ for all $k \in \{1, \ldots, K\}$, then we say G_C is **separable**.

Contiguity allows us to say, under a common preference assumption, that the partition is contiguous with respect to this preference. The second condition states that in order to donate via mode in some partition element k, the agent receiving via that mode must belong to the associated partition of agents. Here, we have a connection between acyclicity and partition separability:

Proposition 6. *If G_C is partition separable, then for $k \neq k'$, G is acyclic with respect to any $m_k \in \mathcal{M}_k$ and $m_{k'} \in \mathcal{M}_{k'}$.*

Proof. See Appendix A.3. □

This result finds that between two modes belonging to different partitions, the compatibility graph restricted to those modes ensures that every preference digraph is acyclic (irrespective of which mode is m_1 or m_2).

For $k \in \{1, \ldots, K\}$, fix ϕ_k a mechanism with respect to $G_C(A_k)$. We will now define a mechanism on the G_C in terms of ϕ_k. Assume G_C is partition separable, and let $\{(\mathcal{M}_k, A_k)\}_{k=1}^K$ be described as above. Fix a preference profile $\succ \in \mathcal{R}^N$, and let ψ be a mechanism such that given $\succ$, it maps to the output of the following algorithm. Let $M = \emptyset$.

1. Process through $k \in \{1, \ldots, K\}$:
   (a) Remove unwilling $\mathcal{M}_k^1$ agents: for each agent $i \in \mathcal{I}$
   $$\emptyset \succ_i \mathcal{M}_k^1 \implies \mathcal{I} \leftarrow \mathcal{I} - \{i\}$$
   (b) Consider $\bar{G}_k = G_C^{\mathcal{M}_k}(\mathcal{I} \cap \cup_{l \in \mathcal{M}_l} A_l)$. Find a matching via ϕ_k:
   $$M \leftarrow M \cup \phi_k(\succ, \bar{G}_k)$$
   (c) Remove matched agents: $\mathcal{I} \leftarrow \mathcal{I} - \mathcal{I}(M)$.
   (d) Process through $l \in \{1 + \sum_{p=1}^{k-1} |\mathcal{M}_p|, \ldots, \sum_{p=1}^{k} |\mathcal{M}_p|\}$:
      i. Process through $o \in \{1 + \sum_{p=1}^{k} |\mathcal{M}_p|, \ldots, N\}$
         A. Remove unwilling m_l agents from $A_o \cap \mathcal{I}$: for each agent $i \in A_o \cap \mathcal{I}$
         $$\emptyset \succ_i m_l \implies \mathcal{I} \leftarrow \mathcal{I} - \{i\}$$
         B. Consider $\bar{G}_{l,o} = G_C((A_l \cup A_o) \cap \mathcal{I})$. Find a maximum bipartite match between A_l and A_o within $\bar{G}_{l,o}$:
         $$M \leftarrow M \cup \texttt{BipartiteMatch}(\bar{G}_{l,o}|\Pi, A_l, A_o)$$
         C. Remove matched agents: $\mathcal{I} \leftarrow \mathcal{I} - \mathcal{I}(M)$.

Intuitively, the mechanism operates as follows. First assume for simplicity that G_C is separable. We order agents by risk, treating those that require lower risk donations to have higher priority. In the first stage, we compute an efficient match between all agents in the highest priority class. In doing so, these agents get their best option, as they not only receive the organ required but also donate via the least risky mode. We proceed to prioritize these agents by considering an individually rational bipartite match between agents in the highest priority class with those in the second highest priority class. From the perspective of agents in the highest priority class, as all opportunities to match with another top priority class agent have been exhausted (since the match in the first stage was efficient), then this is their second best option. On the other hand, agents in the second highest priority class that are matched get their best choice. As we repeat this procedure of bipartite matching until we reach the lowest priority class, though we have prioritized the highest class, each class on the other side of the bipartite matching get their best option. Thus they cannot be improved upon. At this point, we have exhausted all feasible matching opportunities for the highest class, and we can repeat this procedure by replacing the highest with the second highest class in order to exhaust their opportunities. In the case where G_C is partition separable, a similar idea applies, except when applying ϕ_k it is over multiple modes (in $\mathcal{M}_k$) and thus we do not need to do bipartite matchings within this step.

As the mechanism matches agents via different modes in a way associated with the common preference structure, this allows us to find a mechanism with ideal properties:

Theorem 3 (Partition Separability is Sufficient)**.** *Suppose G_C is partition separable and for each $k \in \{1, \ldots, K\}$, with respect to $G_C(A_k)$, there exists ϕ_k that is individually rational.*

1. *If ϕ_k is Pareto efficient for each k, then ψ is Pareto efficient.*
2. *If ϕ_k is strategy-proof for each k, then ψ is strategy-proof.*
3. *If ϕ_k is weak-core stable for each k, then ψ is weak-core stable.*

Proof. See Appendix A.8. □

In the case where G_C is **separable**, we can directly construct a mechanism:

Corollary 1. *If G_C is separable, then there exists a desirable mechanism.*

Proof. For every exchange, we can observe that `MaxMatch` is efficient, strategy-proof and IR as preferences are effectively over being matched or unmatched. Thus we can directly apply Theorem 3 with $\phi_k =$ `MaxMatch`. □

We state our result for each criteria separately to emphasize the following. Although impossibility results abound, such as with liver exchange with multiple donors, we may be willing forgo certain criteria in different parts of our integrated exchange. By separately specifying that each property can hold independently, it might be possible for the exchange to satisfy certain properties like efficiency everywhere, but for example lose strategy-proofness with respect to certain modes. How to choose which properties to give up can be determined by what is necessary in practice. For example, when certain organ groups may have high urgency in getting a transplant due to the lack of alternatives[16], we may expect that strategizing is unlikely and thus choose to not require strategy-proofness.

Partition separability is not necessary. For example, consider the case where there are only two partitions with a single mode each, which corresponds to a dual-mode environment. Weak acyclicity is weaker than partition separability, which is stronger than acyclicity, hence it is not necessary for the existence of desirable mechanism. In the subsequent section, we identify a class of environments where this structure holds.

[16]For example, patients requiring a liver do not have an equivalent option to dialysis for kidney patients.

Robustness to New Technologies. The utility of this formulation is in its robustness to future medical developments. Given that new technologies in medicine are continuously being developed, allowing for novel donation modalities, and thus more ways by which individuals can receive or donate an organ, this poses an issue in paired exchange when risks differ across modes. We have already identified how this occurs between exchanges, for example kidney and liver, and within exchanges, such as with left and right lobe liver donation. Market designers working in this domain must often attend to the specific structure of the problem, such as biological compatibility, to create desirable matching mechanisms. By allowing designers to focus on individual organ exchanges and, under the assumption of a common risk-order, not on the integration of multiple organ exchanges, our mechanism can easily develop alongside new technologies. An example of early stage research on new donation modes includes intestine and pancreas transplants. Though currently not commonly done due to the increased donor risks, should the technology become sufficiently safe for donors and effective for patients, it is likely to be objectively riskier for donors than donating a kidney or liver. As such, this would satisfy assumptions within our model. Future work should study how to relax our risk-ordering assumption, which is less likely to hold as more modes are introduced. The following section describes major applications where our separability and partition separability structures are satisfied.

5.2 Application: Multiple Organ Exchanges

Where does this partition separability appear? In this section we show that this property naturally emerges in the integration of exchanges for different organs. Let a family of organ exchanges $\{\mathbb{E}_\alpha\}_{\alpha\in\mathcal{A}}$ for some ordered finite set $\mathcal{A}$ induce an exchange $\mathbb{E}_\mathcal{A}$ as follows:

1. An agent i belonging to $\mathbb{E}_\alpha$ with type T_i has new type $T_i - \alpha$ in $\mathbb{E}_\mathcal{A}$.

2. The set of modes in $\mathbb{E}_\mathcal{A}$ is $\mathcal{M}_\mathcal{A} = \cup_{\alpha\in\mathcal{A}}\mathcal{M}_\alpha$.

3. If an agent i can donate to an agent j via $m \in \mathcal{M}_\alpha$, then $j \in \mathbb{E}_\alpha$.

4. If $m \in \mathcal{M}_\alpha$ and $m' \in \mathcal{M}_{\alpha'}$ for $\alpha > \alpha'$, then $m \succ m'$.

We interpret α as the unique organ required by a patient in exchange $\mathbb{E}_\alpha$. To understand the third condition, we can note that in this context, a mode from $\mathcal{M}_\alpha$ can be thought of as a way of donating an organ α. As patients belong exactly to the exchange corresponding to the organ they require, then this condition must always hold. The following straightforward corollary identifies how this integrated exchange fits into the requirements of mechanisms we have designed.

