

Multi-Organ Exchange: The Whole is Greater than the Sum of its Parts

John P. Dickerson and Tuomas Sandholm

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Kidney exchange, where candidates with organ failure trade incompatible but willing donors, is a life-saving alternative to the deceased donor waitlist, which has inadequate supply to meet demand. While fielded kidney exchanges see huge benefit from *altruistic* kidney donors (who give an organ without a paired needy candidate), a significantly higher medical risk to the donor deters similar altruism with livers. In this paper, we begin by proposing the idea of *liver* exchange, and show on demographically accurate data that vetted kidney exchange algorithms can be adapted to clear such an exchange at the nationwide level. We then explore cross-organ donation where kidneys and livers can be bartered for each other. We show theoretically that this *multi-organ* exchange provides linearly more transplants than running separate kidney and liver exchanges; this linear gain is a product of altruistic kidney donors creating chains that thread through the liver pool. We support this result experimentally on demographically accurate multi-organ exchanges. We conclude with thoughts regarding the fielding of a nationwide liver or joint liver-kidney exchange from a legal and computational point of view.

Introduction

The transplantation of organs from a deceased donor to a needy living candidate first occurred nearly sixty years ago, but only became popular in the 1970s due to the introduction of immunosuppressants that help prevent the rejection of foreign organs in a patient’s body. Since then, the majority of transplantation has occurred through a deceased donor waiting list consisting of needy patients who wait for any willing donors to die, resulting in the harvesting and subsequent transfer of a compatible organ from the donor’s cadaver to the living patient. There is a great supply shortage of cadaveric organs in most societies (including the US), and the imbalance between supply and demand keeps growing. As of December 2013, there were 98,881 patients waiting for a kidney, 15,863 waiting for a liver, and 8,793 for another organ (e.g., pancreas, joint pancreas-kidney, heart, lung, intestine) in the US alone.

In recent years, *live* donation of organs has significantly increased the total number of organ transplants. In live donation, a donor gives one of his two kidneys, one of his

two liver lobes, or a part of an intestine, etc., to the patient so both the donor and patient can live. The effect of live donation has been most prominent in kidney donation, where a recent advance—kidney *exchange* (Roth, Sönmez, and Ünver 2004)—has provided renewed hope to even “hard to match” patients. In kidney exchange, patients bring willing but incompatible donors to a large waiting pool. Patients can then swap incompatible donors with other patients. Matching a candidate to a donor is difficult for a variety of reasons, including blood (ABO) type, tissue (HLA) type, age, and—due to the limitations of current medical knowledge—unknown exogenous factors. Nevertheless, kidney exchanges on the regional and national scale have seen marked success over the last few years.

In this paper, we explore the creation of living donor exchanges involving organs other than kidneys. We first propose *liver* exchange, which is similar to kidney exchange in some ways, but remains unexplored.¹ The major difference between kidney and liver exchange rests in the increased risk to live donors, with very high rates of donor morbidity (24%), “near-miss” events in surgery (1.1%), and mortality (0.2%) compared to live donor kidney transplantation (Cheah et al. 2013). Fielded kidney exchanges derive significant value from *altruistic* donors, who enter the exchange without a paired needy candidate and trigger long “chains” of donations within the pool. With such a high risk of complication from surgery in liver transplantation, we expect significantly fewer (or no, if deemed unethical by the medical community) altruistic donors in liver exchange.

With this in mind, we propose *multi-organ* exchange, where candidates in need of either kidneys or livers can swap donors in the same pool. We show theoretically that this combination provides linearly more transplants than running separate kidney and liver exchanges; this linear gain is a product of altruistic kidney donors creating chains that thread through the liver pool. We support this result experimentally on demographically accurate kidney, liver, and cross-organ exchanges. We conclude with thoughts regarding the fielding of a nationwide liver or joint liver-kidney exchange from a legal and computational point of view.

¹A notable exception is that in Korea, 16 candidates hand-swapped willing donors in a single hospital over the course of six years. All swaps were arranged by hand. This shows the feasibility of the idea at a small scale (Hwang et al. 2010).

This paper provides the first foray into the theory and computational methods necessary to set the groundwork for a fielded nationwide liver or multi-organ exchange. It is clear that such exchanges would be highly beneficial for sustaining life and creating value in society.

Preliminaries

In order to develop a nationwide liver or multi-organ exchange, we must first accurately model the realities of such an exchange and design optimal, scalable clearing algorithms for it. In this section, we describe the creation of a *compatibility graph* representing the space of possible swaps among n candidate-donor pairs, based on traits of the candidates and donors. We then describe the *clearing problem*, a formalization of the process used to determine an optimal set of swaps.

Compatibility Graph

We begin by encoding an n -patient organ exchange as a directed graph. Construct one vertex for each incompatible candidate-donor pair. Add an edge e from one candidate-donor vertex v_i to another v_j , if the candidate at v_j can take a liver lobe or kidney from the donor at v_i . This process creates a compatibility graph for the general concept of barter exchange, where participants can swap items with each other. Within the compatibility graph, a cycle c represents a possible swap, with each vertex in the cycle obtaining the item of the next vertex. A matching is a collection of *disjoint* cycles; no vertex can give out more than one item (e.g., more than one kidney or liver lobe). Cycles ensure that donors give items if and only if their patients receive organs.

Fielded kidney exchanges also gain great utility through the use of *chains* (Rees et al. 2009). An altruistic donor initiates a chain by donating his organ to a patient, whose paired donor donates her organ to another patient, and so on. Due to the significantly increased medical risk to living donors of other organs, we do not expect an influx of (or possibly any) altruistic donors outside of kidney exchanges (Cheah et al. 2013).

The Clearing Problem

The *clearing problem* is that of finding a maximum-cardinality matching consisting of disjoint chains and cycles of length at most some small constant L . The cycle-length constraint is crucial since all operations in a cycle have to be performed simultaneously. Were this not the case, a donor might back out after his incompatible partner has received an organ. This backing out is legal because, in nearly all countries including the US, it is illegal to form a binding contract over the exchange of organs. The availability of operating rooms, doctors, and staff causes long cycles to be unexecutable. As is the practice in the US-wide kidney exchange and most other real kidney exchanges, we let $L = 3$. Chains need not be limited in length (and typically are not in practice); were a donor to renege before giving an organ but after his paired patient had received the organ, then no remaining pair in the pool has lost its “bargaining chip”—although the collapse of the chain is not desired.

Denote the set of all (uncapped length) chains and all cycles of length no greater than L by $C(L)$. Let $|c|$ represent the number of candidate-donor pairs in a cycle or chain c . Then, given binary indicator variables $\forall c \in C(L)$, we must solve the following integer linear program:

$$\max \sum_{c \in C(L)} |c| x_c \quad s.t. \quad \sum_{c: v_i \in c} x_c \leq 1 \quad \forall v_i \in V$$

The clearing problem with any fixed $L > 2$ is NP-complete (Abraham, Blum, and Sandholm 2007). (The cases $L = 2$ with no chains and $L = \infty$ can be solved in polynomial time.) Significantly better (i.e., higher cardinality) results are found with $L = 3$ over $L = 2$, so solving the NP-complete version of the problem is necessary in practice (Roth, Sönmez, and Ünver 2007). The problem, at least with respect to kidneys, can be solved optimally in practice at the steady-state nationwide scale using a specialized tree search algorithm based on the branch-and-price framework for integer programming (Abraham, Blum, and Sandholm 2007). We will later discuss this algorithm in more detail as well as enhancements to it for liver exchange and multi-organ exchange.

Combining Exchanges Results in Linearly More Matches

In this section, we show that combining independent liver and kidney exchanges leads to a linear gain in the aggregate number of matches. We show this in an adapted version of a recent random graph model for kidney exchange due to Ashlagi et al. (2012). They adapt sparse Erdős-Rényi graphs to a model of kidney exchange with two classes of candidate: those with many incoming edges and those with very few incoming edges (intuitively, “easy-to-match” and “hard-to-match” candidates). That model mimics the basic structure of compatibility graphs seen in fielded kidney exchanges.

