Multi-Organ Exchange

John P. Dickerson   JOHN@CS.UMD.EDU
Department of Computer Science
University of Maryland
College Park, MD 20742 USA

Tuomas Sandholm   SANDHOLM@CS.CMU.EDU
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213 USA

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 exploring the idea of large-scale *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 propose 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; it exists even when only a small but constant portion of the donors on the kidney side of the pool are willing to donate a liver lobe. 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.

1. 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 donor 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 July 5, 2015, there were 101,257 patients waiting for a kidney, 15,268 waiting for a liver, and 9073 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 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* (Rapaport, 1986; Roth, Sönmez, & 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 explore large-scale *liver* exchange, which is similar to kidney exchange in some ways, but remains unexplored from a computational point of view.[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, Simpson, Pomposelli, & Pomfret, 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. The lack of altruistic donation, along with novel characteristics of the (in)compatibility of participants in a demographically-accurate liver exchange, leads to different matching behavior in theory and in practice.

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, and is present even when only a small but constant fraction of one side of the combined pool is willing to donate to a pair on the other side. 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.

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.

2. 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

---

1. Recently, small-scale liver exchanges have been manually arranged by medical professionals. In Korea, 16 candidates swapped *by hand* willing donors in a single hospital over the course of six years (Hwang, Lee, Moon, Song, Ahn, Kim, Ha, Jung, Kim, Choi, Park, Yu, Choi, Park, & Ha, 2010); similarly, in Hong Kong, 2 candidates hand-swapped willing donors (Chan, Lo, Yong, Tsui, Ng, & Fan, 2010). This shows the feasibility of the idea at a small scale (Segev & Montgomery, 2010).

determine an optimal set of swaps. We also overview recent related work in multi-hospital kidney exchange.

2.1 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).[2] 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, Kopke, Pelletier, Segev, Rutter, Fabrega, Rogers, Pankewycz, Hiller, Roth, Sandholm, Ünver, & Montgomery, 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 significantly increased medical risk to living donors of livers, we do not expect many (or possibly any) altruistic donors outside of kidney exchanges (Cheah et al., 2013).

Figure 1 gives an example organ exchange compatibility graph, where pairs on the left, shown with a dark boundary, have patients in need of a kidney while pairs on the right, shown with a light boundary, have pairs in need of a liver. Possible cycles exist entirely in either the kidney or liver pool (for example, the 2-cycle $\langle (d_2 \to p_3), (d_3 \to p_2) \rangle$ and 2-cycle $\langle (d_5 \to p_6), (d_6 \to p_5) \rangle$, respectively) or between the two pools (for example, the 3-cycle $\langle (d_2 \to p_5), (d_5 \to p_3), (d_3 \to p2) \rangle$, which involves a single liver pair and two kidney pairs). Additionally, a single altruistic donor a exists in the pool and is willing to give his or her kidney—but not liver—to a patient, whose paired donor will then either donate a kidney *or* liver to a compatible patient in the pool (for example, via the chain $\langle (a \to p_1), (d_1 \to p_4), (d_4 \to p_7), (d_7 \to \cdot) \rangle$, with the final donor d_7 either donating to the deceased donor waiting list or remaining in the pool as a future altruistic donor).

2.2 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

---

2. This is a more general notion of matching than the traditional idea of simply pairing items.
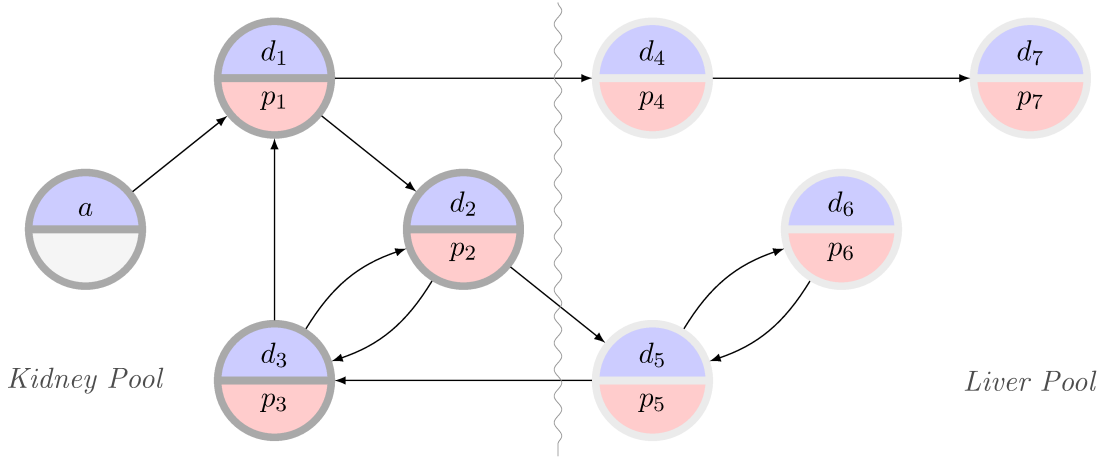
Figure 1: An example joint liver-kidney compatibility graph.

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.

In the small example compatibility graph shown in Figure 1, with $L = 3$, a maximum cardinality matching *without* chains includes five pairs via the 3-cycle and 2-cycle:

$$\{\langle(d_1 \to p_2), (d_2 \to p_3), (d_3 \to p_1)\rangle, \langle(d_5 \to p_6), (d_6 \to p_5)\rangle\}.$$

With chains—but maintaining two separate pools instead of combining the pools into a joint exchange—we can achieve a maximum cardinality matching with the same number of pairs but with a lower cycle cap ($L = 2$ instead of $L = 3$) via one chain and two 2-cycles:

$$\{\langle(a \to p_1), (d_1 \to \cdot)\rangle, \langle(d_2 \to p_3), (d_3 \to p_2)\rangle, \langle(d_5 \to p_6), (d_6 \to p_5)\rangle\}.$$

This lower length cap may have practical advantages. For example, shorter cycles are less likely to fail after the algorithmic matching but before transplantation, and thus may lead to improved matching in practice (Li, Kalbfleisch, Song, Zhou, Leichtman, & Rees, 2011; Dickerson, Procaccia, & Sandholm, 2018). Some fielded kidney exchanges explicitly favor shorter cycles over longer ones (?). While our model could support such extra constraints or preferences, we do not include them in this paper.