Proposition 7. *Consider this family of organ exchange problems $\{\mathbb{E}_\alpha\}_{\alpha\in\mathcal{A}}$, where $\mathcal{A}$ is a finite ordered set. Let G be the compatibility graph induced by $\mathbb{E}_\mathcal{A}$.*

1. *If each exchange $\mathbb{E}_\alpha$ has a single mode for all $\alpha \in \mathcal{A}$, then G is separable.*

2. *If each exchange $\mathbb{E}_\alpha$ has (potentially) multiple modes for all $\alpha \in \mathcal{A}$, then G is partition separable.*

Proof. Contiguity is satisfied by how $\mathbb{E}_\mathcal{A}$ is constructed. Let A_k to be given by the set of agents in $\mathbb{E}_\alpha$, where α is the k-th element of $\mathcal{A}$. Similarly defined $\mathcal{M}_k$. Let $\mathcal{I} = \uplus_{k=1}^K$ for $K = |\mathcal{A}|$, and consider $i, j \in \mathcal{I}$ such that $i \to_m j$ for $m \in \mathcal{M}_k$ for some k. By the definition of $\mathbb{E}_\mathcal{A}$, we have that $j \in A_k$. Thus G is partition separable. It is clear that if $|\mathcal{M}_k| = 1$, we have that G is separable. □

This observation allows us to see that when families of organ exchange problems have a common ordering, then our previous results highlight the existence of a desirable mechanism.

Illustrative Example. A simple environment we can study would be integrating kidney and liver exchanges:

Example 2 (Kidney-Liver Exchange)**.** An agent i's type is given by $X_i - Y_i - O_i$ where $X_i, Y_i \in \mathcal{B}\times\mathcal{S}$ is the patient and donors blood-size type (as in the liver exchange model), and $O_i \in \{K, L\}$ refers to the organ required by the patient. There are two modes of donation, $m_1 = K$ and $m_2 = L$. Hence $i \to_{m_1} j$ if $X_i \leq_\mathcal{B} X_j$ and $O_i = K$, and $i \to_{m_2} j$ if $X_i \leq_\mathcal{B} X_j$, $Y_i \leq Y_j$ and $O_i = L$. △

Our mechanism has the following simple structure when applied to Example 2. The following is an informal description:

1. Compute a maximum match between all kidney agents (and remove all matched agents)

2. Compute a maximum bipartite match between kidney and liver agents (and remove all matched agents)

3. Compute a maximum match between liver agents.

A motivation for integrating exchange pools is to improve the number of transplants and reduce the risks taken by donors. We can characterize this by comparing our mechanism with a baseline efficient matching mechanism for each exchange pool separately.

For $o \in \{K, L\}$, let b^o be an efficient matching mechanism for agents in $\mathcal{I}_o$ that uses some priority order. Denote the joint mechanism by b, and let f denote our kidney-liver mechanism. We consider the following metrics. For $R \in \mathcal{R}$, $A \subseteq \mathcal{I}$, $D \subseteq \{K, L, \emptyset\}$, and ϕ a mechanism, $\mathcal{I}_\phi : \mathcal{R} \to \mathbb{R}$ and $N_\phi : \mathcal{R} \to \mathbb{R}$ are defined as

$$\mathcal{I}_\phi^{A,D}(R) = \{i \in A | \phi_m[R](i) \in D\}$$

and $N_\phi^{A,D}(R) = |\mathcal{I}_\phi^{A,D}(R)|$. In words, $\mathcal{I}_\phi^{A,D}(R)$ is the set of agents in A who are matched via some mode $m \in D$. When $A = \mathcal{I}$ or $D = \{K, L\}$, we will suppress reference to these variables. When comparing mechanisms, we consider the following criteria:

1. If for all $R \in \mathcal{R}$, $N_\phi(R) \geq N_\psi(R)$, then we say that ϕ **weakly increases the number of transplants** over ψ.

2. If for all $R \in \mathcal{R}$,

   (a) $N_\phi^{\mathcal{I}^L,L}(R) \leq N_\psi^{\mathcal{I}^L,L}(R)$ (less donors in the liver pool donate livers),

   (b) $N_\phi^{\mathcal{I}^L,K}(R) \geq N_\psi^{\mathcal{I}^L,K}(R)$ (more donors in the liver pool donate kidneys), and

   (c) $N_\phi^{\mathcal{I}^K,K}(R) \geq N_\psi^{\mathcal{I}^K,K}(R)$ (more donors in the kidney pool donate kidneys),

   then we say that ϕ **weakly reduces (unnecessary) donor risks** over ψ.

To interpret the second comparison criteria, we say that unnecessary donor risks are reduced when there are less liver donations and more kidney donations from pairs in the liver pool. We use the phrase unnecessary as indicating that more liver pairs can receive a liver while undergoing a safe donation. Furthermore, note that kidney patients only donate livers after exhausting their kidney donation opportunities, hence any liver donation on their part is necessary.

Proposition 8. *Let f and b use the same priority order. Then f weakly increases the number of transplants and reduces donor risks over b.*

Proof. See Appendix A.9. □

Clearly there are environments where our claim holds strictly. Further analysis on kidney-liver exchanges, including impossibility results, can be found in the Supplementary Appendix (See Appendix B).

Integrating Future Exchanges. One of the benefits of our mechanism is how adaptable it is to new exchanges as they arise. The following example considers an environment where living donor pancreas exchange has become more common. Though living pancreas donation is relatively rare, it is becoming more common due to the lack of deceased pancreases. As such, it is plausible that there is a future where such an exchange is possible. Due to the comparative difficulty of living pancreas donation, it is reasonable to assume that it is the most risky of kidney, left, and right lobe donation. Using our results, the existence of desirable mechanism in this environment with four modes can be readily established. In particular, both kidney and pancreas exchange, as they are a single mode, have as a desirable mechanism the maximum matching mechanism. Furthermore, Ergin et al. (2020) provide a desirable algorithm for liver exchange. Hence we have the following straightforward result that leverages Theorem 3:

Corollary 2. *There is a desirable mechanism for a joint kidney, liver and pancreas exchange.*

We can similarly consider other combinations including integrating kidney exchange with two donors and a simultaneous kidney-liver transplant exchange (Ergin et al., 2017)[17].

5.3 Simulations

In this section, we study in simulation the welfare impact of our proposal to integrate more donation modes into paired donor exchange. Our main welfare metric of interest when evaluating a matching is the number of pairs matched. More sophisticated mechanisms may be able to account for other welfare-relevant criteria, however that is beyond the scope of this work. We primarily focus on integrating kidney and liver markets, and for simplicity focus on left-lobe only donation.

We consider aggregate population statistics that determine biological compatibility from South Korean patients, detailed in Table 1, as in Ergin et al. (2020). For the same reason as these authors, we consider this population due to the country being a world leader in living liver transplantation. This allows us to illustrate the gains from integration in a realistic setting.

To construct our simulated population, we randomly sample n kidney (patient-donor) pairs and m liver pairs according to the liver-to-kidney patient ratio listed in Table 1. We only consider patients that are incompatible with their donors, and their sexes are drawn randomly from the listed probabilities. We also assume tissue-type compatibility between all agents for simplicity. Given this, their blood types and relevant sizes are drawn randomly, the latter of which is drawn from a normal distribution with the mean and standard deviation as given. Finally, there is some independent probability p of an agent being willing to donate a liver. For different choices of n[18] and willingness probability p we give a heatmap of the

[17]Note that in this case, the exchange is not balanced in the sense that one side transplants more organs than the other in the paired exchange. This is not done in practice, and has primarily been studied in early-stage theoretical work (Feigenbaum, 2021).

[18]We use population sizes similar to Roth et al. (2007).

Parameter	Value
Blood Type Probabilities (ABO)	O: 0.37 A: 0.33 B: 0.21 AB: 0.09
Mean Sizes (cm)	Female (F): 157.40 Male (M): 170.70
Standard Deviation of Sizes (cm)	Female (F): 5.99 Male (M): 6.40
Sex Probabilities	Female (F): 0.3 Male (M): 0.7
Liver to Kidney Ratio	$6274/32016 \approx 0.196$

Table 1: Configuration Parameters from Ergin et al. (2020).

relative increase (Figure 7) and absolute increase (Figure 8) of our integrated mechanism over the baseline mechanism, i.e. maximum matches in each organ market separately. We report the average values over 100 random simulations for these metrics. More details on the results, such as the improvements for kidney and liver groups separately, can be found in section A.10 of the Appendix. Simulation code is publicly available at `https://github.com/AnandS29/multi-modal`.

aWe note some qualitative observations from our results. First note that, for an given number of kidney pairs n, the absolute and relative increase in the total matches is increasing in the willingness probability p. This is intuitive as when there are more agents who are willing to donate a liver, then there should be more opportunities for matches to be executed. However, this intuition doesn't hold for any fixed p when varying n in the case of the relative increase. This can be explained by noting that as the size of an exchange grows, the likelihood of finding a compatible match increases. Thus more agents are matched in larger exchanges, reducing any relative gains from integration.