They build a random directed compatibility graph $D(n, \lambda, t(n), p_L, p_H)$ with n candidate-donor pairs, $t(n)$ altruistic donors, a fraction $\lambda < 1$ of the n candidate-donor pairs—representing *lowly-sensitized*, easy-to-match patients—who have probability p_L of an incoming edge from each vertex in the pool, and a fraction $1 - \lambda > 0$ of the n candidate-donor pairs—representing *highly-sensitized*, hard-to-match patients—who have probability p_H of an incoming edge from each vertex in the pool. We assume $p_L > 0$ is constant, and $p_H = \frac{c}{n}$ for some constant $c > 1$; thus, the graph induced by only those $1 - \lambda$ fraction of (sensitized) vertices with incoming edge probability p_H is sparse.

We assume, for kidney exchange compatibility graphs D_K , $t(n) > 0$; however, for liver exchange graphs D_L , $t(n) = 0$ (i.e., there are no altruistic liver donors). Finally, define the graph join operator $D = \text{join}(D_K, D_L)$ between a kidney exchange graph D_K and liver exchange graph D_L as follows: add directed edges between candidate-donor pairs in both pools in accordance with each pair’s associated probability (p_L or p_H); do not add edges from the $t(n)$ altruistic donors in D_K to vertices in D_L (since altruistic kidney donors are unwilling to donate a liver).²

²For the sake of clarity, we assume that the p_L (resp. p_H) for

In the following theoretical results, we consider cycles of length at most some constant but chains of any length; this mimics current practice in kidney exchange, and would likely mimic that of fielded liver exchange. Thus, an *efficient matching* allocates the maximum number of transplants in cycles of size no more than some constant and chains of any length. Both results build on the work of Ashlagi et al. (2012), which considers only a single kidney exchange.

Proposition 1 assumes a linear (in the number of candidate-donor pairs) number of altruistic donors, while Proposition 2 works with just a constant number of altruistic donors. We omit the proof of Proposition 1 due to space, and contrast both theoretical results at the end of this section.

Proposition 1. *Consider $\beta > 0$ and $\gamma > 0$, kidney compatibility graph D_K with n_K pairs and $t(n_K) = \beta n_K$ altruistic donors, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. Then any efficient matching on $D = \text{join}(D_K, D_L)$ matches $\Omega(n_K)$ more pairs than the aggregate of any such efficient matchings on D_K and D_L (with probability approaching 1 as n_K approaches ∞).*

Proposition 2. *Consider $\gamma > 0$, kidney compatibility graph D_K with n_K pairs and constant $t > 0$ altruistic donors, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. Then there exists $\lambda' > 0$ such that for all $\lambda < \lambda'$, any efficient matching on $D = \text{join}(D_K, D_L)$ matches $\Omega(n_K)$ more pairs than the aggregate of any such efficient matchings on D_K and D_L (with constant positive probability).*

Proof sketch. For small enough λ and large enough c , with high probability there exists a set S_K (of size at least $n_K/2$) of highly-sensitized pairs in D_K that are “too far” away from lowly-sensitized pairs in D_K to be matched in a cycle of capped length and must be matched in a chain triggered by an altruist a or not matched at all (Ashlagi et al. 2012). By similar reasoning, there exists a larger set $S_{K\&L}$ of highly-sensitized pairs in the combined kidney and liver graph (of size at least $(n_K + n_L)/2 = (1 + \gamma)n_K/2$) that must be matched by an a -initiated chain or not at all.

We apply a general result on sparse random directed graphs from Krivelevich, Lubetzky, and Sudakov (2013): a directed path of length approaching $|S_K|$ in S_K and $|S_{K\&L}|$ in $S_{K\&L}$ exists (as $c \rightarrow \infty$). With constant probability there exists an edge from a to one of the vertices in the first half of the directed path in S_K (Ashlagi et al. 2012); thus, the size of this a -initiated chain approaches at least $n_K/4$ as $c \rightarrow \infty$. Similarly, with the same constant probability there exists an edge from a to one of the vertices in the first half of the directed path in $S_{K\&L}$ (since in expectation $1/(1 + \gamma)$ fraction of the first half of this path are in the original kidney graph D_K , and $\gamma/(1 + \gamma)$ in expectation are in D_L and thus have probability 0 of an incoming edge from a), resulting in a chain of length approaching at least $(1 + \gamma)n_K/4 > n_K/4$ in expectation (as $c \rightarrow \infty$).

Thus, by combining pools, we see an increase of at least $(1 + \gamma)n_K/4 - n_K/4 = \gamma n_K/4$, which is $\Omega(n_K)$. This is a linear increase in overall efficiency since $n_L = \gamma n_K$. \square

D_K equals the p_L (resp. p_H) for D_L . This is without loss of generality; all that matters is that p_L be constant and $p_H = \frac{c}{n}$ for $c > 1$.

Intuitively, Propositions 1 and 2 shows the theoretical efficacy of combining kidney exchange with alternate organ exchanges (where altruistic donation is less likely to be popular or deemed ethically acceptable).³ We will support Proposition 1 empirically in the coming sections.

On the dense model for organ exchange. Initial research on random graph models for organ exchange adapted *dense* (constant probability of an edge existing) Erdős-Rényi graphs to kidney exchange (Ashlagi and Roth 2011; Dickerson, Procaccia, and Sandholm 2012b). Fielded exchanges have proven to be sparse in practice—as in the theory above—and thus actual pools and their optimal matchings do not align with these dense models (Ashlagi et al. 2012; Ashlagi, Jaillet, and Manshadi 2013; Dickerson, Procaccia, and Sandholm 2013; 2014). Still, we note that the efficiency results in the dense model without chains and with chains (Proposition 5.1 of Ashlagi and Roth (2011) and Theorem 1 of Dickerson, Procaccia, and Sandholm (2012b), respectively) can be applied directly to independent liver exchange and multi-organ exchange to yield efficient matchings with linear expected overall gain from combining pools (given a linear number of altruists) for large enough compatibility graphs.⁴

Generating and Clearing Demographically Accurate Pools

In this section, we describe our method for generating organ exchange graphs. We then describe the standard kidney exchange clearing algorithm and, motivated by generated realistic liver and kidney exchange graphs, present a tweak to this algorithm to decrease liver exchange solution time.

Data Generation

In order to create an at-scale nationwide liver or multi-organ exchange, we first have to develop a compatibility graph generator with which we can run simulations. First, we draw data from reliable sources (here, specific to the US). Second, this data is fed into a graph creation algorithm that probabilistically determines the existence of compatible and incompatible candidate-donor pairs, as well as compatibility constraints between different candidate-donor pairs. In the large, with high probability, graphs generated by this algorithm will mimic the demographics that would prevail in a large-scale fielded exchange in the US. (Plugging different raw data (e.g., age, weight, blood type distributions) into the generator algorithm would provide realistic generation of non-US compatibility graphs.)

³While Proposition 2 may seem like a stronger result due to its relaxed reliance on a constant number of altruistic kidney donors (instead of the linear number in Proposition 1), the numerator c in $p_H = c/n$ may be required to be quite large (although still constant), the λ sensitivity constant quite small, and the result also holds with constant positive probability instead of holding with probability approaching one. We feel this makes Proposition 1 a more relevant result overall than Proposition 2 for the composition (in terms of pool sensitization and number of altruistic donors) of currently fielded kidney exchanges.

⁴The proofs of Proposition 1 and the statement on dense liver and multi-organ exchange are given in the attached appendix.

We generate kidney exchange compatibility graphs in accordance with Saidman et al. (Saidman et al. 2006); however, the compatibility of a potential liver donor with a candidate differs from that of a potential kidney donor in three critical ways. While a donor and candidate must be blood type (ABO) compatible, (a) they need not be HLA-compatible,⁵ (b) the age of the donor and candidate makes a significant difference in transplant success (Egawa et al. 2004), and (c) the donor must be heavier than the candidate (or else the donor’s liver, which must be cut in two before transplantation, will not be large enough to support the donor and candidate).