Finally, if we allow a combined liver-kidney exchange, the cardinality of the maximum matching increases to seven pairs. This is achieved by "threading" a chain started by the altruist a through a kidney-needing pair into the greater liver pool to match two previously unmatchable pairs, as well as by using the same two 2-cycles as before:

$$\{\langle(a \to p_1), (d_1 \to p_4), (d_4 \to p_7), (d_7 \to \cdot)\rangle, \langle(d_2 \to p_3), (d_3 \to p_2)\rangle, \langle(d_5 \to p_6), (d_6 \to p_5)\rangle\}.$$

While the clearing problem can be solved easily on very small graphs like that in Figure 1, it quickly becomes intractable on larger graphs. The standard algorithmic method for optimally clearing kidney exchanges is integer programming (Roth, Sönmez, & Ünver, 2007). Formally, 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 $x_c \in \{0, 1\}\ \forall c \in C(L)$, we must solve the following integer linear program:

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

The clearing problem with any fixed $L > 2$ is NP-complete (Abraham, Blum, & 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 et al., 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 et al., 2007; Dickerson, Manlove, Plaut, Sandholm, & Trimble, 2016). We will later discuss this algorithm in more detail as well as enhancements to it for liver exchange and multi-organ exchange.

2.3 Related Work in Multi-Hospital Kidney Exchange

We are not the first to consider combining exchanges in general; rather, we are the first to consider combining *exchanges corresponding to different organs*, and we are the first to approach the intricacies of that combination from a theoretical and empirical point of view. Indeed, fielded centralized *kidney* exchanges typically consist of the merged pools of multiple participating hospitals. For example, the United Network for Organ Sharing (UNOS) kidney exchange includes 143 hospitals in the US. Each of these hospitals enters (possibly some subset of) the set of candidate-donor pairs and altruistic donors that have registered at their center into the centralized exchange, and then the exchange recommends a matching based on its clearing engine. Each hospital can be seen as having its own private exchange; then, questions can be asked about the possible gains in match efficacy based on the number or size of participating hospitals, or the truthfulness of their reporting.

Ashlagi, Fischer, Kash, and Procaccia (2015) look at the multi-hospital exchange problem from a game-theoretic point of view, where participating hospitals can manipulate the exchange by misreporting their private set of candidate-donor pairs. They show that in the worst case no deterministic strategy-proof mechanism can provide more than 1/2 of the truthful maximum matching, and no randomized strategy-proof mechanism can provide more than 7/8 of the truthful maximum matching. That bound was tightened for the two hospital case by Caragiannis, Filos-Ratsikas, and Procaccia (2015). That model looks at 2-cycles only (represented as an undirected graph), while ours looks at 2- and 3-cycles and altruist-initiated chains. They also operate in a dense theoretical model only, while we give results in both dense and an arguably more realistic sparse model.

Ashlagi and Roth (2014) and Toulis and Parkes (2015) also analyze the multi-hospital exchange problem from a game-theoretic point of view. Ashlagi and Roth (2014) show that, in general, a lack of participation of all hospitals can be very costly (although it is possible, at the cost of a few "lost" matches, to guarantee individual rationality and strategy proofness). Those results are in a dense model with both 2- and 3-cycles, but no chains. Toulis and Parkes (2015) present a new multi-hospital mechanism for multi-hospital exchange and compares against their results.

As we do not consider incentive issues in the present work, perhaps most related to our work is a general matching result by Toulis and Parkes (2015). They show, in a dense kidney exchange model (which we overview in Section 3.2), that given m transplant centers, each with n candidate-donor pairs, the expected gain in number of matches to an individual hospital by entering a centralized exchange with full participation is roughly $\Omega(\sqrt{n})$. This result is in a model without chains; indeed, we will show that the inclusion of chains results in a linear gain in number of matches.

We are now ready to present the results of the paper. Section 3 addresses liver and multi-organ exchange in adaptations of the two most common theoretical models of kidney exchange, and proves results in both models regarding the efficiency gains of multi-organ exchange over independent single-organ exchanges. Section 4 moves from theory to practice and presents our method of generating demographically accurate kidney, liver, and joint kidney-liver compatibility graphs. (Appendix A provides more detail about this process, as well as a quantitative comparison of the resulting graphs against the status quo kidney exchange generator.) Section 4 also presents the clearing algorithm we use to solve the multi-organ exchange problem in practice. Section 5 shows experimental results on liver and multi-organ exchanges, and gives strong empirical support to our earlier theoretical results showing that multi-organ exchange results in a greater number of matches than two independent exchanges. (Appendix B presents additional experimental results.) We conclude in Section 6 with some thoughts on fielding a liver or multi-organ exchange, as well as future research directions.

3. Combining Exchanges Yields Linearly More Matches

In this section, we prove that combining independent liver and kidney exchanges leads to a linear gain in the aggregate number of matches. We do this in multi-organ adaptations of the two standard models of kidney exchange. Section 3.1 works in a newer *sparse* model adapted from Ashlagi, Gamarnik, Rees, and Roth (2012), while Section 3.2 works in an older *dense* model like that used by Ashlagi and Roth (2014).[3] We obtain similar—but not identical—theoretical results in both models.

3.1 Sparse Model

We begin by adapting to the multi-organ exchange case a version of a recent random graph model for kidney exchange due to Ashlagi et al. (2012). They adapt sparse Erdős-Rènyi 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

---

3. While the most recent publication date of the work by Ashlagi and Roth (2014) is after that of the work by Ashlagi et al. (2012), the former paper appeared in 2011 as a conference paper, while the latter appeared as a conference paper in 2012 and is still under submission as a final journal paper.

pairs—representing *highly-sensitized*, hard-to-match patients—who have probability p_H of an incoming edge from each vertex in the pool. They 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 with n_K pairs, $t(n_K) > 0$; however, for liver exchange graphs D_L with n_L pairs, $t(n_L) = 0$ (i.e., there are no altruistic liver donors). Furthermore, we introduce an additional constant probability $p_{K \to L} > 0$ to address the likelihood that some paired kidney donors will be unwilling to donate a liver lobe instead of a kidney. For notational simplicity, we will not introduce the complementary probability $p_{L \to K}$, which would represent the probability that a paired liver donor would be willing to donate a kidney if matched, because in practice we believe any potential liver donor would prefer donating the vastly "easier" kidney to donating a liver (i.e., $p_{L \to K} = 1$ in practice). Still, were this not to be the case, the qualitative results to follow would still hold. We do not assume that p_L (resp. p_H) for D_K equals p_L (resp. p_H) for D_L. Formally, let $p_{L,\{K,L\}} \in (0,1]$ be different constants and $p_{H,\{K,L\}} = c_{\{K,L\}}/n_{\{K,L\}}$ for D_K and D_L and positive constants c_K and c_L. When the usage is obvious from context, for expositional ease we will still use p_H, p_L, and c.

Now, define the graph join operator $D = join(D_K, D_L, p_{K \to L})$ between a kidney exchange graph D_K and liver exchange graph D_L as follows. Flip a $p_{K \to L}$-weighted coin for each patient-donor pair in D_K; if heads, this pair is willing to give a liver to a pair in D_L if matched (or a kidney to a different pair in D_K), otherwise the paired donor is only willing to give a kidney. Next, add directed edges between candidate-donor pairs in both pools in accordance with each pair's associated probability (e.g., $p_{L,L}$ from any kidney pair to a lowly-sensitized liver pair or $p_{H,K}$ from any liver pair to a highly-sensitized kidney pair), *except* for those vertices with paired kidney donors who are unwilling to donate livers. Do not add any edges from the $t(n_K)$ altruistic donors in D_K to vertices in D_L (since altruistic kidney donors are unwilling to donate a liver lobe).