6 Discussion

6.1 Larger Exchanges.

As we note in Section 3, though pairwise exchanges are commonly studied when considering novel environments, exchanges with large-cycles are practiced. Gains may be diminishing for sufficiently large exchanges (Roth et al., 2007), however it is plausible that improvements

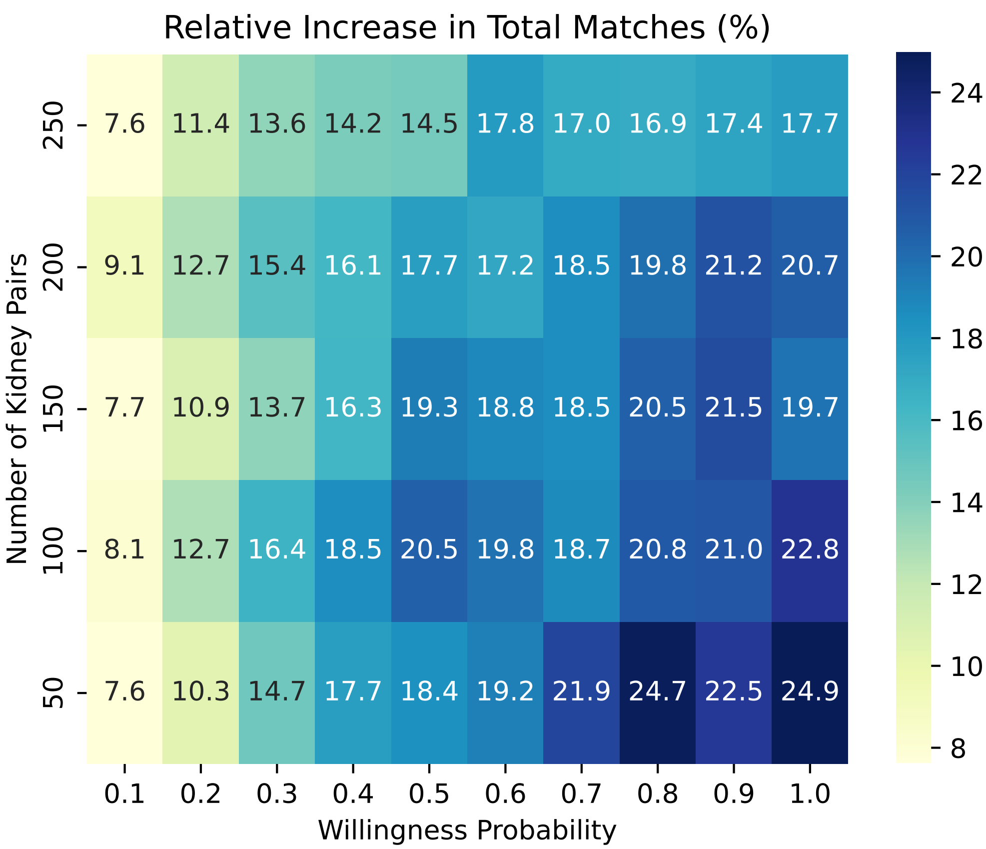


Figure 7: Relative Increase in Transplants

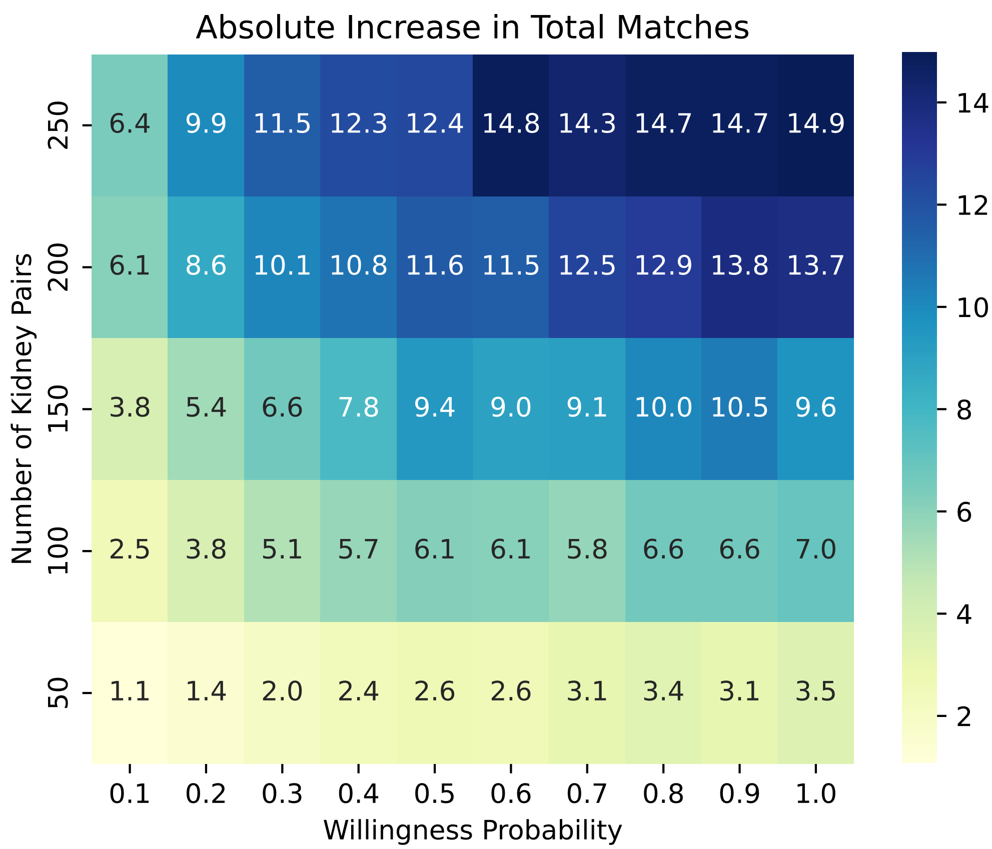


Figure 8: Absolute Increase in Transplants

in metrics such as the number of transplants can outweigh the logistical cost of large cycles. Furthermore, we study the most minimal setting, that is of pairwise exchange, hence small increases are likely to realize in net gains. One of the main challenges with this approach comes from a theoretical and algorithmic standpoint. When attempting to maximize the size of a pairwise exchange, which corresponds to having as many transplants via pairwise exchange as possible, it is very computationally tractable as it corresponds to finding a maximum matching in a graph[19]. However hardness results abound when requiring bounded cycles of length greater than two, posing a challenge for this extension[20]. Approaches based in approximation techniques have been studied (Abraham et al., 2007), however their interaction with strategic behavior, efficiency, and stability as we study is an understudied area that holds promise for future work.

6.2 Preferences over Received Organs.

One may contend that in practice, pairs have preferences not only over what they donate, but who donates to them. The main reason for this is likely to be that patients prefer to receive from healthier donors. In general, it is reasonable to assume that all patients have the same observable preference over the organ received. As our work aims to present a general model for multi-modal exchanges, we simplify these aspects and view this as an extension of models with dichotomous preferences like in Roth et al. (2005). Furthermore, in the context of kidneys, it has been noted that the recommendation of surgeons is to be indifferent over living donor kidneys (Roth et al., 2005; Yılmaz, 2011). As such, it is not clear to what degree such preferences exist. We have motivated that different donation modes can have different adverse consequences for donors, and thus we believe that our insights are valuable in practice albeit not a complete model.

Should we nevertheless want to consider preferences over organs received, a useful assumption in the literature is that such preferences are observable. For example, Ergin et al. (2020) assume the existence of a "public information received-graft preference relation", where graft refers to the organ transplanted. The assumption that this can be observed is motivated by objectivity of the quality of organs. However, formulating joint preferences over organs received and modes of donation beyond lexicographic preference (as in Ergin et al. (2020)) can make it challenging to construct desirable mechanisms, and thus we leave this for consideration in future work.

[19]This is well-known to be possible in polynomial time.

[20]Holyer (1981) shows that finding a maximum cycle packing with a restriction to cycles of length 3 is NP-hard, a classic notion from computer science that characterizes the difficulty of finding solutions. This result applies to any length greater than 2 and strictly less than the number of nodes.

6.3 Weak-Core Stability.

Why is weak-core stability an attractive property in our environment, even though it is not commonly studied in the organ exchange literature? In practice, kidney paired exchange is not completely centralized as it requires participation by hospitals and is geographically separated. As a result, Ashlagi and Roth (2014) study a setting where there are strategic incentives for hospitals to misreport the set of available patients to the centralized mechanism. This can occur because hospitals have an incentive to maximize the number of their own patients that receive a transplant, and thus they can choose to internally match pairs that are "easier" to match, and use the centralized mechanism as a means to match their "harder" to match patients. Given this, we can view hospitals as coalitions of patient-donor pairs with the potential for deviating. The weak-core helps to mitigate these incentives, however it is insufficient as a solution to this problem because it requires a strict improvement for all agents. This suggests an avenue for further study on leveraging properties given by different core notions as a solution.