Graph generation is performed as follows. For each candidate and donor, we draw a gender (from the 2010 US Census Report⁶); conditioned on gender, we then draw candidate blood types from the OPTN (Organ Procurement and Transplantation Network⁷) distribution and donor blood types from the overall US population.⁸ We sample ages (dependent on gender) for candidates from the OPTN pool and for the donors from the 2010 US Census at a granularity level of one year. Then, given the age and gender (generated separately from OPTN data for candidate and US Census data for donors, as described earlier), we sample from a fine-grained table of weights recently released by the Center for Disease Control (McDowell et al. 2008). For candidates requiring a kidney, HLA is sampled from the OPTN databases. During edge generation, we include an exogenous “failure factor” $f \in [0, 1]$ that randomly determines an edge failure even in the case of a compatibility success. This factor is common in the kidney literature (Ashlagi et al. 2011), and is used to account for incompleteness of medical knowledge and, during simulation, temporal fluctuations in candidate-donor compatibility.

The Clearing Algorithm

We now briefly discuss the most scalable optimal kidney exchange clearing algorithm (Abraham, Blum, and Sandholm 2007), which is also used in the US-wide kidney exchange; we adapt that algorithm for our liver and multi-organ exchange experiments based on characteristics of the graphs generated using the algorithm described above. At a high level, given a compatibility graph $G = (V, E)$, the algorithm enumerates all chains and cycles of length at most L and chooses the optimal disjoint set of these cycles and chains according to the objective function of maximizing cardinality of the matching.

In reality the number of cycles is prohibitively large (cubic in $|E|$ for $L = 3$, and exponential in $|E|$ for unbounded chains) to write down in memory. Therefore, solving this problem hinges on a technique called *branch-and-*

price (Barnhart et al. 1998), a method for incrementally generating only a small part of the model during tree search, yet guaranteeing optimality by proving that all the promising variables have been incorporated into the model. The actual solver uses several additional techniques to make kidney exchange clearing scalable for memory and time (Abraham, Blum, and Sandholm 2007). It uses empirically and theoretically motivated heuristics to seed the initial cycle (i.e., variable) set used on the model, and then incrementally brings cycles into the model depending on their shadow price, a quantitative estimate of a cycle’s utility given the current model. Optimality is proven when no cycles can possibly increase the objective. The algorithm also uses specific branching heuristics and primal heuristics to construct feasible initial integral solutions at each branch. If these integral solutions match the (restricted, possibly fractional) LP solution, then the subtree can be pruned and optimality potentially proven.

A Liver-Specific Cycle Seeding Heuristic. The selection of the initial seed columns—representing individual cycles—is a heuristic process. The prior algorithm uses the cycles from two heuristically-generated feasible solutions (very few such cycles) and hundreds of thousands of randomly selected cycles from $C(L)$. Since enumerating $C(L)$ in its entirety is a costly ordeal, their sampling relies on a series of random walks. Starting at a randomly chosen vertex, a random walk takes steps to new vertices. At each step, if an edge exists leading back to the initial vertex, the corresponding cycle is added to the set of seed cycles and a new start vertex is chosen. This results in a randomized, but not uniformly random, sampling of all cycles.

We define a different sampling method for the cycle seeding problem. Our generated liver compatibility graphs tended to have many more vertices with low out-degree than the corresponding kidney exchange graphs. These candidates are difficult to match. With this in mind, we conduct a biased random walk sampling in the same spirit as the prior algorithm, except weighting the selection of the randomized start vertex inversely proportional to its out-degree. This biased sampling of the set of all cycles motivates the solver to branch on hard-to-match candidate-donor pairs. This can be done efficiently through an initial sorting of the vertices by out-degree, a process whose one-time $O(|V| \log |V|)$ run-time is overshadowed by the NP-hard clearing problem.

Experimental Results

We now provide computational results for a hypothetical nationwide liver or multi-organ exchange, using the realistic data generated above. First, we describe timing and matching results in the *static* case, where the algorithm sees the problem in its entirety up front. Second, we describe results for the *dynamic* case, where candidate-donor pairs arrive in the pool over time and are either matched or die waiting. We show results at sizes mirroring an estimated steady-state size of a US-wide liver exchange. Finally, we explore the possibility of a multi-organ exchange, where both liver- and kidney-needing candidates can swap donors in the same pool. This results in more lives being saved than were the nation to run separate liver and kidney exchanges.

⁵In kidney exchange, tissue type (HLA antibodies and antigens) are an important determinant of compatibility. A candidate and donor sharing antigen encodings on the same locus are more likely to result in a rejected kidney. Due to the use of suppressant drugs, this is a negligible factor in liver transplantation.

⁶<http://www.census.gov/compendia/statab/cats/population.html>

⁷<http://optn.transplant.hrsa.gov/data/>

⁸http://bloodcenter.stanford.edu/about_blood/blood_types.html

Static Liver Exchange Experiments

In the static case, the generator outputs a single graph and the optimization engine solves the clearing problem on this graph exactly once. Figure 1 shows timing results on liver exchange graphs of various sizes $|V|$ and exogenous failure rates f . Intuitively, when f is low (or zero), the optimizer must consider many more edges than when f is high, resulting in longer runtimes for denser graphs. As expected, the computation time increases drastically with graph size.

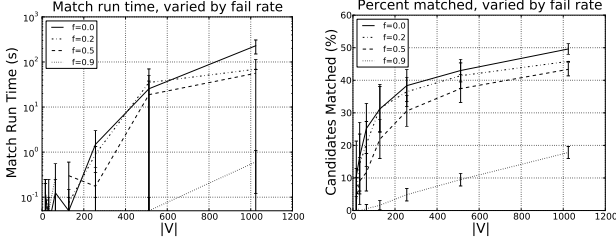


Figure 1: Match runtime (left) and percentage of candidates matched (right), varying failure rate f and graph size $|V|$.

Figure 1 also shows the percentage of candidates matched (the number of candidates matched by the algorithm divided by the total number of candidates in the pool) as a function of compatibility graph size $|V|$ and exogenous failure rate f . Intuitively, when f is held low, the percentage of candidates matched is higher than when the failure rate is high. Of interest is the match behavior as $|V|$ increases. Regardless of f , the percentage of candidates matched increases with the size of the underlying compatibility graph. This behavior is similar to that seen in kidney exchange and motivates the need for a large (i.e., nationwide) liver exchange.

Addressing the needs of society. The estimated steady-state monthly size of the nationwide kidney exchange is 10,000 candidate-donor pairs (Abraham, Blum, and Sandholm 2007). The rate of live liver donation is 1/8th of the rate of live kidney donation (5% of all liver transplants involve live donors, compared to 40% for kidneys (Brown 2008)), although this number would hopefully increase due to the publicity of a successful exchange—we will conservatively estimate a factor of 1/2 as many live liver donors than kidney donors in steady-state. With 98,881 candidates currently waiting for a kidney and 15,863 candidates waiting for a liver in the US—and half as many live donors available—the steady-state for a US-wide liver exchange can be estimated at approximately half of $15,863 / 98,881 \approx 8\%$ of 10,000, or roughly 800 candidates. So, our clearing algorithm should be able to handle batch runs of a nationwide liver exchange.

Dynamic Liver Exchange Experiments

In the dynamic case, a variable number of candidates enter and leave the pool over a period of multiple time units. While the fielded nationwide kidney exchange currently operates under the static paradigm described earlier, recent work in the kidney exchange community has shown that optimizing in the dynamic setting leads to both more realistic and higher cardinality matchings over time (Awasthi and Sand-

holm 2009; Ünver 2010; Dickerson, Procaccia, and Sandholm 2012a).

We start with a pool of $|V| = 800$ candidates assumed to be highly sensitized patients who built up in the system over time. These are matched myopically. Given a matched cycle by the algorithm, we then simulate that transplant actually succeeding in real life via an exogenous parameter set to $f = 0.7$. If any edge in a cycle fails, that entire cycle fails, and all candidates are returned to the pool (with the failed edge removed). We simulate candidates leaving the pool (either through finding a transplant or dying). On expectation $|V_{new}| = 226$ new candidates arrive in the pool per month, and the algorithm continues. We test over 24 months.

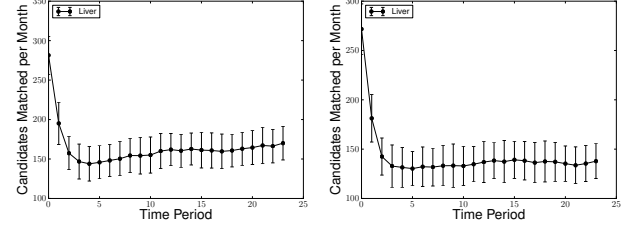


Figure 2: Number of candidates matched per time period in a dynamic setting over $T = 24$ months, with an expected lifetime per candidate of 10 years (left) or 1–2 years (right).