In the following theoretical results, we consider cycles of length at most some constant but chains of any length; this models current practice in kidney exchange, and would model potential future fielded liver exchanges. Thus, an *efficient matching* allocates the maximum number of transplants in cycles of size no more than some constant and chains of any length. Our 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 contrast both theoretical results at the end of this section.

Proposition 1. *Consider $\beta > 0$ and $\gamma > 0$, sparse kidney compatibility graph D_K with n_K pairs and $t(n_K) = \beta n_K$ altruistic donors, and sparse liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. Then for any constant cycle cap and $p_{K \to L} > 0$, any efficient matching on $D = join(D_K, D_L, p_{K \to L})$ 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. The proposition follows from the proof of Theorem 5.4 by Ashlagi et al. (2012), which directly supports a similar result as Theorem 5.2 by Ashlagi et al. (2012). In that Theorem 5.4 (which assumes a kidney exchange graph similar to ours, with no altruistic

645

donors), they show that there are a linear in n ($n_{\{K,L\}}$ for us) number of "good cycles" of some constant length z. These "good cycles" have a single vertex u in the lowly-sensitized portion of $D_{\{K,L\}}$ that is *only* connected to a single vertex v_1 in the highly-sensitized portion of $D_{\{K,L\}}$ (and possibly other vertices in the lowly-sensitized portion). From v_1 there then exists a path $\langle v_1, \ldots, v_{z-1} \rangle$ of highly-sensitized vertices with out- and in-degree one such that v_{z-1} connects back to u. Finding that path $\langle v_1, \ldots, v_{z-1} \rangle$ relies on a well-known result (see, e.g., Janson, Luczak, & 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 z and augmenting it to include enough of these "good cycles" of length at most $z + 1$—of which there are linearly many in n—resulting in the gain.

Recall that we assume a constant cycle cap of z and no chain cap. Note that regardless of cycle cap, any efficient matching will match all lowly-sensitized pairs (w.h.p. as n grows), via direct application of well-known matching results on dense Erdős-Rènyi graphs (see, e.g., Janson et al., 2011). 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 z (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 z and thus cannot be matched. By gluing the two pools D_L and D_K together, these isolated vertices gain access—through chains that start in D_K, of which there are linearly many in n_K—to a linear number of altruists who, as in Theorem 5.2 by Ashlagi et al. (2012), act as the u vertex in "good cycles" of length greater than z that are now no longer required to connect back to u.

Formally, fix an efficient matching M_K^* in D_K alone and an efficient matching M_L^* in D_L alone, which is disjoint from M_K^* by construction. The aggregate size $m_I = |M_K^* \cup M_L^*| = |M_K^*| + |M_L^*|$ of these two matchings is the size of the efficient matching under the setting of independent liver and kidney exchanges. We will show that by combining exchanges, all pairs matched in $M_I^* = M_K^* \cup M_L^*$ can be matched *while also matching a linear number of previously unmatched pairs.*

Figure 2 overviews the augmented matching we will construct. On the left side is a linear-in-n_K number a of chains created from the $t(n_K)$ kidney altruists and only pairs in D_K; these structures exist and are already in the efficient matching M_K^*—and thus in M_I^* (Ashlagi et al., 2012). Now, for each of the a chains, consider the final kidney pair in the chain. By assumption, with constant positive probability $p_{K \to L}$, that pair is willing to participate in a donation that crosses into the liver pool. This preference is determined independently across all pairs, so in expectation $(p_{K \to L})a$ chains will be willing to thread into the liver pool. We will extend these chains in our combined matching, and will leave the remaining $(1 - p_{K \to L})a$ in expectation kidney-only chains allocated as they were in the original matching M_K^*. These estimates are concentrated about their mean via standard concentration bounds.

Given the original constant cycle cap z, for any larger constant integer $z' > z$, there exist $b = \beta n_L \in \Omega(n_L)$ isolated paths in the liver pool as well, where β is a positive constant (Ashlagi et al., 2012); these are shown on the right side of the graph in Figure 2. These consist entirely of pairs that are not matched in M_L^*, and thus are not matched in M_I^*, either. Now, for each of those $(p_{K \to L})a$ kidney pairs at the end of a chain in M_K^* who
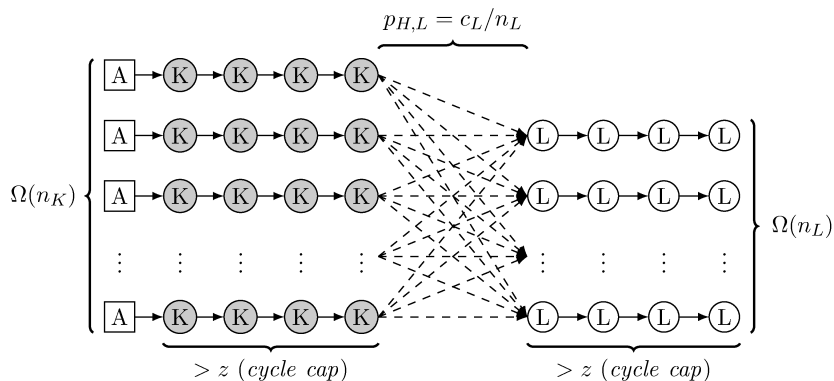
Figure 2: A linear number of chains threading into the liver pool. Altruistic donors are shown as a boxes, while pairs in D_K and D_L are shown as circles with inscribed Ks and Ls, respectively. Pairs that are matched *only* when exchanges are combined are shown in white.

are willing to donate a liver lobe, the probability p_1 that at least one edge exists from that pair to at least one liver pair at the head of one of the βn_L isolated paths of length z' in D_L is

$$p_1 = 1 - (1 - c_L/n_L)^{\beta n_L} =_{n_L \to \infty} 1 - e^{-c_L \beta} \in O(1). \tag{1}$$

This connectivity is determined independently for each of the $(p_{K \to L})a$ kidney pairs, leading to $p_1(p_{K \to L})a \in \Omega(n_K)$ in expectation pairs with at least one connection to a liver pair at the head of an isolated path in D_L. Edges that may exist between the two sets of paths—the chains in M_K^* and the isolated unmatched paths in D_L—are shown as dashed lines in Figure 2. Then all that is left to do is lower bound the size of a maximum bipartite matching on the graph induced by those dashed edges that actually exist, which we denote by G. We prove this bound with a balls and bins argument.