6.4 Dual Equipoise and Asymmetric Risks.

An ethical concern with our model stems from two potential issues. One is referred to as dual equipoise[21], which maintains that risks from a donor donating must be balanced with the gains that their patient indirectly receives from being able to participate in paired exchange. As we allow for multiple ways of donating, there is a range of risks that vary with the donation modality. The second issue lies in the difference in risks undertaken by different donors that are paired together. For example, a donor may donate a kidney whereas another donates a liver. The gains from donation are different, and so are the risks. These issues can appear in the settings of Ergin et al. (2017) and Andersson and Kratz (2020) as well, and we find that it is necessary to consider the benefits of such approaches while accounting for their moral consequences. In some cases, integration has been practiced in the medical field (CMU, 2019). Nevertheless, it is not a common practice, and thus there is a larger conversation to be had about such tradeoffs which are beyond the scope of this paper. Nevertheless, we believe our findings will provide helpful information about the available possibilities.

7 Conclusion

In this work we consider a range of environments in paired organ exchange that share the feature of incorporating new donation or transplantation technologies. We begin with a

[21]Ergin et al. (2020) provide a helpful discussion of this in the context of left and right-lobe liver donation.

dual-mode environment - that is one with two technologies - and derive a necessary and sufficient condition for such environments to admit an efficient, stable, and strategyproof mechanism. We build on this by considering a set of exchange problems that model the integration of multiple risk-ordered exchanges, and show how independently designed, desirable mechanisms for each exchange can be themselves integrated together in a way that preserves their properties. The design of such modular mechanisms is a novel approach to scalable integration of different markets, and we see this as a promising direction for future exploration. A simulation test of our mechanism in the simple, but practical, environment of kidney-liver integration provides some foundation that such an approach can have positive welfare effects in application. Future work should explore this from an empirical lens to understand how incorporating richer preference information, along with exchange integration and new modes of donation, can improve donor participation, increase the number of transplants, and mitigate donor risk.

References

Abraham, David J., Avrim Blum, and Tuomas Sandholm (2007), "Clearing algorithms for barter exchange markets: enabling nationwide kidney exchanges." In *Proceedings of the 8th ACM conference on Electronic commerce*, 295–304, ACM, San Diego California USA, URL `https://dl.acm.org/doi/10.1145/1250910.1250954`. 30

Andersson, Tommy and Jörgen Kratz (2020), "Pairwise kidney exchange over the blood group barrier." *The Review of Economic Studies*, 87(3), 1091–1133. 2, 3, 4, 6, 10, 11, 16, 17, 31

Ashlagi, Itai and Alvin E Roth (2014), "Free riding and participation in large scale, multi-hospital kidney exchange." *Theoretical Economics*, 9(3), 817–863. 3, 31

Ashlagi, Itai and Alvin E Roth (2021), "Kidney exchange: An operations perspective." *Management Science*, 67(9), 5455–5478. 10

Balbuzanov, Ivan (2020), "Short trading cycles: Paired kidney exchange with strict ordinal preferences." *Mathematical Social Sciences*, 104, 78–87. 6

CMU (2019), "Computer science idea triggers first kidney-liver transplant swap." `https://www.cmu.edu/news/stories/archives/2019/may/kidney-liver-swap.html`. Accessed: 2024-05-26. 2, 31

Dickerson, John P. and Tuomas Sandholm (2017), "Multi-organ exchange." *J. Artif. Int. Res.*, 60(1), 639–679. 2, 7

Egawa, Hiroto, Hideki Ohdan, and Kazuhide Saito (2023), "Current status of abo-incompatible liver transplantation." *Transplantation*, 107(2), 313–325. 12

Ergin, Haluk, Tayfun Sönmez, and M Utku Ünver (2017), "Dual-donor organ exchange." *Econometrica*, 85(5), 1645–1671. 2, 6, 26, 31

Ergin, Haluk, Tayfun Sönmez, and M Utku Ünver (2020), "Efficient and incentive-compatible liver exchange." *Econometrica*, 88(3), 965–1005. 2, 3, 4, 6, 8, 10, 12, 13, 14, 17, 19, 26, 27, 30, 31, 35, 36

Feigenbaum, Itai (2021), "Unbalanced matchings in kidney exchange." *Available at SSRN 3885977*. 26

Gabow, Harold N. (2017), "The weighted matching approach to maximum cardinality matching." URL `https://arxiv.org/abs/1703.03998`. 8

Gaston, Robert S, Vineeta Kumar, and Arthur J Matas (2015), "Reassessing medical risk in living kidney donors." *Journal of the American Society of Nephrology*, 26(5), 1017–1019. 2

Gilon, Oren, Tal Gilon, and Assaf Romm (2019), "Kidney exchange with multiple donors." *Available at SSRN 3451470*. 6

Heo, Eun Jeong, Sunghoon Hong, and Youngsub Chun (2022), "The top-trading cycles and chains solution for kidney exchange with immunosuppressants." *International Journal of Economic Theory*, 18(1), 77–91. 6

Holyer, Ian (1981), "The NP-Completeness of Some Edge-Partition Problems." *SIAM Journal on Computing*, 10(4), 713–717, URL https://doi.org/10.1137/0210054. _eprint: https://doi.org/10.1137/0210054. 30

Kamada, Yuichiro and Yosuke Yasuda (2025), "Robust exchange under single-peaked preferences." *Working paper*. 6

Kumar, Rajnish, Kriti Manocha, and Josué Ortega (2022), "On the integration of shapley–scarf markets." *Journal of Mathematical Economics*, 100, 102637. 4

Lee, Sung-Gyu (2010), "Living-donor liver transplantation in adults." *British Medical Bulletin*, 94(1), 33–48. 2

Micali, Silvio and Vijay V Vazirani (1980), "An o (v— v— c— e—) algoithm for finding maximum matching in general graphs." In *21st Annual symposium on foundations of computer science (Sfcs 1980)*, 17–27, IEEE. 8

Mishra, Ashish, Alexis Lo, Grace S Lee, Benjamin Samstein, Peter S Yoo, Matthew H Levine, David S Goldberg, Abraham Shaked, Kim M Olthoff, and Peter L Abt (2018), "Liver paired exchange: Can the liver emulate the kidney?" *Liver Transplantation*, 24(5), 677–686. 2

Nicolo, Antonio and Carmelo Rodriguez-Alvarez (2017), "Age-based preferences in paired kidney exchange." *Games and Economic Behavior*, 102, 508–524. 6

Okumura, Yasunori (2014), "Priority matchings revisited." *Games and Economic Behavior*, 88, 242–249. 8

Roth, Alvin E, Tayfun Sönmez, and M Utku Ünver (2004), "Kidney exchange." *The Quarterly journal of economics*, 119(2), 457–488. 5, 6

Roth, Alvin E, Tayfun Sönmez, and M Utku Ünver (2007), "Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences." *American Economic Review*, 97(3), 828–851. 6, 10, 26, 27

Roth, Alvin E., Tayfun Sönmez, and Utku Ünver (2005), "Pairwise kidney exchange." *Journal of Economic Theory*, 125(2), 151–188, URL `https://www.sciencedirect.com/science/article/pii/S0022053105001055`. 2, 6, 10, 30, 42

Turner, Jonathan (2015), "Faster maximium priority matchings in bipartite graphs." URL `https://arxiv.org/abs/1512.09349`. 8

Watanabe, Moyuru (2022), "Multiorgan exchange: A liver-kidney paired exchange perspective." *SSRN*. 6

Yılmaz, Özgür (2011), "Kidney exchange: An egalitarian mechanism." *Journal of Economic Theory*, 146(2), 592–618. 30

A Appendix

A.1 Proof of Proposition 1

Proof. Consider a pairwise stable matching $M \in \mathbf{M}$. Assume for contradiction that M is not weak core stable. Then there is a set of agents U and a matching M' that is strictly improving for all agents in U. Let i be an agent matched in M'. Note that there must be some agent matched since M was an IR matching and all agents in U were strictly improved upon. Thus there are at least two agents i and j who are matched to each other in M'. Since i and j strictly preferred to be matched each other than their matching in M, they form a blocking pair. This contradicts the pairwise stability of M. Thus, M is weak-core stable. □

A.2 Proof of Proposition 2

Proof. Assume $G(\mathcal{I})$ is weakly acyclic. For contradiction, assume there is a cycle C in $\tilde{J}_K$ (using the notation of Ergin et al. (2020)), and thus there does not exist a topological order. Otherwise the operation `TopOrder` is well-defined and the algorithm of Ergin et al. (2020) applies. Denote G' as the graph. Since G' is a subgraph of $G(\mathcal{I})$, then all cycles in the former are also cycles in the latter. Given that $G(\mathcal{I})$ is weakly acyclic, then any cycle contains an m_1 mutually compatible pair. Hence there exists $i, j \in C \cap (\tilde{J}_K)$, and thus $i, j \notin J_K$. This is a contradiction as, assuming $i < j$ in the Π_L order, then i would not be transformed to m_2 since it would be matchable with j. Thus G' is acyclic, and there exists a topological order. □