Figure 2 shows the number of candidates matched at each time period. This is the number of candidates matched by the algorithm, but before the virtual failures are taken into account. On the left, 12% of candidates will be alive after 10 years, corresponding to the expected lifetime of a kidney patient on dialysis waiting for a kidney (USRDS 2007). On the right, the probability of a candidate dying is set to an expected life of 1–2 years. This mimics the urgency of needing a liver transplant. While dialysis can be used to keep a patient with failed kidneys alive, no such treatment exists for livers. This corresponds to a drop in the number of candidates matched, due to the decreased number of candidates in the pool at each time period. (Note that a large number of candidates are matched per month in the beginning when the exchange goes live because there is a large pool that has accumulated. Soon thereafter a steady state is reached.)

Dynamic Bi-Organ Exchange Experiments

In this section, we expand beyond simulating a dynamic liver exchange to the novel concept of multi-organ exchange. In the long run, one could imagine exchanges of multiple different kinds of organs. However, to our knowledge, only kidneys and livers have ever been swapped (and only separately). Therefore, in this section we will focus on kidneys and livers. We show that combining an independent nationwide liver exchange with a nationwide kidney exchange into a joint kidney-liver exchange results in a statistically significant increase in the number of organ transplants, which aligns with Proposition 1.

We simulate a demographically accurate bi-organ exchange featuring candidates in need of either a kidney or a liver who can swap donors in a combined candidate-donor pool. Approximately 85% of the candidates in the simulated

pool need kidneys, while the other 15% need livers, as determined by the most recent OPTN waitlist data. We mimic the experiments in the previous section, with a starting pool size of $|V| = 800$ candidates who are highly sensitized and are assumed to have built up in the pool over time; we also include 100 altruistic kidney donors who enter the combined pool at an expected constant rate. We use the same exogenous transplant failure parameter ($f = 0.7$) as in the previous section, and simulate candidate-donor pairs entering and exiting the pool in a similar fashion. To generate the candidates, we draw from the two different US distributions based on whether the candidate needs a kidney or a liver. Naturally, donors are drawn from the same US distribution in the two cases. We test over 24 months.

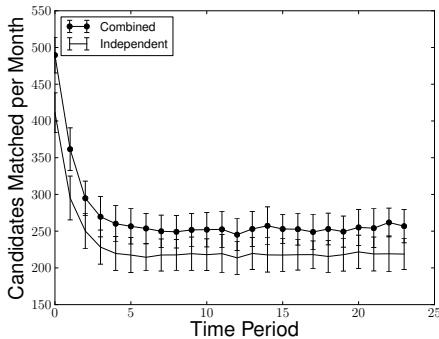


Figure 3: Number of matches in independent liver and kidney exchanges and a combined multi-organ exchange, per time period, in a dynamic setting over $T = 24$ months.

Figure 3 shows the number of candidates matched each month in the combined bi-organ exchange, as well as the aggregate number of candidates matched while keeping both liver- and kidney-needing candidates in separate pools. Clearly evident is the loss of life resulting from keeping both the liver and kidney pools independent, with the bi-organ exchange matching nearly 40 more candidates per month when compared to the two independent exchanges.

When we compare the *total* number of matches made over the entire period simulated above, the difference in lives saved between two independent pools and the combined bi-organ pool is more stark. In our experiments, the combined bi-organ pool produced 16.8% more matches than the sum of the two independent organ pools. An independent samples *t*-test revealed that the difference between the aggregate number of lives saved using independent, simultaneous liver and kidney exchanges and using a combined multi-organ exchange was significant, $t(46) = 31.37, p \ll 0.0001$.

Conclusions and Future Work

We explored the possibility of extending large-scale organ exchange to include liver lobes, either in conjunction with or independently of presented fielded kidney exchange. On demographically accurate data, vetted kidney exchange clearing algorithms (with a small tweak) can also clear liver exchanges at a projected US nationwide size. We explored the prospect of multi-organ exchange, where candidates needing either a liver or kidney can swap willing donors in the same

pool. We showed that such a combination matches linearly more candidates than maintaining two separate exchanges; this linear gain is a product of altruistic kidney donors creating chains that thread through the liver pool. This result is supported experimentally on demographically accurate multi-organ exchanges with high statistical significance.

This paper is intended as a first foray into automated liver and multi-organ exchange. As such, there is much room for future research (much of which is applicable to other organ exchange and even to barter exchanges beyond organs), and is motivated by experiences fielding the nationwide kidney exchange. One direction of future work is to take on the slow and politics-laden task of founding a liver exchange, or including livers in currently fielded kidney exchanges. Another is to develop scalable computational methods for the dynamic problem. Even for kidneys, the best current techniques are for simplified models (Ünver 2010; Ashlagi, Jaillet, and Manshadi 2013; Anshelevich et al. 2013) or face computational challenges (Awasthi and Sandholm 2009; Dickerson, Procaccia, and Sandholm 2012a).

Even for the static problem, scalability problems tend to get worse with the inclusion of a recent innovation in kidney exchange—donation chains started by altruistic donors. The cycle length cap L no longer applies to chains since they do not require simultaneous execution. Recent work explores this innovation, and hits computational limits experimentally with long chains (Ashlagi et al. 2012; 2011; Dickerson, Procaccia, and Sandholm 2012a; 2012b; Gentry and Segev 2011; Gentry et al. 2009). We do not expect altruistic donors in liver exchange due to increased risk for the donor compared to kidney donation, complicating the ethical considerations of even allowing altruistic donors in the pool (Woodle et al. 2010). However, that remains to be seen. In any case, one could include chains started by kidney-donating altruists into a bi-organ exchange—if the scalability challenges of chains can be adequately addressed.

Finally, this paper (and most papers on kidney exchange) deals with optimizing algorithmic organ matches; in reality, most algorithmic matches in fielded kidney exchanges do not result in an actual transplant. We expect this would be the case in liver and multi-organ exchange as well, although the exact failure rates for liver and multi-organ exchanges would be different than the observed failure rates in currently fielded kidney exchanges due to the medical and logistical differences in the organs and the transplant processes. Making organ exchange failure-aware is a critical step toward improving yield; recent work explores this notion (Blum et al. 2013; Dickerson, Procaccia, and Sandholm 2013) to both theoretically and empirically maximize the expected number of actual transplants (possibly with respect to some fairness constraints (Dickerson, Procaccia, and Sandholm 2014)) stemming from an algorithmic match.

Regardless, the urgent societal need for liver exchange is there today, and we hope to be able to address it through a dedicated or combined liver- or multi-organ exchange.

References

Abraham, D.; Blum, A.; and Sandholm, T. 2007. Clearing algorithms for barter exchange markets: Enabling nation-