As a lower bound on the size of that matching, assume that each of the $p_1(p_{K \to L})a$ kidney pairs with at least one edge crossing into the liver pool has *exactly* one edge crossing into the liver pool, selected uniformly at random from its true set of edges. This induces an injective mapping of kidney pairs to liver pairs, and also a subgraph $G' \subset G$ of the full bipartite graph over which we are performing a maximum matching; thus, the cardinality of a maximum matching in this reduced subgraph G' is a lower bound on the cardinality of a maximum matching in the fully realized bipartite graph G. Furthermore, the size of a maximum matching in this subgraph G' is equal to the number of pairs on the liver side with at least one incoming edge. We calculate that now, treating each kidney pair as one of $p_1(p_{K \to L})a$ balls being dropped uniformly at random into one of the βn_L bins (representing a liver pair with at least one incoming edge in the full bipartite graph G).

Formally, index the liver pairs $[I] = \{1, 2, \ldots, \beta n_L\}$, let random variable Y represent the number of liver pairs with zero incoming edges in the subgraph G', and let X_i be a binary random variable that is set to 1 if pair $i \in [I]$ has no incoming edges. Then $\mathbb{E}[X_i] = (1 - 1/\beta n_L)^{p_1(p_{K \to L})a} \in O(1)$. Thus, with constant $\mathbb{E}[X_i] < 1$, and $\mathbb{E}[Y] = \mathbb{E}[\sum_{i \in [I]} X_i] =$

$\sum_{i \in [I]} \mathbb{E}[X_i]$ by linearity of expectation, we have $E[Y] = \beta n_L E[X_i]$, a constant fraction (strictly less than one) of the liver pairs.

Finally, let random variable Z represent the number of liver pairs with at least one incoming edge in the subgraph. Then $\beta n_L = Z + Y$, and $\mathbb{E}[Z] = \beta n_L - E[Y] = \beta n_L (1 - E[X_i]) \in \Omega(n_L)$. With $\mathbb{E}[Z]$ a constant positive fraction of the liver pairs, we have a maximum matching of size $\Omega(n_L)$ in G', and thus a maximum matching of size $\Omega(n_L)$ in G.

We now construct our final matching M_C^* in the combined pool. Take $M_C^* = M_I^*$. For those kidney-only chains in $M_K^* \subseteq M_I^*$ ending in (i) a kidney pair willing to donate a liver lobe with (ii) at least one edge into the liver pool, add edges in accordance with a fixed maximum matching in G. Each of these $\Omega(n_L)$ edges connects to an isolated path of length $z' > z \geq 0$ of *previously unmatched* pairs in the liver pool. Add each of these pairs to M_C^*, resulting in a linear gain in matching size over M_I^*.

$\square$

As shown in Proposition 1, the presence of a linear number of altruistic kidney donors in a multi-organ exchange results in a linear gain in the overall number of pairs matched in an efficient matching relative to the aggregate of efficient matchings in independent kidney and liver exchanges. This is realized specifically by giving those highly-sensitized liver pairs who are unmatchable using only cycles in the liver pool access to more flexible, longer altruist-initiated chains that thread out of the kidney pool. In the following Proposition 2, we restrict the number of altruistic kidney donors to a constant and show that even in this constrained setting, with constant positive probability, chains that thread out of the kidney pool into the liver pool allow for a linear gain in overall number of matches.

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

Proof. For small enough λ and large enough c_K, where $p_{H,K} = c_K/n_K$, 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, for large enough c_L, where $p_{H,L} = c_L/n_L$, there exists a set S_L of pairs in D_L that cannot be matched in a cycle of constant capped length and would have to be matched in a chain or not matched at all—since we assume that no altruists willing to directly donate a liver lobe exist, these pairs would go unmatched if exchanges for different organs operated independently.

To aid the reader, Figures 3 and 4 accompany the statements in this part of the proof; Figure 3 corresponds to our status quo case of two separate kidney and liver exchanges, while Figure 4 corresponds to the setting of this proof, where the exchanges are combined via the *join* operator. We will make use of a general result on sparse random directed graphs by Krivelevich, Lubetzky, and Sudakov (2013): as $c_{\{K,L\}}$ increases, a directed path of length approaching $|S_{\{K,L\}}|$ in $S_{\{K,L\}}$ exists.
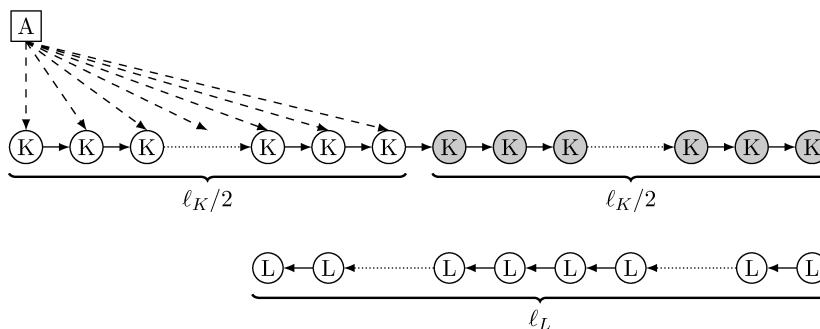
Figure 3: Relevant portion of the maximum matching for $S_K \subseteq D_K$ and $S_L \subseteq D_L$ in the *independent* exchanges case. An altruist is shown as a box, while pairs in S_K and S_L are shown as circles with inscribed Ks and Ls, respectively. Pairs that are matched with constant positive probability are shown in gray. Note that no pairs in S_L are matched.
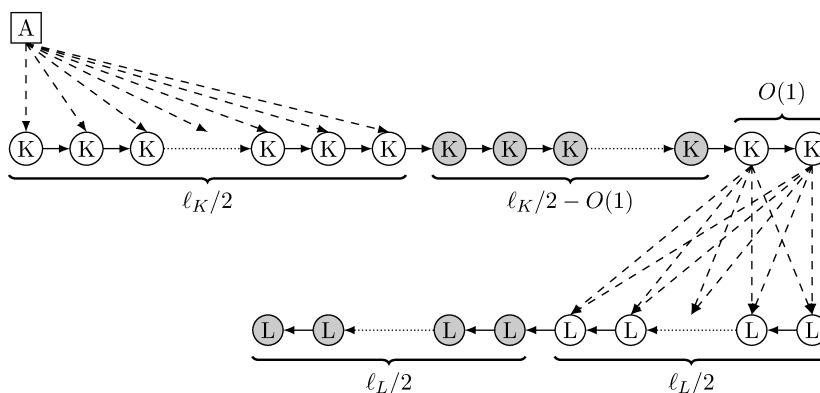

Figure 4: Relevant portion of the maximum matching for $S_K \subseteq D_K$ and $S_L \subseteq D_L$ in the *combined* exchange case. An altruist is shown as a box, while pairs in S_K and S_L are shown as circles with inscribed Ks and Ls, respectively. Pairs that are matched with constant positive probability are shown in gray. Note that a linear portion of the chain in S_L is matched with constant positive probability.