A.3 Proof of Proposition 6

Proof. Fix $m_k \in \mathcal{M}_k$ and $m_{k'} \in \mathcal{M}_{k'}$ for $k \neq k'$. Consider the graph $G_{m_i,m_j}(\mathcal{I})$. Assume for contradiction that there is a cycle such that $C = (i_0, \dots, i_{K-1})$. For any $k \in \mathbb{N}$, $i_{k \mod K} \to i_{(k+1) \mod K}$ implies, by separability, that $i_{k \mod K} \to_{m_i} i_{(k+1) \mod K}$ and $i_{(k+1) \mod K} \to_{m_j} i_{k \mod K}$. By separability, $i_{(k+1) \mod K} \in A_i$ and $i_{k \mod K} \in A_j$. Since this is true for arbitrary k, then we have that for any $k \in \{0, ..., K-1\}$, $i_k \in A_i \cap A_j$. This is a contradiction since $\{A_n\}$ is a partition of $\mathcal{I}$, and thus $A_i \cap A_j = \emptyset$. Thus there cannot exist a cycle. □

A.4 Proof of Theorem 1

Proof. Let ϕ be the Preference Adaptive algorithm of Ergin et al. (2020). IR, strategyproofness and efficiency directly follows from the proof in Ergin et al. (2020).

Recall that individual rationality and pairwise stability are a necessary and sufficient condition for a matching to be weak-core stable. Assume for contradition that there is a preference profile $\succ$ such that $M = \phi(\succ)$ is not weak-core stable. Then there is a pair i, j and $m, m' \in \mathcal{M}$ such that i is compatible via m with j, j is compatible via m' with i, $m \succ_i M_{\mathcal{M}}(i)$, and $m' \succ_j M_{\mathcal{M}}(j)$.

Note that we do not need to consider the case where either agent receives a match via m_1, that is $M_{\mathcal{M}}(i) = m_1$ or $M_{\mathcal{M}}(j) = m_1$, as there is no means by which either agent can strictly improve their match.

Consider the case where $M_{\mathcal{M}}(i) = M_{\mathcal{M}}(j) = \emptyset$. Given that they were not matched in M, and the final step output a maximum matching, it must be that either i or j are unwilling to be matched via m_2. Without loss of generality, assume it is i. Then $m_2 \prec_i \emptyset = M_{\mathcal{M}}(i)$, so it must be that $m = m_1$. Since a maximum match via m_1 bilaterally was found in the first step, it must be that $m' = m_2$ and $i \notin \mathcal{I}_{m_1 \leftrightarrow m_1}$. Thus both agent are in $\mathcal{I} - \mathcal{I}_{m_1 \leftrightarrow m_1}$. Since $i \to j$, it must be that j is transformed first. When i is reached when processing the topological order, it must be that in M, i is removed from the graph given that they are unwilling and unmatched. However, i is matchable with j, and there cannot be any promises that cannot be any promises that restrict this as both agents are unmatched at the end of the algorithm. Hence i would be promised an exchange via m_1, which is a contradiction.

Now consider the case where $M_{\mathcal{M}}(i) = m_2$ and $M_{\mathcal{M}}(j) = \emptyset$. Since both agents strictly improve their outcome, it must be that $m = m_1$ and $m' \in \{m_1, m_2\}$. If $m' = m_1$, then i

and j are both matchable via m_1 with each other while keeping all promises in the second step. This is a contradiction as they would have both been promised a match via m_1. Now consider $m' = m_2$. Since i is not matched via m_1, they must have not been promised a match in the second step. Since j is unmatched and $i \in \mathcal{I} - \mathcal{I}_{m_1 \leftrightarrow m_1}$, then j would be transformed to an m_2 agent, since they must be willing for m' to be individually rational, and i would be matchable with them via m_1. This is a contradiction as i is not matched via m_1. Note that a symmetric argument applies for $M_{\mathcal{M}}(j) = m_2$ and $M_{\mathcal{M}}(i) = \emptyset$.

Finally, consider the case where $M_{\mathcal{M}}(i) = M_{\mathcal{M}}(j) = m_2$. Then it must be that $m = m' = m_1$. By the same logic in the previous paragraph, this is not possible. As in all cases it is impossible for there to be a blocking pair, it cannot be that there is a blocking pair. Hence the mechanism is pairwise stable, and consequently weak-core stable. □

A.5 Proof of Theorem 2

Proof. Fix $\mathcal{E} \in \mathbf{E}$. Consider a simple cycle $C = (i_1, \ldots, i_n)$. Note that it must be that $n > 2$, since otherwise we have that $i_1 \to i_2 \to i_1$ (and we assume no self-loops), which is a contradiction as i_1 can donate to i_2 via m_1 but the latter cannot donate to the former via m_2. Assume for contradiction that there is a desirable mechanism f. Note that C contains agents of distinct types. Let $\mathcal{I} = C$, and assume all agents find m_2 feasible. By efficiency, there are two agents that are matched. One case is that there are some agents i_k and i_l matched mutually via m_2, or there are some two agents succesive in the cycle that are matched. If it is the latter case, let this be i_1 and i_2. Furthermore, i_1 donates via m_1 and i_2 donates via m_2 since no agents are mutually compatible via m_1, and if they were both matched via m_2 there would be a Pareto improvement by having one agent donate via m_1 since by assumption, that is possible. If i_2 report that m_2 is not IR, then i_2 is unmatched in C' by strategy-proofness. If i_3 is unmatched, then there would be a Pareto improvement by matching i_2 and i_3. Thus i_3 must be matched with i_4. Similarly, if it is the case that there are two agents i_k and i_l matched via m_2, first note that they are non-adjacent in the cycle. Without loss let $k = 1$ and $l > 2$. By the same argument we require that if i_1 reports to be unwilling, then i_2 is matched with some other agent (possibly adjacent or not). Inductively continuing this process of progressive misreporting, we find that if i_{n-1} misreports, then i_n must be matched somehow. However as there are no m_1 mutually-compatible matches, and i_1 is not willing to match via m_2, then i_n must be matched via an m_2 mutually-compatible match. However all $i_1, \ldots, i_{n-1}$ do not find m_2 individually rational. Hence to be efficient, i_{n-1} and i_n must be matched, with the former using m_1 and the latter using m_2. This contradicts what is required by strategyproofness. □

A.6 Proof of Proposition 3

Proof. Note that if $i \to j$, then j can donate to i but i can't donate to j. For simplicity, and to be in line with previous definition of the direction of the arrow, we flip the direction of every arrow to mean that i can donate to j but j can't donate to i. This preserves all cycles. First we show that the above cycle is the only cycle in the graph. Consider some cycle $C = (i_1, i_2, ...)$. Since i_2 cannot donate to i_1 via m_1, then it must be that i_2's donor A, B or AB. Furthermore this must be true for every agent's donor since this a cycle. Since i_1's donor is A, B or AB, then consider the following cases: i_1's donor is

- A: i_2 is A or AB
- B: i_2 is B or AB
- AB: i_2 is AB

If i_2 is

- A: i_3's donor is B or AB
- B: i_3's donor is A or AB
- AB: not possible as any type can donate to AB

Hence we can conclude that no agent has type AB. Thus agents can only have type A or B, since there are no donors that can donate to O agents. If there is an AB donor, there must be an AB agent. Hence there are no AB donors either. Thus if i_1 is

- A: i_2's donor is B
- B: i_2's donor is A

Consider i_1 having patient-donor type $A - A$. Then i_2 must be either $A - B$. i_3 must then be $B - B$, i_4 is $B - A$, and i_5 is $A - A$. Since all possible types are in this cycle, and their successor is uniquely determined, this is the only cycle. Since there is a cycle, the graph is not acyclic. To see that it is weakly acyclic, note that $B - A$ and $A - B$ are mutually compatible. Hence all cycles in this graph have m_1-mutually compatible. □

A.7 Proof of Proposition 4

Proof. Consider some cycle $C = (i_1, i_2, ...)$. Since i_1 is not compatible with i_2's first donor, then i_2 must be of type A, B or O since if they were AB, then they would be compatible any donor. Since this is a cycle, then average agent is of type A, B or O. Furthermore, given that i_1's first donor can donate to i_2, then i_1's first donor must be of type A, B or O. Again, this must be true for all agents. Since i_2's first donor is not compatible with i_1, then i_2's first donor cannot be O as they are compatible with all agents. This implies that i_1's first donor cannot be O, hence i_2 cannot be O either. Thus every patient and every first donor is either A or B. Consider the patient-donor tuple $A - A - X$, where X is the the type of the second donor. If corresponded to i_1, then i_2 must be $A - B - X$. Furthermore, i_3 will be $B - B - X$, i_4 will be $B - A - X$, and i_5 will be $A - A - X$. All cycles must have this form as any agent will have a type in this cycle, and their successor will have the same form. Clearly such a cycle exists (as given by the previous example), hence the digraph is acyclic. To see that this is weakly acyclic, note that i_2 and i_4 are m_1-mutually compatible since they can donate to each other via their first donor. Hence all cycles in this graph have m_1-mutually compatible agents. □

A.8 Proof of Theorem 3

Proof. Clearly the mechanism is individually rational.