- wide kidney exchanges. In *ACM Conference on Electronic Commerce (EC)*, 295–304.
- Anshelevich, E.; Chhabra, M.; Das, S.; and Gerrior, M. 2013. On the social welfare of mechanisms for repeated batch matching. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 60–66.
- Ashlagi, I., and Roth, A. E. 2011. Individual rationality and participation in large scale, multi-hospital kidney exchange. In *ACM Conference on Electronic Commerce (EC)*, 321–322.
- Ashlagi, I.; Gilchrist, D. S.; Roth, A. E.; and Rees, M. 2011. Nonsimultaneous chains and dominos in kidney-paired donation—revisited. *American Journal of Transplantation* 11(5):984–994.
- Ashlagi, I.; Gamarnik, D.; Rees, M.; and Roth, A. E. 2012. The need for (long) chains in kidney exchange. NBER Working Paper No. 18202.
- Ashlagi, I.; Jaillet, P.; and Manshadi, V. H. 2013. Kidney exchange in dynamic sparse heterogeneous pools. In *ACM Conference on Electronic Commerce (EC)*, 25–26.
- Awasthi, P., and Sandholm, T. 2009. Online stochastic optimization in the large: Application to kidney exchange. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, 405–411.
- Barnhart, C.; Johnson, E.; Nemhauser, G.; Savelsbergh, M.; and Vance, P. 1998. Branch-and-price: column generation for solving huge integer programs. *Operations Research* 46:316–329.
- Blum, A.; Gupta, A.; Procaccia, A. D.; and Sharma, A. 2013. Harnessing the power of two crossmatches. In *ACM Conference on Electronic Commerce (EC)*, 123–140.
- Brown, R. S. 2008. Live donors in liver transplantation. *Gastroenterology* 134(6):1802–1813.
- Cheah, Y. L.; Simpson, M. A.; Pomposelli, J. J.; and Pomfret, E. A. 2013. Incidence of death and potentially life-threatening near-miss events in living donor hepatic lobectomy: A world-wide survey. *Liver Transplantation* 19(5):499–506.
- Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2012a. Dynamic matching via weighted myopia with application to kidney exchange. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 1340–1346.
- Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2012b. Optimizing kidney exchange with transplant chains: Theory and reality. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 711–718.
- Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2013. Failure-aware kidney exchange. In *ACM Conference on Electronic Commerce (EC)*, 323–340.
- Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2014. Price of fairness in kidney exchange. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*.
- Egawa, H.; Oike, F.; Buhler, L.; Shapiro, A.; Minamiguchi, S.; Haga, H.; Uryuhara, K.; Kiuchi, T.; Kaihara, S.; and Tanaka, K. 2004. Impact of recipient age on outcome of ABO-incompatible living-donor liver transplantation. *Transplantation* 77(3):403.
- Gentry, S. E., and Segev, D. L. 2011. The honeymoon phase and studies of nonsimultaneous chains in kidney-paired donation. *American Journal of Transplantation* 11(12):2778–2779.
- Gentry, S. E.; Montgomery, R. A.; Swihart, B. J.; and Segev, D. L. 2009. The roles of dominos and nonsimultaneous chains in kidney paired donation. *American Journal of Transplantation* 9(6):1330–1336.
- Hwang, S.; Lee, S.-G.; Moon, D.-B.; Song, G.-W.; Ahn, C.-S.; Kim, K.-H.; Ha, T.-Y.; Jung, D.-H.; Kim, K.-W.; Choi, N.-K.; Park, G.-C.; Yu, Y.-D.; Choi, Y.-I.; Park, P.-J.; and Ha, H.-S. 2010. Exchange living donor liver transplantation to overcome ABO incompatibility in adult patients. *Liver Transplantation* 16(4):482–490.
- Janson, S.; Luczak, T.; and Rucinski, A. 2011. *Random Graphs*. Wiley Series in Discrete Mathematics and Optimization. Wiley.
- Krivelevich, M.; Lubetzky, E.; and Sudakov, B. 2013. Longest cycles in sparse random digraphs. *Random Struct. Algorithms* 43(1):1–15.
- McDowell, M.; Fryar, C.; Ogden, C.; and Flegal, K. 2008. Anthropometric reference data for children and adults: United States, 2003–2006. *Nutrition* 10(10):1–45.
- Rees, M.; Kopke, J.; Pelletier, R.; Segev, D.; Rutter, M.; Fabrega, A.; Rogers, J.; Pankewycz, O.; Hiller, J.; Roth, A.; Sandholm, T.; Ünver, U.; and Montgomery, R. 2009. A non-simultaneous, extended, altruistic-donor chain. *New England Journal of Medicine* 360(11):1096–1101.
- Roth, A.; Sönmez, T.; and Ünver, U. 2004. Kidney exchange. *Quarterly Journal of Economics* 119(2):457–488.
- Roth, A.; Sönmez, T.; and Ünver, U. 2007. Efficient kidney exchange: Coincidence of wants in a market with compatibility-based preferences. *American Economic Review* 97:828–851.
- Saidman, S. L.; Roth, A.; Sönmez, T.; Ünver, U.; and Delmonico, F. 2006. Increasing the opportunity of live kidney donation by matching for two and three way exchanges. *Transplantation* 81(5):773–782.
- Takahashi, K. 2007. Recent findings in ABO-incompatible kidney transplantation: classification and therapeutic strategy for acute antibody-mediated rejection due to ABO-blood-group-related antigens during the critical period preceding the establishment of accommodation. *Clinical and Experimental Nephrology* 11(2):128–141.
- Ünver, U. 2010. Dynamic kidney exchange. *Review of Economic Studies* 77(1):372–414.
2007. United States Renal Data System (USRDS). <http://www.usrds.org/>.
- Woodle, S.; Daller, J.; Aeder, M.; Shapiro, R.; Sandholm, T.; Casingal, V.; Goldfarb, D.; Lewis, R.; Goebel, J.; and Siegler, M. 2010. Ethical considerations for participation of nondirected living donors in kidney exchange programs. *American Journal of Transplantation* 10:1460–1467.

Appendix A: A Parameterized, Realistic Compatibility Graph Generator

In order to create an at-scale nationwide liver or multi-organ exchange, we first have to develop a compatibility graph generator with which we can run simulations. First, we draw data from reliable sources (here, specific to the US). Second, this data is fed into a graph creation algorithm that probabilistically determines the existence of compatible and incompatible candidate-donor pairs, as well as compatibility constraints between different candidate-donor pairs. In the large, with high probability, graphs generated by this algorithm will mimic the demographics that would prevail in a large-scale fielded exchange in the US. (Plugging different raw data (e.g., age, weight, blood type distributions) into the generator algorithm would provide realistic generation of non-US compatibility graphs.) We then conclude the section with a comparison of liver exchange graphs generated by our algorithm to kidney exchange graphs generated by the standard generator of Saidman et al. (2006). Our generator is a generalization of (i.e., more powerful than) that current standard.

Sampling from Real-World Data

Current medical knowledge is incapable of exactly predicting the compatibility of a particular donor and candidate. However, many attributes are known that can guide doctors—and algorithms—toward a realistic quantification of the chance of organ rejection. In this section, we describe these factors and the open source data sets that our algorithm uses to realistically sample the US population. In the discussions ahead, we use “OPTN” to refer to the data available from the Organ Procurement and Transplantation Network.⁹ All OPTN data is current as of November 11, 2011.

Gender While a donor of one gender can donate an organ to a candidate of another gender, we must take gender into account during graph generation. This is because other traits that affect the probability of a transplant’s success (e.g., weight or age) depend on a person’s gender. We draw candidate genders from the OPTN data set, and donor genders from the greater US population through the 2010 US Census report.¹⁰ Figure 4 shows the distributions of liver-needing candidates and the natural US population as donors. Men are very over-represented in the candidate pool. (Note that similar distributions can be obtained for kidney-needing candidates, and used in a multi-organ generator.)

	Male	Female
Candidate	61.71	38.29
Donor	48.53	51.47

Figure 4: Distribution of (liver) candidate and donor genders, drawn from OPTN and 2010 US Census data, respectively.

Blood Type A candidate and donor must be ABO blood type compatible (e.g., an A-type donor is compatible with A- and AB-type candidates), although blood type suppression through drugs is a recent advance that has the potential to remove this constraint (Takahashi 2007). We draw candidate blood types from the OPTN distribution (dependent on gender), and donor blood types from the overall US.¹¹ The OPTN distribution is roughly equal across genders, and both distributions are roughly equal to each other. Nevertheless, it is important to have this parameterized capability in the generator in the event that, for instance, some “harder” blood type (e.g., AB) gets over-represented in the candidate pool. Figure 5 shows the exact distribution and the ABO-compatibility matrix, with percentages shown for liver-needing candidates.

Donor	Candidate				
	ABO	O	A	B	AB
O		x	x	x	x
A			x		x
B				x	x
AB					x

	Male		Female	
ABO	Cand.	Donor	Cand.	Donor
O	47.83	44	48.91	44
A	38.39	42	37.08	42
B	11.37	10	11.41	10
AB	2.40	4	2.58	4

Figure 5: Top: ABO blood type compatibility matrix. Marks indicate a donor (row) as ABO-compatible with a candidate (column). Bottom: ABO percentages for candidates and donors.

Age Age plays a role in transplantation, but we were unable to find any specific quantification of the amount by which increased donor or candidate age (or, in the case of children, decreased candidate age) affects this success rate. Even without this information, age is important to model because it will allow us to generate a realistic distribution of candidate and donor *weights*, a trait whose effect is easily quantified. We sample ages (dependent on gender) for candidates from the OPTN pool and for the donors from the 2010 US Census at a granularity level of one year. To save space, Figure 6 does not separate the population into one-year segments as rows, while our generator does. In our generator we also take into account the constraint that organ donors must be 18 years old, and we normalize the distributions accordingly.