We first show the existence of long paths of vertices that must be matched by chains in each of D_K and D_L. Take the set $S_K \subseteq D_K$ of kidney pairs that must be matched via chains or not at all; then $|S_K| \geq n_K/2 \in \Omega(n_K)$ (Ashlagi et al., 2012). Similarly, take the set $S_L \subseteq D_L$ of liver pairs that must be matched via chains (threaded, in this case, through a willing kidney pair as determined by $p_{K \to L}$); then $|S_L| \geq n_L/2 \in \Omega(n_K)$. Via the results of Krivelevich et al. (2013), there exists a directed chain C_K of length $\ell_K \in \Omega(|S_K|)$ in S_K, and similarly there exists a directed chain C_L of length $\ell_L \in \Omega(|S_L|)$ in S_L. Then there exists constants α_K and α_L such that $\ell_K = \alpha_K n_K$ (the length of the chain in S_K) and $\ell_L = \alpha_L n_L$ (the length of the chain in S_L), respectively.

Take the first $\ell_K/2$ pairs in the head of C_K; then, the probability p_1 that a given altruistic donor has at least one outgoing edge to a pair in that head is

$$p_1 = 1 - (1 - c_K/n_K)^{\ell_K/2} = 1 - (1 - c_K/n_K)^{\alpha_K n_K/2} =_{n_K \to \infty} 1 - e^{-c_K \alpha_K/2} \in O(1). \quad (2)$$

Then, with constant positive probability p_1, the altruistic donor in the *independent exchanges* case matches m_I pairs, where $\ell_K/2 < m_I \leq \ell_K$; exactly 0 pairs in the chain C_L are matched. That matching is visualized in Figure 3.

In the *combined exchanges* case, a given pair in C_K has an outgoing edge to a given pair in C_L with probability $p_{K \to L} c_L/n_L > 0$. Take a positive constant $t > 0$ number of pairs in the tail of the C_K chain. Then, the probability p_2 that at least one of the t pairs in that tail has at least one outgoing edge to at least one pair in the first $\ell_L/2$ pairs at the head of the liver chain C_L is

$$\begin{aligned} p_2 &= 1 - \left[(1 - p_{K \to L} c_L/n_L)^{\ell_L/2}\right]^t = 1 - \left[(1 - p_{K \to L} c_L/n_L)^{\alpha_L n_L/2}\right]^t \\ &=_{n_L \to \infty} 1 - \left[e^{-p_{K \to L} c_L \alpha_L/2}\right]^t \in O(1). \end{aligned} \quad (3)$$

The independence assumptions above are valid because (i) the willingness of a kidney pair to give a liver is determined independently via the $p_{K \to L}$ parameter and (ii) the initial edge compatibility check between pairs in D_K and D_L is independent of the results of the $p_{K \to L}$ coin flip. Then, with constant positive probability $p_1 p_2$, the altruistic donor in the *combined exchanges* case matches m_C pairs, where $m_C > m_I + \ell_L/2 - t$, the guaranteed matched pairs in C_K minus a constant t pairs in that tail plus the guaranteed matched pairs in the tail of C_L. That matching is visualized in Figure 4.

Recall $\ell_L/2 \in \Omega(|S_L|)$, and with constant t, $\ell_L/2 - t \in \Omega(|S_L|)$. So, the gain in matches between the matching in Figure 3 and that in Figure 4 is

$$m_C - m_I > \ell_2/2 - t = \alpha_L n_L/2 - t = \alpha_L \gamma n_K/2 - t \in \Omega(n_K) \quad (4)$$

which occurs with constant probability at least $p_1 p_2 > 0$, where p_1 and p_2 are given in Equations 2 and 3, respectively. $\qquad \square$

3.1.1 DISCUSSION OF THEORETICAL RESULTS IN THE SPARSE MODEL

Intuitively, Propositions 1 and 2 show the theoretical efficacy of combining kidney exchange with alternate organ exchanges (where altruistic donation is less likely to be popular or deemed ethically acceptable). While 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 probability of not being highly sensitized constant quite small, and the result also holds with merely constant positive probability instead of holding with probability approaching one. We thus consider Proposition 1 to be 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.

3.2 Dense Model

Initial research on random graph models for organ exchange adapted *dense*—that is, constant probability of an edge existing—Erdős-Rènyi graphs to kidney exchange (Ashlagi & Roth, 2014; Dickerson, Procaccia, & Sandholm, 2012b). Fielded exchanges have proven to be somewhere in between dense—as we discuss now—and sparse—as in the theory above— in practice, and thus actual pools and their optimal matchings do not align with these dense models (Ashlagi et al., 2012; Ashlagi, Jaillet, & Manshadi, 2013; Dickerson, Procaccia, & Sandholm, 2014b; Dickerson et al., 2018; McElfresh & Dickerson, 2018). Still, we show that the efficiency results in the dense model with chains (Dickerson et al., 2012b, Thm. 1) can be applied to independent liver exchange and multi-organ exchange to yield efficient matchings with linear expected overall gain from combining the pools (given a linear number of altruists) for large enough compatibility graphs. We derive these results now.

We begin by overviewing the *dense* model of kidney exchange (Roth et al., 2004; Roth, Sönmez, & Ünver, 2005a, 2005b). This model concentrates on blood types of donors and patients. At a high level, human blood can be of four types—O, A, B, and AB—based on the presence or absence of type A and type B proteins. Ignoring other potential reasons for incompatibility, 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. Usually O-patients are the hardest to match because only O-type kidneys can be given to them. Similarly, O-donors are usually 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 (e.g., a pair with A-type patient and B-type donor, or vice versa), 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. 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. The dense model 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. We will use $V_{\{K,L\}}^{X\text{-}Y}$ to refer to the subset of vertices with patient and donor of blood type X and Y, respectively, in the kidney and liver compatibility graphs, and $V_{\{K,L\}}^{X}$ for the subset of vertices with altruistic donors with blood type X in the kidney and liver compatibility graphs.

Under the realistic assumptions on blood type distributions stated above, but assuming no chains and only patients who need kidneys, Proposition 5.1 by Ashlagi and Roth (2014) states that an efficient allocation exists (with high probability) that uses only cycles of length at most 3. That result is proven only with respect to cycles, that is, it assumes there are no altruistic donors. Dickerson et al. (2012b, Thm. 1) extends that result into a pool that also has 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 the set $S \in \{V_{\{K,L\}}^{X\text{-}Y}\} \cup \{V_{\{K,L\}}^{X}\}$ for any blood types X and Y will be very close to its expectation as $|S| \to \infty$.

3.2.1 ONLY LIVERS

We first look at *liver* exchange in the dense model. 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 (2014) can be applied directly to liver-only compatibility graphs. If altruistic liver donors existed in a liver-only compatibility graph, then the result of Dickerson et al. (2012b) would be directly applicable instead.

3.2.2 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 distribution and ordering as Ashlagi and Roth (2014) 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 constant $\gamma > 0$. As motivated earlier 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. Each kidney pair is willing to give to a compatible liver pair with some constant probability $p_{K \to L} > 0$. Thus, there are no outgoing edges in D from altruistic kidney donors to pairs needing a liver, but potentially some edges from kidney pairs to compatible liver pairs.