We now prove Pareto efficiency. Assume ϕ_i are efficient and individually rational for all i. Assume for contradiction that there is a matching M' that Pareto dominates the $M = \psi(\succ)$. Let i be such that $M'_{\mathcal{M}}(i) = m' \succ m = M_{\mathcal{M}}(i)$. We proceed by strong induction on $\mathcal{I}_{m_k \leftrightarrow m_l}$ - the set of agents matched via an m_k-m_l swap for m_k and m_l in different elements of the partition $\{\mathcal{M}_i\}$ -, $\mathcal{I}_{\mathcal{M}_k}$ - the agents in A_k matched under ϕ_k, and $\mathcal{I}_{\emptyset \leftarrow \mathcal{M}_k}$ - the set of agents in A_k that were unmatched in M -, to show that no i in any of these set can be strictly improved upon. Consider the following order $\rhd$:

$$\mathcal{M}_1, (\mathcal{M}_1^1, \mathcal{M}_{>1}^1), \ldots, (\mathcal{M}_1^1, \mathcal{M}_{>1}^{-1}), (\mathcal{M}_1^2, \mathcal{M}_{>1}^1), \ldots, (\mathcal{M}_1^{-1}, \mathcal{M}_{>1}^1), \ldots,$$
$$(\mathcal{M}_1^{-1}, \mathcal{M}_{>1}^{-1}), \ldots, \mathcal{M}_k, (\mathcal{M}_k^1, \mathcal{M}_{>k}^1), \ldots, (\mathcal{M}_k^{-1}, \mathcal{M}_{>k}^{-1}), \ldots, \mathcal{M}_{-1}$$

where -1 represents the last element of a set. In other words, it is generated by the following process: $L = \{\}$

- For $k \in \{1, \ldots, |\mathcal{M}|\}$:
  - $L \leftarrow L \cup \{\mathcal{M}_k\}$
  - For $l \in \{i, \ldots, |\mathcal{M}_k|\}$
    - For $m \in \cup_{n>k} \mathcal{M}_n$:
      - $L \leftarrow L \cup \{(\mathcal{M}_k^l, m)\}$

We interpret this as the type of match one can be part of, in order that our algorithm does them in. We say that $a, b \in L$, $a \rhd b$ if a is after b in L.

We strongly induct on $a \in L$ to show that there is no agent i that can be strictly improved to match via a without making some agent worse off. Note that if i is ever strictly improved from a match a, it must be to a match b such that $a \rhd b$ or $a = b = \mathcal{M}_k$. To see why, assume the latter case does not hold. For M, let $i \to_m j$ and $j \to_n i$, and in M', $i \to_{m'} k$ and $k \to_{n'} i$. Note that $m \in \mathcal{M}_r = a$ and $m' \in \mathcal{M}_s = b$ for $s < r$. Assume that $i \in A_l$.

1. $l = s$: This means that $i, k \in A_l$ by partition separability. Observing the construction of L, we have that the "internal" match for $\mathcal{M}_l$ is done prior to the "external" matches with less preferred modes. That is $\mathcal{M}_l \rhd (m, n)$ for $n \in \mathcal{M}_r$ for $r > s$. Hence $a \rhd b$.

2. $l < s$: Note that this implies a cross partition match in M'. From the construction of L, we have that $(m', q_1) \rhd (m', q_2) \rhd (m, q_3) \rhd (m, q_4)$ for all $q_1 \succ q_2$ and $q_3 \succ q_4$ (such that these are well-defined to be in L). Hence $a \rhd b$.

3. $l > s$: This also implies a cross partition match in M'. Observe that $(q_1, r_1) \rhd (q_2, r_2)$ for all $q_1 \succ q_2$ (such that this is well-defined to be in L), and $(q, r) \rhd \mathcal{M}_t$ for $q \succ \mathcal{M}_t$. Hence $a \rhd b$.

The base case is $a = \mathcal{M}_1$. Consider i, j such that $i \to_{m_i} j$ and $j \to_{m_j} i$ in M. Assume that i can be strictly improved in a new match M', where $i \to_{m'_i} k$ and $j \to_{m'_j} l$. By partition separability, $i, j, k, l \in A_1$. Note that all agents in A_1 matched in M using a mode in $\mathcal{M}_1$ must be weakly improved in M', hence they must again be matched via a mode in $\mathcal{M}_1$ when in M'. Recall that that the match restricted to agents in A_1 corresponds to those matched using ϕ_1. However the existence of M' implies that some agent (i) can be strictly improved while all other agents in A_1 are weakly improved. This contradicts the efficiency of ϕ_1.

Fix some $a \in L$. Assume that the inductive hypothesis holds for all a' such that $a \rhd a'$. We now show that it holds for a. Assume it does not, that is there is some agent i that can be strictly improved to match from being matched via a in M' without making any other agent worse off. Let M' be such a match, and let b be the type of match i participates in within M'. From our previous observation, either $a = b = \mathcal{M}_k$ for some k, or $a \rhd b$. First assume the latter case. Because $a \rhd b$ means that the matches corresponding to b were executed earlier than the matches corresponding to a, then i must have been unmatched at the start of the step corresponding to b. Let j be the agent that i is matched with in M'. As j must be weakly improved from M in M', they similarly must be unmatched at the start of the step corresponding to b. If $b = (m, m')$ then as `BipartiteMatch` match finds a maximum (bipartite) match then they would have been matched if j was not matched at this step. Similarly, if $b = \mathcal{M}_l$ for some l then this would also hold by efficiency of ϕ_l. Thus j must have been matched to some k at this step. Due to the efficiency of ϕ_l or `BipartiteMatch`, we can find that there must be some agent originally matched at step b in M who must be strictly improved (otherwise there is a Pareto improvement at step b). This would contradict the inductive hypothesis, hence it must be that $a = b = \mathcal{M}_k$ for some k. Thus there is some i who strictly improves to donating via $m \in \mathcal{M}_k$ to $m' \in \mathcal{M}_k$, in M and M' respectively. However, again by efficiency of ϕ_k there must be some agent j' that strictly improves to $m'' \succ \mathcal{M}_k$. Let c be the match they improve to, and note that by partition separability and the fact that j' was matched in $b = \mathcal{M}_k$ (hence $j' \in A_k$), we have that $c = (n, n') \lhd a = b$. j' must have been unmatched at the step correspond to c, and by a similar argument to before we have that there is some j'' that must be strictly improved (by maximality/efficiency of `BipartiteMatch`). This contradicts the inductive hypothesis, and as this exhausts the possible cases it must be that there is no agent matched via a that can be strictly improved.

By our inductive argument, there is no agent matched via $a \in L$ that can be strictly improved in a Pareto improvement. If i were unmatched at the end of the algorithm but can be strictly improved in a Pareto improvement, we can apply the same arguments as before to show that some j originally matched via some $a \in L$ would be strictly improved, and thus find a contradiction. Hence we can conclude that no agent i can be strictly improved without making another agents strictly worse off, which means that M is Pareto efficient.

To see strategyproofness, observe that any misreport outside of the agent's partition element to a more preferred mode does not increase any agent's chance to be matched, and any misreport to a less preferred more either does not change the outcome for the agent, or causes them to be matched via a mode less preferred to being unmatched. Thus we can conclude on strategyproofness of the mechanism with respect to these deviations. Furthermore, misreports within the agent's partition element does not affect the algorithm until the the mechanism corresponding to that partition is used. Since the mechanism is strategyproof, there is no profitable deviation. Hence the overall mechanism is strategyproof.