Weight Unlike in kidney exchange, the physical weight of both the candidate and donor play an enormous role in the feasibility of liver transplantation.¹² Intuitively, the size of a liver is generally proportional to the size of the person who grew it. In live liver donation, the donor’s liver is cut in two

⁹ <http://optn.transplant.hrsa.gov/data/>

¹⁰ <http://www.census.gov/compendia/statab/cats/population.html>

¹¹ http://bloodcenter.stanford.edu/about_blood/blood_types.html

¹² Large weight differences between donor and candidate can factor into kidney exchange as well, but this has not been taken into account in either the current state of the art generator or the weighting algorithms used in the fielded US-wide kidney exchange.

Age	Male		Female	
	Candidate	Donor	Candidate	Donor
<1	0.259	—	0.465	—
1–5	0.837	—	1.220	—
5–10	0.568	—	1.075	—
11–17	0.717	—	1.444	—
18–34	4.193	31.883	5.554	29.357
35–49	14.851	27.798	14.976	26.617
50–64	64.851	25.066	57.079	25.053
≥65	13.725	15.252	18.186	18.972

Figure 6: Probability distribution of ages, respective of candidate and donor gender.

(one lobe is removed). For both donor and candidate to remain healthy, the slice of liver left in the donor must be large enough to maintain her life, and the slice of liver given to the candidate must be large enough to maintain his. Thus, a general rule of thumb that the donor must weigh as much as (or more than) the candidate is in place in live liver donation. We adopt that convention for liver exchange.

Given the age and gender (generated separately from OPTN data for candidate and US Census data for donors, as described earlier), we sample from a fine-grained table of weights recently released by the Center for Disease Control (McDowell et al. 2008). This data, given on a by-year basis until age 20 and in increments of 5 years thereafter, includes mean weights, sample errors, and sample sizes. From this, we calculate a standard deviation and sample from a normal distribution with this mean and standard deviation. While there are issues with this method—most notably that the candidate weights may be drawn from a different distribution than the general US public, and that human weights are not distributed normally but are skewed toward weighing more—we feel that this sampling approach provides a reasonable starting point for future generation techniques. The full table of weights is omitted due to space.

HLA Antibodies and Antigens In kidney exchange, tissue type (HLA antibodies and antigens) are another very important determinant of compatibility. A candidate and donor sharing antigen encoding on the same locus possibly results in a positive *virtual crossmatch* across antigens. A positive virtual crossmatch means that the system can detect incompatibility. In kidney exchange graph generation, this is quantified by the probability that the candidate is not tissue-type compatible with a randomly drawn donor. This probability is called %PRA for panel reactivity antibody (Saidman et al. 2006). Furthermore, tissue type can change over time, resulting in the need for contingency plans after the time of algorithmic matching but before the surgery. For example, if the candidate comes down with a cold or flu days before surgery, the surgery may need to be rescheduled or permanently canceled.

In liver exchange, %PRA plays less of a role due to the use of suppressant drugs. As such, while the generator supports %PRA (and can use sampled data from the OPTN databases¹³), we exclude %PRA in our liver experiments.

¹³The relationship (e.g., spousal, parent-child) between candi-

However, %PRA is included in our multi-organ experiments for kidney candidates.

Generator Algorithm

We now give the method for generating the compatibility graph from data sampled from the sources given in the previous section. Note that the probability distributions from the previous section (and the organs to which they pertain) can be swapped without affecting the correctness of the algorithm beyond the “is compatible” checks described below.

Algorithm 1: Compatibility graph generator

Input: Integer n , real number f , real-world data
Output: Compatibility graph $G = (V, E)$ s.t. $|V| = n$
begin
 $G := (V = \emptyset, E = \emptyset)$
 while $|V| < n$ **do**
 $c = \text{candidate}, d = \text{donor}$
 $c.\text{drawOrganType}()$
 $\{c, d\}.\text{drawGender}()$
 $\{c, d\}.\text{drawBlood}(\text{gender})$
 $\{c, d\}.\text{drawAge}(\text{gender})$
 $\{c, d\}.\text{drawTissueType}(\text{gender})$
 $\{c, d\}.\text{drawWeight}(\text{gender}, \text{age})$
 if $\neg \text{isCompatible}(c, d)$ **then**
 $V = V \cup \{v^{c,d}\}$
 for $v_i, v_j \in V$ s.t. $V_i \neq V_j$ **do**
 if $\text{isCompatible}(v_j^c, v_i^d)$ and $x \in U[0, 1] > f$ **then**
 $E = E \cup \{(v_i, v_j)\}$
 return directed compatibility graph G

Algorithm 1 gives a two-step process for generating a compatibility graph $G = (V, E)$, given a number n , such that $|V| = n$. First, sample from real-world data until n incompatible candidate-donor pairs are generated. When generating a liver exchange, one would set the algorithm to sample from the liver data given above; however, when generating a multi-organ exchange consisting of livers and kidneys, one would include the proper proportions of kidney and liver candidates and sample from the appropriate real-world data per organ. As of the writing of this paper, the kidney waitlist is 5.84 times longer than the liver waitlist, which would be reflected in this algorithm.

If needed, the algorithm can easily be augmented to keep track of any compatible candidate-donor pairs generated. As is common practice in kidney exchange, these pairs are assumed to match on their own, and do not enter the pool.¹⁴ Other additions could be made to the algorithm as data becomes available (e.g., correlating donor and candidate characteristics under the assumption that a donor may likely come from the candidate’s family).

date and donor can yield information on HLA compatibility, and is supported by the generator of Saidman et al. and our generator.

¹⁴Recent kidney exchange research suggests that incentivizing even compatible pairs to join a nationwide exchange could result in better matchings (Rees et al. 2009; Ashlagi and Roth 2011).

After n incompatible candidate-donor pairs are generated, the algorithm steps through each pair v_i, v_j of candidate-donor pairs and, if the latter's candidate v_j^c is compatible with the former's donor v_i^d , then a directed edge is added from v_i to v_j . Note the inclusion of an exogenous "failure factor" $f \in [0, 1]$ that, if prescribed, randomly determines an edge failure even in the case of a compatibility success. This factor is common in the kidney literature (Ashlagi et al. 2011), and is used to account for incompleteness of medical knowledge and, during simulation, temporal fluctuations in candidate-donor compatibility.

Algorithm 1 calls a function *isCompatible(c, d)*. In the liver case, this checks whether two patients are ABO-compatible and whether the donor's weight is greater than or equal to the candidate's weight. In the kidney case, this checks whether two patients are ABO-compatible and whether a virtual crossmatch based on tissue type returns negative. As better medical knowledge and data become available, this function can be generalized to take new compatibility aspects into account.

Comparison to Kidney Exchange

We now compare our generator to the current state of the art (kidney) exchange generator (Saidman et al. 2006). While the generators and data are similar in spirit, the medical differences between kidney and liver compatibility create distinctly different compatibility graphs both at the small and large scale. We will discuss those differences below.

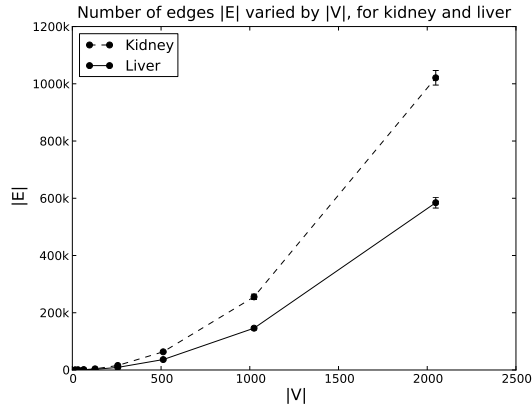


Figure 7: #Edges (in thousands) in generated liver and kidney compatibility graphs (100 graphs per $|V|$). The kidney graphs are denser than the liver graphs.