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. We achieve this linear gain via a similar high-level strategy to what was used in Propositions 1 and 2, which threaded kidney-altruist-initiated chains through willing non-altruist kidney pairs into the liver pool.

Formally, let V_K^X be the subset of vertices in D_K containing only altruistic kidney donors of blood type $X \in \{O, A, B, AB\}$.

Proposition 3. *Consider $\beta^A = \mu_A \mu_{AB}$, $\beta^B = \mu_B \mu_{AB}$, constants $\gamma > 0$ and $p_{K \to L} > 0$, dense kidney compatibility graph D_K with n_K pairs, and dense liver compatibility graph D_L with $n_L = \gamma n_K$ pairs. If at least one of $|V_K^A| > \beta^A n_K$ or $|V_K^B| > \beta^B n_K$, then any efficient matching on $D = join(D_K, D_L, p_{K \to L})$ 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 $n_K \to \infty$).*

Proof. We begin by adopting vocabulary from Ashlagi and Roth (2014); specifically, if a vertex v participates in an exchange with some under-demanded vertex v', then we say *v helps v'*. Pairs denoted by X-Y have X-type patients and 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 under-demanded pairs, respectively, 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 (2014) and apply it to D_L. Only (some) under-demanded liver vertices remain unmatched. Second, apply the efficient matching result of Dickerson et al. (2012b) to D_K. Again, only (some) under-demanded kidney vertices remain unmatched.

Figure 5 provides a visual representation of the full allocation we will construct by augmenting the two allocations mentioned above.[4] As in the work of Dickerson et al. (2012b), since applying the two initial matchings results in all over-demanded, self-demanded, and reciprocally-demanded pairs being matched (assuming $|S|$ approaches its expectation as $|S| \to \infty$ for any set $S \in V_{\{K,L\}}^{X\text{-}Y}$, $X, Y \in \{O, A, B, AB\}$), we must only exhaustively consider all ways of matching under-demanded pairs. We do this in the list below: bolded items trigger a linear gain in the combined efficient match, while all other items show no efficiency loss. This guarantees a linear gain overall.

- *AB-altruists:* Altruistic AB-donors can only help over- and self-demanded (AB-AB) pairs, both of which are matched entirely already in the separate exchanges.

- **A-altruists:** Of the under-demanded pairs, altruistic A-donors can only help A-AB pairs.[5] In the matching of Dickerson et al. (2012b), A-donors donate to the A-AB pairs until one of the two sets is exhausted. Under our assumption, $|V_K^A| > \mu_A \mu_{AB} n_K = |V_K^{A\text{-}AB}|$, 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 the remainder of under-demanded A-AB liver pairs.

  Given constant probability $p_{K \to L} > 0$ of a non-altruist kidney donor being willing to donate a liver lobe, and constant probability $\bar{p} > 0$ of an otherwise blood type

---

4. Figure 5 shows the full allocation up to symmetries between A-B and B-A pairs. By assumption, $\mathbb{E}[|V_K^{A\text{-}B}|] = \mathbb{E}[|V_K^{B\text{-}A}|]$ and $\mathbb{E}[|V_L^{A\text{-}B}|] = \mathbb{E}[|V_L^{B\text{-}A}|]$, but it could be the case in practice that one subgraph is larger than the other. We assume WLOG in Figure 5 that $|V_K^{A\text{-}B}| \geq |V_K^{B\text{-}A}|$ and $|V_L^{A\text{-}B}| \geq |V_L^{B\text{-}A}|$.

5. In the original compatibility graph, altruistic A-donors can also help reciprocal A-B pairs; however, by the earlier applications of the efficient matchings due to Ashlagi and Roth (2014) to the liver pool and the efficient matching due to Dickerson et al. (2012b) to the kidney pool, all reciprocal pairs are already matched.
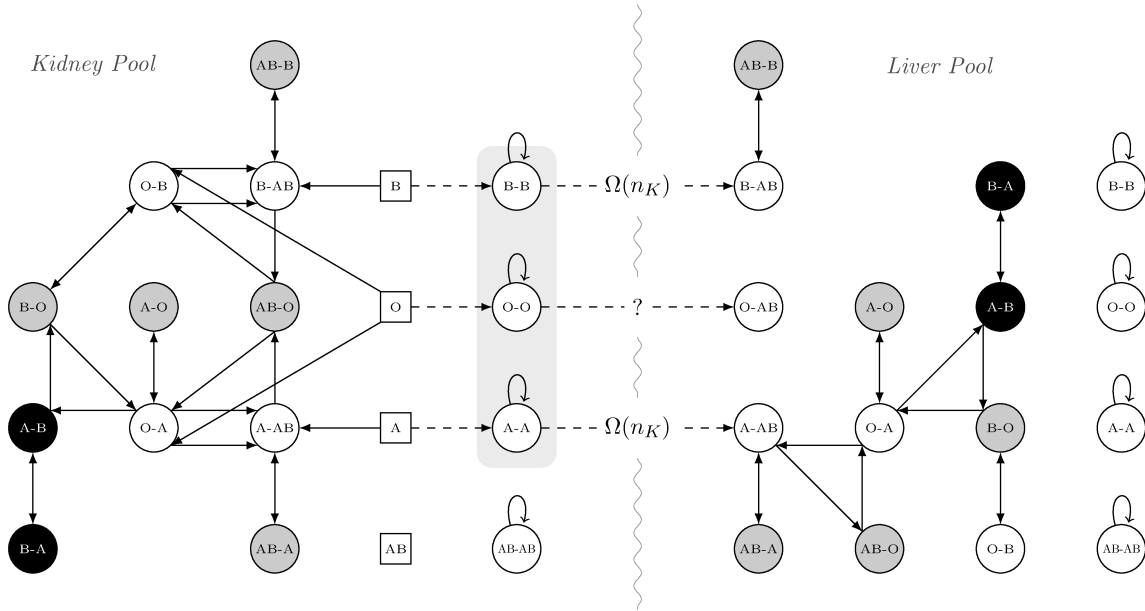
Figure 5: Our constructed matching that directly combines the allocations of Dickerson et al. (2012b) and Ashlagi and Roth (2014)—which it applies initially to the kidney pool and liver pool, respectively—and then threads leftover altruistic kidney donors through the kidney pool into unmatched portions of the liver pool. Altruists are shown as rectangles and candidate-donor pairs as ovals, over-demanded pairs are gray, under-demanded and self-demanded pairs are white, and reciprocal pairs are black. Solid edges represent donations that are in the original allocations, while dashed edges are those added by the allocation we generate for the joint pool.

compatible pair being tissue type incompatible, a constant fraction of the pairs in $V_K^{\text{A-A}}$, specifically $p_{K \to L}|V_K^{\text{A-A}}| = p_{K \to L}\bar{p}\mu_A\mu_A n_K$ pairs in expectation, are willing to give a liver to a liver pair. The use of an A-A kidney pair via an A-altruist-initiated chain results in 0 efficiency loss relative to the initial efficient matching, since there remains a perfect matching in $V_K^{\text{A-A}}$ by well-known results on dense Erdős-Rènyi graphs (see, e.g., Janson et al., 2011). Thus, we gain 1 match for each of the remaining

$$\min\left\{p_{K \to L}|V_K^{\text{A-A}}|, |V_K^{\text{A}}| - |V_K^{\text{A-AB}}|\right\} \tag{5}$$

A-donors. The first input in the minimization in Equation 5 is of size $\Omega(n_K)$, because a constant fraction of a set that is linear in n_K is still linear in n_K. The second input is never negative and is potentially $\Omega(n_K)$ by the theorem statement's assumption on the number of altruistic A-donors or B-donors; we address this uncertainty in the next paragraph.