To see pairwise stability, and thus also weak-core stability by Proposition 1, note that if two agents preferred to be matched to one another over their partner, then by virtue of the algorithm, if they were in the same partition then this would contradict pairwise stability of the corresponding mechanism, and if they were part of different partition elements then they would have been bipartite matched earlier in the algorithm. □

A.9 Proofs of Proposition 8

Proof. Let $\mathcal{I}^b_{K\leftrightarrow K}$ and $\mathcal{I}^b_{L\leftrightarrow L}$ be the agents in $\mathcal{I}^K$ and $\mathcal{I}^L$, respectively, that are matched by b^K and b^L. First note that $\mathcal{I}_{K\leftrightarrow K}$ has the same cardinality as $\mathcal{I}^b_{K\leftrightarrow K}$ as maximal matchings are maximum matchings (Roth et al., 2005). Note that we can construct f (via choice of priority order) such that a pair i in $\mathcal{I}^b_{L\leftrightarrow L}$ not matched in $\mathcal{I}_{L\leftrightarrow L}$ is such that they must be matched to a kidney patient, i.e. $i \in \mathcal{I}_{K\leftrightarrow L}$. Worst case, the matched pair of i in $\mathcal{I}^b_{L\leftrightarrow L}$ is unmatched in f, but for every such case there is a kidney pair is who is matched in f but not in the baseline. Hence there must be at least as many transplants in f then in the baselines.

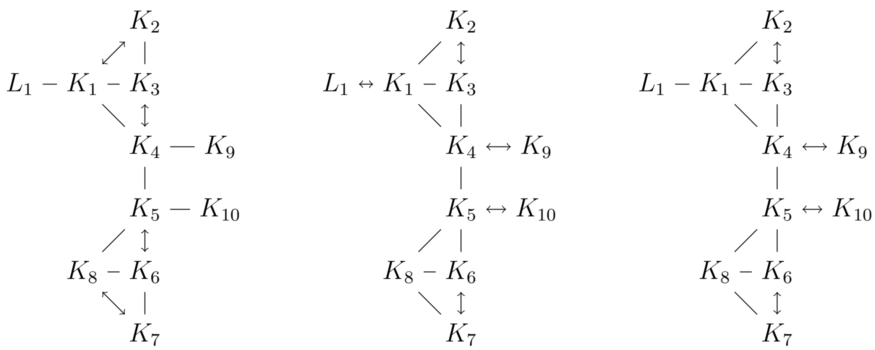
Now we show the second condition. We have already shown that more kidney patients donate kidneys in the previous paragraph. For a similar reason, less liver patients donate livers. If this did not hold, then the matching in the baseline would not have been maximum. Finally, since no liver patients donate kidneys in the baseline, the claim that more liver patients donate kidneys holds trivially. □

A.10 Simulation

In this section, we provide more comprehensive results on our simulation.

		Integrated				Baseline				Increase	
		Avg K	Avg L	Avg T	Std T	Avg K	Avg L	Avg T	Std T	Abs Δ	Rel Δ
p	n										
0.1	**50**	13.58	1.04	14.62	4.30	13.00	0.58	13.58	4.10	1.04	7.66
	100	29.67	2.99	32.66	5.10	28.02	2.28	30.30	5.35	2.36	7.79
	150	47.58	5.54	53.12	8.09	44.66	4.52	49.18	7.77	3.94	8.01
	200	64.25	9.37	73.62	8.21	59.38	7.96	67.34	8.18	6.28	9.33
	250	83.07	11.57	94.64	10.26	77.28	10.62	87.90	9.70	6.74	7.67
0.2	**50**	13.95	1.33	15.28	3.90	12.92	0.74	13.66	3.78	1.62	11.86
	100	31.53	3.73	35.26	5.81	29.02	2.66	31.68	5.82	3.58	11.30
	150	47.08	6.20	53.28	6.95	42.74	4.42	47.16	6.64	6.12	12.98
	200	64.88	9.44	74.32	9.47	58.34	7.16	65.50	9.24	8.82	13.47
	250	82.49	11.13	93.62	9.45	74.78	9.64	84.42	9.63	9.20	10.90
0.3	**50**	14.44	1.42	15.86	4.08	13.26	0.54	13.80	3.97	2.06	14.93
	100	32.69	3.85	36.54	6.22	29.52	2.38	31.90	6.36	4.64	14.55
	150	49.32	6.94	56.26	7.32	43.84	4.92	48.76	7.13	7.50	15.38
	200	66.41	9.65	76.06	9.00	59.08	7.10	66.18	8.90	9.88	14.93
	250	83.33	11.91	95.24	9.55	74.84	9.36	84.20	9.53	11.04	13.11
0.4	**50**	14.89	1.63	16.52	4.10	13.44	0.60	14.04	3.84	2.48	17.66
	100	32.50	4.66	37.16	6.35	28.84	2.72	31.56	5.82	5.60	17.74
	150	49.96	6.82	56.78	8.06	44.14	4.54	48.68	8.35	8.10	16.64
	200	67.18	9.70	76.88	9.99	59.16	7.18	66.34	9.64	10.54	15.89
	250	85.04	12.72	97.76	10.29	75.06	9.74	84.80	10.00	12.96	15.28
0.5	**50**	14.83	1.87	16.70	4.55	13.22	0.90	14.12	4.17	2.58	18.27
	100	31.34	4.40	35.74	6.35	27.60	2.30	29.90	6.04	5.84	19.53
	150	49.29	6.67	55.96	7.64	43.60	4.64	48.24	7.44	7.72	16.00
	200	67.46	10.14	77.60	9.51	59.06	7.12	66.18	8.74	11.42	17.26
	250	85.88	12.10	97.98	10.76	75.86	9.18	85.04	10.64	12.94	15.22
0.6	**50**	14.27	1.89	16.16	4.26	12.54	0.68	13.22	3.81	2.94	22.24
	100	32.21	4.45	36.66	7.13	28.40	2.52	30.92	6.74	5.74	18.56
	150	50.81	6.95	57.76	7.58	44.74	4.52	49.26	7.02	8.50	17.26
	200	67.23	9.77	77.00	8.83	58.80	7.26	66.06	8.42	10.94	16.56
	250	85.68	12.36	98.04	9.47	74.94	9.90	84.84	9.79	13.20	15.56
0.7	**50**	14.71	1.83	16.54	4.06	12.96	0.66	13.62	3.90	2.92	21.44
	100	32.31	4.55	36.86	6.31	28.30	2.32	30.62	5.84	6.24	20.38
	150	49.92	7.74	57.66	8.04	43.08	4.86	47.94	8.17	9.72	20.28
	200	68.44	10.32	78.76	9.11	59.32	7.68	67.00	8.39	11.76	17.55
	250	84.55	12.43	96.98	10.29	73.86	9.80	83.66	9.58	13.32	15.92
0.8	**50**	14.77	1.93	16.70	3.91	12.92	0.52	13.44	3.85	3.26	24.26
	100	32.60	4.84	37.44	5.88	28.34	2.50	30.84	5.46	6.60	21.40
	150	50.08	7.66	57.74	6.80	43.20	4.84	48.04	6.25	9.70	20.19
	200	68.48	10.80	79.28	9.39	58.84	7.48	66.32	8.94	12.96	19.54
	250	87.10	13.64	100.74	9.84	74.82	9.92	84.74	9.38	16.00	18.88
0.9	**50**	15.52	2.24	17.76	4.49	13.44	0.82	14.26	3.93	3.50	24.54
	100	32.69	4.81	37.50	5.95	28.30	2.66	30.96	5.74	6.54	21.12
	150	51.65	7.77	59.42	6.94	44.60	4.46	49.06	7.10	10.36	21.12
	200	70.08	10.92	81.00	8.09	60.22	7.96	68.18	7.56	12.82	18.80
	250	88.01	14.03	102.04	11.22	75.42	10.40	85.82	9.88	16.22	18.90
1.0	**50**	15.52	2.06	17.58	4.14	13.54	0.74	14.28	4.00	3.30	23.11
	100	32.90	4.64	37.54	6.41	28.64	2.24	30.88	6.20	6.66	21.57
	150	51.64	7.40	59.04	7.21	44.54	4.62	49.16	7.34	9.88	20.10
	200	70.78	10.76	81.54	9.06	61.12	7.98	69.10	9.23	12.44	18.00
	250	86.55	12.67	99.22	9.91	74.90	9.24	84.14	11.03	15.08	17.92

Table 2: Simulation results. p is the willingness probability; n is the number of kidney patient-donor pairs; K, L, T is kidney, liver, and total number of matches; Δ is difference (e.g. Abs Δ is the absolute increase from baseline to integrated).



(a) (b) (c)

$\mathcal{I}$	Patient	Donor
K_1	AB, 1	A, 6
K_2	A, 1	B, 1
K_3	AB, 1	A, 1
K_4	A, 1	O, 1
K_5	B, 1	O, 1
K_6	AB, 1	B, 1
K_7	B, 1	AB, 1
K_8	AB, 1	B, 1
K_9	O, 1	A, 1
K_{10}	O, 1	B, 1
L_1	AB, 5	A, 3

(d) Example types

Figure 9: Isolated component, and example types. Assume all agents in the component have preferences $K \succ L \succ \emptyset$, that is they find all donations acceptable. – indicates biological feasibility, and $\leftrightarrow$ indicates edges in a matching.

B Supplemental Appendix

In this section, we describe additional results on the integration of kidney and liver exchanges.