Figure 7 plots the average number of edges in the liver-only compatibility graphs, using the generator in this paper, against the average number of edges in the kidney compatibility graphs generated by the state of the art, as the number of candidate-donor pairs increases. The kidney compatibility graphs are, for graph sizes above 64, denser than comparably-sized liver compatibility graphs. This is interesting because it shows that, even though the liver exchange graphs do not need to take %PRA (i.e., HLA incompatibility) into account, their sensitivity to age and weight distri-

butions proves to be more constricting than HLA sensitivity! Regardless, neither the liver nor the kidney graphs are sparse in the classical sense of the word: at $|V| = 1024$, the number of edges in the liver graph is 26% of the total possible edges in a 1024-clique. This lack of sparsity drives the experimental computational complexity of solving the real-world clearing problem.

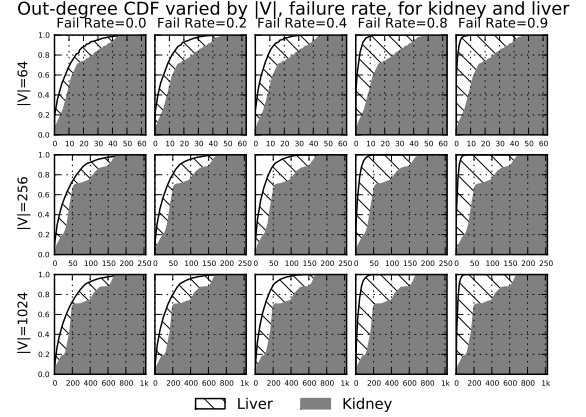


Figure 8: Cumulative distribution functions of the out-degree of vertices as we increase $|V|$ (varies per row) and exogenous failure rate f (varies by column), shown for the liver graphs (in white) and kidney graphs (in gray). Note the divergence between kidney and liver graph as the exogenous failure rate increases, as well as the three qualitative sections in the kidney graphs due to the three different %PRA classes.

Figure 8 enumerates the differences in the out-degree of the vertices in compatibility graphs for liver-only exchange generated using our algorithm (shown in white) and compatibility graphs for kidney exchange from the Saidman et al. generator (shown in gray). The size of the graph, $|V|$, is held constant along the rows, while the exogenous failure rate (f) between two otherwise compatible candidates and donors is held constant in each column. We vary $|V|$ and $f \in \{0.0, 0.2, \dots, 0.9\}$. Note that there is no notion of an exogenous failure rate in the kidney graphs (although the %PRA virtual crossmatch simulation is similar to an exogenous failure rate, but not parameterized); as such, the kidney exchange graphs vary only in terms of cardinality.

The cumulative distribution functions over the out-degrees of vertices, shown in Figure 8, exhibit interesting behavior. For example, there are more vertices with low degree in the liver exchange graphs than in the kidney exchange graphs. More interesting is the behavior exhibited by the kidney exchange graphs as $|V|$ increases. For instance, when $|V| = 1024$, we see three distinct out-degree sections in the kidney exchange graphs. These are an artifact of the somewhat ad-hoc method of doing %PRA virtual crossmatch tests in the Saidman et al. generator. The generator groups pairs into three sensitivity levels ("high", "medium", and "low"). As $|V|$ increases, those patients who are highly sensitized tend toward very few edges, while those at the medium and low sensitivity levels tend toward a medium and high number of edges, respectively. We believe that this is an

artifact of the generator by (Saidman et al. 2006) and is not representative of the real kidney exchange data. Our generator (even if used for kidneys) does not have such coarse artifacting because it can bucket sensitivity into finer-grained classes.

Appendix B: Theoretical Statements

We now support the theoretical claims made in the paper. Specifically, we provide a proof of Proposition 1 in the *sparse* organ exchange model, as stated in the paper. We also formalize (as Proposition 3) our statement at the end of the theory section regarding a linear gain in the size of the efficient matching of a combined *dense* multi-organ exchange.

Sparse Kidney Exchange

We now restate and sketch a proof of Proposition 1.

Proposition 1. *Consider $\beta > 0$ and $\gamma > 0$, kidney compatibility graph D_K with n_K pairs and $t(n_K) = \beta n_K$ altruistic donors, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. Then any efficient matching on $D = \text{join}(D_K, D_L)$ matches $\Omega(n_K)$ more pairs than the aggregate of any such efficient matchings on D_K and D_L (with probability approaching 1 as n_K approaches ∞).*

Proof sketch. The proposition follows from the proof of Theorem 5.4 of Ashlagi et al. (2012), which directly supports a similar result as Theorem 5.2 of Ashlagi et al. (2012). In that Theorem 5.4 (which assumes a kidney exchange graph similar to ours, with no altruistic donors), they show that there are a linear in n (n_K for us) number of “good cycles” of some constant length c . These “good cycles” have a single vertex u in the lowly-sensitized portion of D_K that is *only* connected to a single vertex v_1 in the highly-sensitized portion of D_K (and possibly other vertices in the lowly-sensitized portion). From v_1 there then exists a path $\langle v_1, \dots, v_{c-1} \rangle$ of highly-sensitized vertices with out- and in-degree one such that v_{c-1} connects back to u . Finding that path $\langle v_1, \dots, v_{c-1} \rangle$ relies on a well-known result (see, e.g., (Janson, Luczak, and Rucinski 2011)) that there exist linearly many isolated tree-like structures in a sparse graph (like the one induced by our highly-sensitized vertices). They show an additive linear gain in increasing cycle caps by first taking some optimal cover of cycles of length at most c and augmenting it to include enough of these “good cycles” of length at most $c + 1$ —of which there are linearly many—resulting in the gain.

We assume a constant cycle cap of c and no chain cap, which mimics real-world kidney exchanges and would probably be the case in a fielded liver exchange (if altruistic donors existed). Note that regardless of cycle cap, any efficient matching will match all lowly-sensitized pairs, via direct application of well-known matching results on dense Erdős-Rényi graphs. Under this constant cycle cap assumption, there exist a linear number of highly-sensitized vertices in the liver pool D_L that remain unmatched by an efficient matching of cycles of length at most c (recall there are no chains in the liver pool). These are the linearly many isolated highly-sensitized paths that are part of “good cycles”

of length strictly greater than c and thus cannot be legally matched. By gluing the two pools D_L and D_K together, these isolated vertices gain access to a linear number of altruists who, as in Theorem 5.2 of Ashlagi et al. (2012), act as the u vertex in “good cycles” of length greater than c that are now no longer required to connect back to u . Given a linearly larger combined pool (since the remainder of highly-sensitized liver pairs is linear in the size of D_L , which is linear in the size of D_K), we get an additive linear overall gain via the results of Theorem 5.4 of Ashlagi et al. (2012) by combining pools. \square

Dense Kidney Exchange

At the end of the theory section, we state: “... the efficiency results in the dense model without chains and with chains (Proposition 5.1 of Ashlagi and Roth (2011) and Theorem 1 of Dickerson, Procaccia, and Sandholm (2012b), respectively) can be applied directly to independent liver exchange and multi-organ exchange to yield efficient matchings consisting of cycles (and chains) of length at most 3, with linear expected overall gain from combining pools (given a linear number of altruists) for large enough compatibility graphs.” We formalize and support this statement in this section.

The Dense Model We begin by overiewing the *dense* model of kidney exchange. This model concentrates on blood types of donors and patients. At a very high level, human blood is split into four types—O, A, B, and AB—based on the presence or absence of type A and type B proteins. Ignoring other potential complications, a type O kidney can be transplanted into any patient; type A and B kidneys can be transplanted into A and B patients respectively or an AB patient; and type AB kidneys can only be transplanted into type AB patients. Therefore, some patients are more difficult to match with a random donor than others. O-patients are the hardest to match because only O-type kidneys can be given to them. Similarly, O-donors are the easiest to match.