- **B-altruists:** 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 yields

$$\min \left\{ p_{K \to L} |V_K^{\text{B-B}}|, |V_K^{\text{B}}| - |V_K^{\text{B-AB}}| \right\} \tag{6}$$

  additional matches by threading through willing B-B kidney pairs into the unallocated under-demanded liver pool. By similar logic to the above A-donor case, the first input to the minimization in Equation 6 is again assuredly $\Omega(n_K)$, while the second is never negative and is potentially $\Omega(n_K)$. The theorem statement ensures that at least one of the sets of A-donors or B-donors is large enough—specifically, $|V_K^{\text{A}}| > \mu_{\text{A}} \mu_{\text{AB}} n_K$ or $|V_K^{\text{B}}| > \mu_{\text{B}} \mu_{\text{AB}} n_K$—to trigger a linear gain in at least one of Equations 5 or 6.

- *O-altruists:* In the matching of Dickerson et al. (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). Similarly, because we are not making assumptions on the number of O-donors, if there are so many O-donors in D_K that all under-demanded pairs (e.g., pairs of type O-AB in D_K) are matched, then these O-donors can be threaded directly into the liver pool by way of self-demanded O-O kidney pairs who are willing to give livers (at no efficiency loss, as a perfect matching will remain in $V^{\text{O-O}}$) for a gain of at least 1 under-demanded match in D_L (but this is also not necessary for this proof).

- *Non-altruistic (i.e., paired) vertices:* Self-demanded and reciprocally-demanded pairs cannot help under-demanded pairs without involving over-demanded pairs or altruistic donors. 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 et al. (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 ($|V_K^{\text{A}}| > \mu_{\text{A}} \mu_{\text{AB}} n_K$) or B-donors ($|V_K^{\text{B}}| > \mu_{\text{B}} \mu_{\text{AB}} n_K$) is satisfied by the proposition statement's assumptions, we are guaranteed $\Omega(n_K)$ additional matches by combining both pools by way of Equations 5 and 6. $\qquad \square$

The theoretical results presented in this section motivate the combination of independent kidney and liver exchanges and show that such a joint exchange would allow for the use of altruistic kidney donors at great gain to overall social welfare. Still, both models are significant simplifications of real organ exchange; the push for a fielded liver or multi-organ exchange in reality will require extensive *realistic* simulations showing expected gains in number of matches, among other statistics. We address this in the rest of the paper. Section 4 describes our method for generating and clearing demographically accurate bi-organ compatibility graphs and Section 5 presents experimental results on (i) liver exchange alone and (ii) independent liver and kidney exchanges versus a combined multi-organ exchange.

4. Generating and Clearing Demographically Accurate Pools

In this section, we describe our method for generating realistic organ exchange graphs for programs in steady state, and compare these generated graphs to those produced by the current status quo steady-state kidney exchange generator. We also briefly describe a generator built for early-stage exchanges that have not yet reached steady state; this generator draws from real data from the United Network for Organ Sharing (UNOS) nationwide kidney exchange. We then describe the standard kidney exchange clearing algorithm and, motivated by generated realistic steady-state liver and kidney exchange graphs, present a tweak to this algorithm to decrease liver exchange solution time.

4.1 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.) These graphs will mimic organ exchange in *steady state*; in Section 4.3, we will briefly describe the differences in compatibility graph composition that we have witnessed in the creation of a nascent kidney exchange.

We generate kidney exchange compatibility graphs in accordance with Saidman, Roth, Sönmez, Ünver, and Delmonico (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,[6] (b) the age of the donor and candidate makes a significant difference in transplant success (Egawa, Oike, Buhler, Shapiro, Minamiguchi, Haga, Uryuhara, Kiuchi, Kaihara, & Tanaka, 2004), and (c) the portion of the donor liver that is cut out and transplanted into the candidate must be large enough to keep the candidate alive, while the remainder of the liver in the donor must be large enough to keep the donor alive. A proxy for liver size is the weight of the candidate or donor; intuitively, larger people need larger livers. Thus, we assume a donor must be at least as heavy as his or her matched 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 Report[7]); conditioned on gender, we then draw candidate

---

6. In 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. This is a drastically less important factor in liver transplantation, and is typically not taken into account in liver transplantation in practice or theory.

7. http://www.census.gov/compendia/statab/cats/population.html

blood types from the OPTN (Organ Procurement and Transplantation Network[8]) distribution and donor blood types from the overall US population.[9] 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 released by the Center for Disease Control (McDowell, Fryar, Ogden, & Flegal, 2008). For candidates requiring a kidney, HLA is sampled from the OPTN databases. During edge generation, we include an exogenous "incompatibility 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, Gilchrist, Roth, & Rees, 2011), and is used to account for incompleteness of medical knowledge and temporal fluctuations in candidate-donor compatibility.

Appendix A provides a much more in-depth detailing of the steps we take for data generation, as well as a formal compatibility graph generation algorithm. Next, we compare steady-state liver exchange graphs generated by our algorithm to kidney exchange graphs produced by the standard steady-state generator due to Saidman et al. (2006). Our generator is a generalization of (i.e., more powerful than) that current standard.

4.2 Comparison to Steady-State Kidney Exchange

In empirical kidney exchange research, traits of the family of compatibility graphs used in experiments—like the average in- and out-degree of vertices or number of long paths in the graph—have typically had a large effect on both the performance of clearing engines and the qualitative results obtained (see, e.g., Constantino, Klimentova, Viana, & Rais, 2013; Dickerson et al., 2018, 2014b; Anderson, 2014; Klimentova, Alvelos, & Viana, 2014). With that in mind, we now compare our steady-state generator to the current state of the art kidney exchange generator (Saidman et al., 2006), which was meant to mimic a kidney exchange running in the United States in steady state. 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 6 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 distributions 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 (as exemplified by, e.g., Dickerson et al. (2012b), where the clearing problem with *unbounded* chains was easily solved on sparse real-
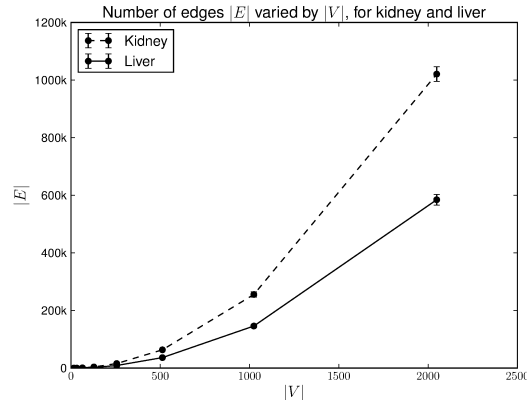

Figure 6: #Edges (in thousands) in generated liver and kidney compatibility graphs (100 graphs per $|V|$). The generated kidney graphs are denser than the liver graphs.

world graphs, but the clearing problem with even *bounded* chains became computationally intractable).