B.1 (Non-)Uniqueness

Is our mechanism unique? Consider the following mechanism g:

1. Identify components in $\mathcal{G}$ as in Figure 9d and match all agents according to Figure 9b if individually rational, and Figure 9c otherwise.

2. Denote matched agents as $\mathcal{I}_{K\leftrightarrow L}$ and remove them from them for $\mathcal{I}$: $\mathcal{I} \leftarrow \mathcal{I} - \mathcal{I}_{K\leftrightarrow L}$.

3. Match $\mathcal{I}$ according to f.

Intuitively, this mechanism matches all subsets of agents that belong to the isolated component and have the described structure, and for the remaining agents implements f. Observe that the matches specified in Figure 9a is a maximum match that would arise in our proposed mechanism f.

Theorem 4. *g is a distinct mechanism from f that satisfies PE, IR, SP, and weak-core stable.*

Proof. Clearly g is individually rational. It is distinct from f by observing that, in the isolated component, the kidney pairs would be matched by f instead of to the liver pairs as in g. Consider agents in the isolated component, if any. If it is individually rational for both pairs of K-L pairs to be matched with one another, then g will match them. Otherwise, they are not matched at this stage and they will be potentially matched under f. Observe that as the L pairs get their best outcome, they have no incentive to misreport. If the K pair misreports to say they are unwilling to match with a L pair, then they will remain unmatched since they belong to an isolated component (hence are mutually compatible with no other pair outside of the component) and their only potential match would be the other K pair, which is already matched. Thus they would remain unmatched, which is strictly worse for them according to their true preferences. Since f is IC, this whole mechanism is IC. The mechanism is PE as the pairs in the isolated component who can be improved, that is the K pairs, can only be improved by unmatching with the L pairs, who have no other matching possibilities. Since f is PE, the whole mechanism is PE. To see pairwise stability, observe that in the isolated component the only agents that K_1 can't form a blocking pair with any agent K_i since they are all matched to some K_j and thus would not be strictly improved upon. □

B.2 Impossibility Results

In this section we state some impossibility results in various cases. We collate the results in the following theorem, whose details are explained subsequently.

Theorem 5. *Consider a pairwise mechanism that is strategyproof, Pareto efficient, individually rational, and pairwise stable. Then it cannot*

1. *hold for heterogeneous preferences,*
2. *maximize the number of transplants, or*
3. *satisfy neutrality*

Heterogeneous Preferences and Pairwise Exchanges. Our solution thus far relies on this assumption of a common risk ordering. A natural question is whether this assumption was necessary. That is, can we find an efficient, IR and strategyproof mechanism when preferences over transplants can be arbitrary? Our following result shows that this is not possible:

Proposition 9. *Let preferences be arbitrary and only allow for pairwise exchanges. Then there is no efficient, IR and strategyproof mechanism.*

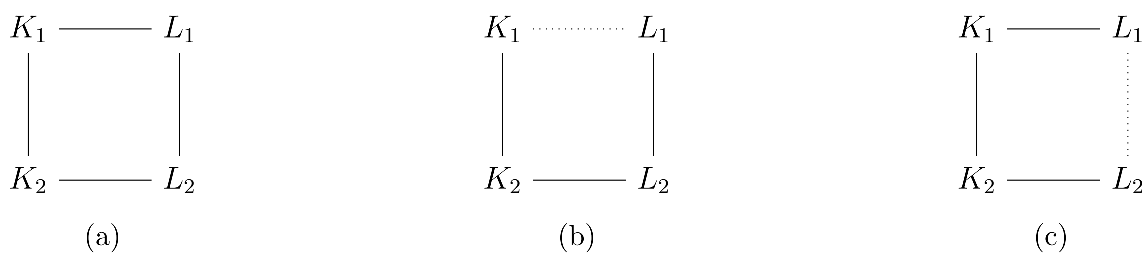


Figure 10: Example compatibility graph observable characteristics.

Proof. Consider Figure 10a. Let the pairs have the following preferences:

$$L \succ_{K_1} K \succ_{K_1} \emptyset$$
$$K \succ_{K_2} L \succ_{K_2} \emptyset$$
$$L \succ_{L_1} K \succ_{L_1} \emptyset$$
$$K \succ_{L_2} L \succ_{L_2} \emptyset$$

There are two possible matchings: $\{(K_1, L_1), (K_2, L_2)\}$ or $\{(K_1, K_2), (L_1, L_2)\}$.

For the first matching, consider the case if L_1 reports $L \succ_{L_1} \emptyset \succ_{L_1} K$. To ensure strategyproofness, we cannot allow our efficient (IR) matching under this new preference profile to match L_1 to L_2. If we match K_2 to L_2, then the former would have an incentive to report $K \succ_{K_2} \emptyset \succ_{K_2} L$, which by efficiency and IR would result in K_2 being matched to K_1. This would be a profitable deviation and thus not possible by strategyproofness. Hence the only possible match is between K_1 and K_2. But this match is not efficient as L_1 and L_2 can also be matched.

For the second matching, consider the case if L_2 reports $K \succ_{L_2} \emptyset \succ_{L_2} L$. As before, we cannot allow our mechanism to match K_2 and L_2 by strategyproofness. If K_1 and K_2 are matched, then K_1 reporting $L \succ_{K_1} \emptyset \succ_{K_1} K$ would be a profitable deviation as, by efficiency and IR, K_1 and L_1 must be matched. Thus the only possible match would be between K_1 and L_1. However, again this match would not be efficient as K_2 and L_2 can also be matched.

Since neither matchings are possible, then no mechanism that satisfies the stated properties exists. □

Note that the restriction to pairwise exchanges is important. In the given example, the cycle (K_1, L_1, L_2, K_2) would result in all agents getting their preferred choice. We explore in later sections how removing limits on cycle size can give positive results when there are heterogeneous preferences.

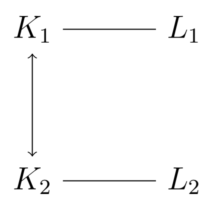
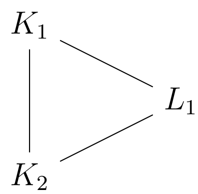


Figure 11: $\mathcal{G}$



(a) $\mathcal{G}$

$\mathcal{I}$	Patient	Donor
K_1	$(B, 2)$	$(A, 3)$
K_2	$(A, 2)$	$(B, 3)$
L_1	$(AB, 2)$	$(O, 1)$

(b) Example observable characteristics

Figure 12: Compatibility graph and example observable characteristics. Note that the patient of all pair is compatible with the donor of other pairs, but not with their own donor.

Transplant Maximization. One of our welfare criteria is transplant maximization. Is it possible to have a mechanism that implements a maximal matching of the compatibility graph, subject to individual rationality?

Proposition 10. *There is no transplant maximal, strategyproof and IR mechanism, nor a transplant maximal, pairwise stable and IR mechanism.*

Proof. Observe that the transplant maximal matching given by $\{(K_1, L_1), (K_2, L_2)\}$ is not strategy proof or pairwise stable. □

Equal Treatment of Exchange Pools. Implicitly, we are favoring kidney pairs over liver pairs as a consequence of the risk ordering. This could be viewed unfavourably, and we may want to consider a mechanism that treats either exchange pool equally. We say a mechanism g satisfies neutrality if swapping organ labels (without changing preferences) does not change the outcome. Note that f is not neutral:

Example 3. Consider the example in Figure 12, where the set of agents is $\mathcal{I} = \{K_1, K_2, L_1\}$ and the compatibility graph in Figure 12a is generated by the observable characteristics in Figure 12b. Assume all donations are individually rational. Then f results in the following matching: K_1 and K_2 matched via kidney donation. However if we swapped the organ labels, then for some $K \in \{K_1, K_2\}$, the following is the matching: K_1 and K are matched. △

We can see that there is no mechanism satisfying our desiderata while also being organ-anonymous:

(a) $\mathcal{G}$ (b) $\mathcal{G}'$

$\mathcal{I}$	Patient	Donor
K_1, L'_1	$(B, 2)$	$(A, 2)$
L_1, K'_1	$(A, 2)$	$(B, 2)$
L_2, K'_2	$(B, 2)$	$(A, 2)$

(c) Example observable characteristics

Figure 13: Compatibility graphs and example observable characteristics. Note that the patient of all pairs is not compatible with their own donor.

Proposition 11. *There is no IC, IR, neutral and PE mechanism, and there is no IR, neutral, PE and pairwise stable mechanism.*

Proof. Consider the environments in Figure 13. An IR, PE and either IC or pairwise-stable mechanism must match (K_1, L_1) in $\mathcal{G}$ and (K'_1, K'_2) in $\mathcal{G}'$. Otherwise, pairs can misreport their preferences to force a better match. Alternatively, there are blocking pairs. Note that swapping labels maintains the same compatibility graph, but the outcomes are different. Hence such a mechanism cannot be neutral. □