An *under-demanded* pair is any pair such that the donor is not ABO-compatible with the patient. If an under-demanded pair contains only type A and B blood, it is called *reciprocal*. Any pair in the pool such that the donor is ABO-compatible with the candidate is called *over-demanded*. Furthermore, if a donor and candidate share the same blood type, they are a *self-demanded* pair. Under-demanded and reciprocal pairs are intuitively “harder” to match than over-demanded and self-demanded pairs. Note that this is not necessarily the case if sensitization, the probability of matching with a random donor, is considered. For example, an A-type patient who is lowly-sensitized is typically easier to match than an O-type patient who is highly-sensitized; however, the dense model does not consider different degrees of sensitization. The dense model critically assumes that a donor and patient who are blood type compatible are tissue type incompatible with *constant* probability \bar{p} . This differs from the model we used in Propositions 1 and 2, where lowly-sensitized patients had a constant edge probability while highly-sensitized patients did not (which more closely mimics reality). It also denotes by μ_X the frequency of blood type X , and assumes $\mu_O < 3\mu_A/2$ and an ordering $\mu_O > \mu_A > \mu_B > \mu_{AB}$. The

United States national blood type distribution satisfies these constraints.

Dense Theoretical Results Under the realistic assumptions on blood type distributions stated above, assuming no chains and only patients who need kidneys, Proposition 5.1 of Ashlagi and Roth (2011) states that an efficient allocation exists (with high probability) that uses only cycles of length at most 3. Theorem 1 of Dickerson, Procaccia, and Sandholm (2012b) extends this result in a pool with chains (but still only patients who need kidneys), stating that an efficient allocation exists (with high probability) using only cycles and chains of length at most 3. Both of these results are “in the large” and rely on the fact that the size of a set S will be very close to its expectation as $|S| \rightarrow \infty$. Figure 9 reproduces the efficient allocation from Dickerson, Procaccia, and Sandholm as an aid to the reader.

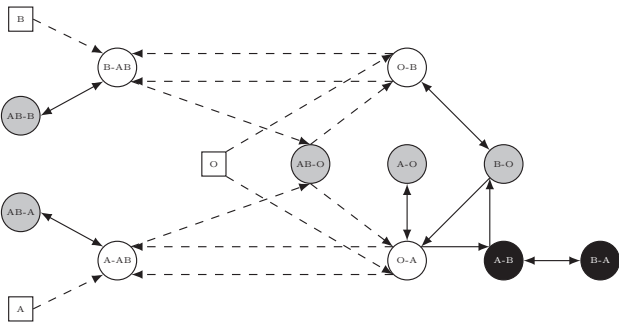


Figure 9: Efficient allocation presented in Dickerson, Procaccia, and Sandholm (2012b). Altruistic donors willing to give kidneys are shown as rectangles; candidate-donor pairs as ovals. Over-demanded pairs are gray, under-demanded are white, and reciprocal pairs are black. Dashed edges are those that may be matched in cycles or chains, while complete edges are those that are matched in cycles only.

Only Livers. We first look at dense *liver* exchange in this model. We note that the blood type distributional requirement is satisfied by patients in need of livers, just as it is with patients who need kidneys. Thus, under the dense model, a liver-only compatibility graph looks exactly the same as a kidney-only compatibility graph (albeit with no chains). Thus, the efficiency result of Ashlagi and Roth (2011) can be applied directly to liver-only compatibility graphs. If an altruistic liver donor existed in a liver-only compatibility graph, then the result of Dickerson, Procaccia, and Sandholm (2012b) would be directly applicable instead.

Multi-Organ Exchange. Next, we consider dense *multi-organ* exchange in this model. In this model, there will exist altruistic donors willing to give a kidney but not a liver, as motivated earlier in our paper.

We assume the same blood type distributions of Ashlagi and Roth (2011) for both liver and kidney patients and donors. We also assume a directed multi-organ dense compatibility graph D , with n_K pairs needing a kidney and $n_L = \gamma n_K$ pairs needing a liver, for some $\gamma > 0$. As motivated in our paper, altruistic kidney donors will not donate

directly to liver patients, but may trigger chains that result in a kidney pair donating to a liver pair. Thus, there are no outgoing edges in D from altruistic kidney donors to pairs needing a liver.

In Proposition 3, we show that if there are enough altruistic kidney donors, the size of an efficient matching on D is larger by an additive linear fraction than the size of the aggregate of efficient matchings on D_L and D_K , the subgraphs induced by only the vertices consisting of pairs needing livers and kidneys, respectively. Formally, let D_K^X be the subgraph induced by only altruistic kidney donors of blood type $X \in \{O, A, B, AB\}$.

Proposition 3. Consider $\beta^A = \mu_A \mu_{AB}$, $\beta^B = \mu_B \mu_{AB}$, and $\gamma > 0$, kidney compatibility graph D_K with n_K pairs, and liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. If at least one of $|D_K^A| > \beta^A n_K$ or $|D_K^B| > \beta^B n_K$, then any efficient matching on $D = \text{join}(D_K, D_L)$ matches $\Omega(n_K)$ more pairs than the aggregate of any such efficient matchings on D_K and D_L (with probability approaching 1 as $n_K \rightarrow \infty$).

Proof sketch. From Ashlagi and Roth (2011), if a vertex v participates in an exchange with some under-demanded vertex v' , then v helps v' . Let pairs of type $X-Y$ have X -type patients and Y -type donors, for $X, Y \in \{O, A, B, AB\}$. Note that AB-altruists cannot help under-demanded pairs, A- and B-altruists can only help A-AB and B-AB, respectively, under-demanded pairs, and O-donors can trigger two types of chains of length 3 containing under-demanded pairs: O-altruist, O-A pair, A-AB pair, or O-altruist, O-B pair, B-AB pair.

First, take the efficient matching result of Ashlagi and Roth (2011) and apply it to D_L . Only (some) under-demanded liver vertices remain unmatched. Let D_L^U represent these vertices. Second, apply the efficient matching result shown in Figure 9 to D_K . Again, only (some) under-demanded kidney vertices D_K^U remain unmatched.

As in Dickerson, Procaccia, and Sandholm (2012b), since applying these two matchings results in all over-demanded, self-demanded, and reciprocally-demanded pairs being matched (assuming $|S|$ approaches its expectation as $|S| \rightarrow \infty$ for any set S), we must only exhaustively consider all ways of matching under-demanded pairs. Bolded items in the list trigger a linear gain in the combined efficient match, while all other items show a loss in efficiency of at most zero. This guarantees a linear gain overall.

- **AB-donors:** Altruistic AB-donors can only help over- and self-demanded (AB-AB) pairs, both of which are matched entirely.
- **A-donors:** Of the under-demanded pairs, altruistic A-donors can only help A-AB pairs. In the matching in Figure 9, A-donors donate to the A-AB pairs until one of the two sets is exhausted. Under our assumption, $|V_K^A| > \mu_A \mu_{AB} n_K = |V_K^{A-AB}|$, so the A-AB set will be exhausted, leaving some A-donors unallocated. These remaining A-donors can be threaded into the liver pool through A-A kidney pairs to match with remainder under-demanded A-AB liver pairs. This use of an A-A results in 0 efficiency loss, since there remains a perfect matching in

V^{A-A} ; thus, we gain 1 match for each of the remaining $|V_K^A| - \mu_A \mu_{AB} n_K$ A-donors, which is a gain of $\Omega(n_K)$.

- **B-donors:** Of the under-demanded pairs, altruistic B-donors can only help B-AB pairs. Under a symmetric argument as the A-donors above, combining pools gains $|V_K^B| - \mu_B \mu_{AB} n_K$ extra matches by threading through B-B kidney pairs into the unallocated under-demanded liver pool, which is a gain of $\Omega(n_K)$.
- *O-donors:* In the matching of Dickerson, Procaccia, and Sandholm (2012b), some O-donors may be used in 2-chains with remaining under-demanded pairs in D_K . It is possible that these O-donors could be threaded through an under-demanded kidney pair into an under-demanded liver pair to form a 3-chain at utility gain of 1 (but not necessary for this proof).
- *Non-altruistic vertices:* Self-demanded and reciprocally-demanded pairs cannot help under-demanded pairs without involving altruistic donors or over-demanded pairs. AB-O vertices are the only pairs that can help at most two under-demanded pairs (either O-A and A-AB, or O-B and B-AB). In the Dickerson, Procaccia, and Sandholm (2012b) allocation, most AB-O pairs are used in 3-cycles with two under-demanded pairs; however, some may be used in 2-cycles with a single under-demanded pair. Reallocating these are not necessary for this proof.

Since at least one of the minimum size constraints on the set of altruistic A-donors or B-donors is satisfied by the proposition statement's assumptions, we are guaranteed $\Omega(n_K)$ additional matches by combining both pools. \square