Figure 7 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 incompatibility rate (f) between two otherwise compatible candidates and donors is held constant in each column. We vary $|V|$ and $f \in \{0.0, 0.2, 0.4, 0.8, 0.9\}$. Note that there is no notion of an exogenous incompatibility rate in the kidney graphs (although the %PRA virtual crossmatch simulation is similar to an exogenous incompatibility 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 7, 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. 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.
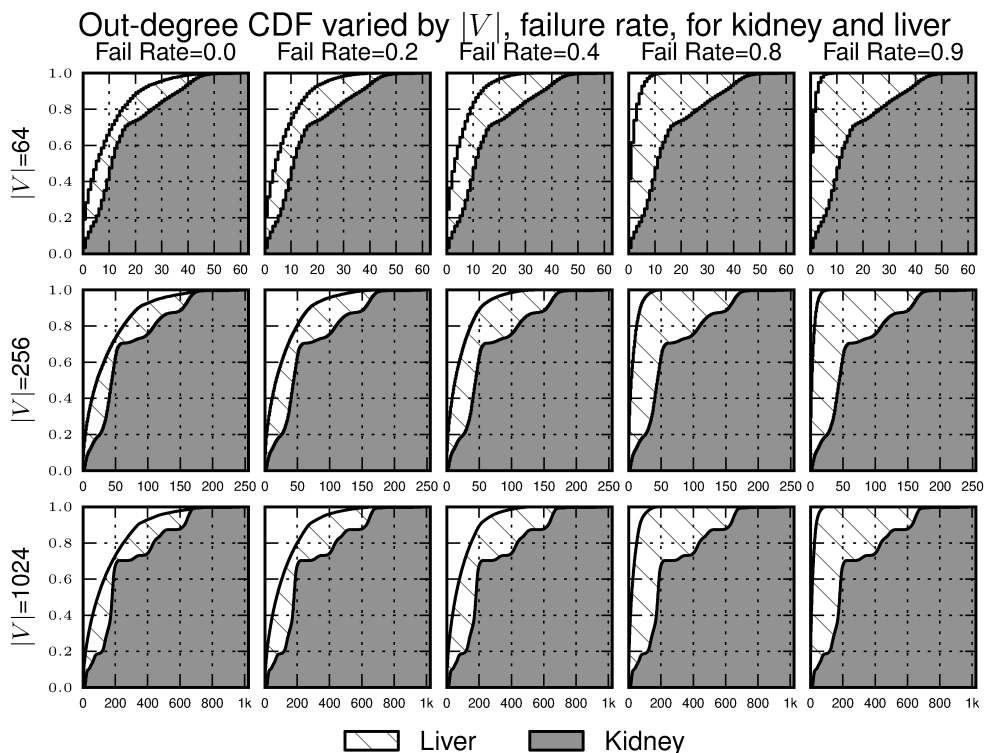
MULTI-ORGAN EXCHANGE

Figure 7: Cumulative distribution functions of the out-degree of vertices as we increase $|V|$ (varies per row) and exogenous incompatibility 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 incompatibility rate increases, as well as the three qualitative sections in the kidney graphs due to the three different %PRA classes.

4.3 Sparse Generated Compatibility Graphs

While the transplantation of each organ is unique in its own way, we can draw from the experience of nascent kidney exchanges in the US and abroad when considering the makeup of a hypothetical liver or multi-organ exchange. As discussed in Section 3.2, early theoretical models of kidney exchange ended up behaving substantially differently than fielded exchanges; indeed, as practioners uncovered logistical constraints and medical features of kidney exchange, economic and computational models adapted to better reflect reality.

The data generation process described in Section 4.1 produces a "best guess" at what a steady-state liver or multi-organ exchange would look like. That generation process can easily be adapted to unforeseen features of those exchanges as they arise. Indeed, we address the unforeseen in a general way using the exogeneous incompatibility rate $f \in [0, 1]$. This incompatibility rate affects each pair independently. In reality, some pairs may be much harder to match than others, sometimes for poorly-understood reasons from a medical point

659

of view. This is the case in kidney exchange (see discussions in, e.g., Ashlagi et al., 2012, 2013; Dickerson et al., 2018).

With this in mind, in this paper we also perform experiments directly on compatibility graphs drawn from the United Network for Organ Sharing (UNOS) nationwide kidney exchange, which is a large, fielded real-world kidney exchange that currently includes 143 transplant centers in the US. In other kidney-exchange-specific work, the authors built a compatibility graph generator that accurately mimics the UNOS nationwide exchange (Dickerson, Procaccia, & Sandholm, 2014a). In the present work, we seed this generator with the first 192 match runs (October 2010 through March 2015) of the UNOS exchange and feed those graphs into our static and dynamic organ exchange simulators. To simulate multiple organs, we mark pairs as needing either a kidney or liver using demographic information from the most recent OPTN reports on waiting lists for kidneys and livers, respectively, and attach edges from real-world altruists in the UNOS pool only to those pairs marked as needing kidneys.

Obviously, we would not expect compatibility graphs generated in this manner to adhere at a fine-grained level to those of a fielded multi-organ exchange; indeed, these generated graphs contain no notion of different organs (beside kidneys) beyond a simple coin flip. That said, using these compatibility graphs allows us to remove the dependence on the exogeneous incompatibility factor f; indeed, it has already been taken into account by the real world in these graphs! Thus, by including experimental results on this second distribution of graphs, we hope to show *qualitatively* if not quantitatively that gains from multi-organ exchange still hold under a more intricate notion of exogeneous incompatibility rates—because certainly a fielded multi-organ exchange will encounter unpredictable constraints at run time, as was the case in the nascency of kidney exchange.

In the experimental results of Section 5, we will refer to those dense compatibility graphs (sparsified in a parameterized way using f and $p_{K \to L}$) generated in accordance with Section 4.1 as DENSE, and those exogeneously sparse compatibility graphs (sparsified further only by $p_{K \to L}$) drawn from real data as described in this section as UNOS.

4.4 The Clearing Algorithm

We now briefly discuss a scalable optimal kidney exchange clearing algorithm. An extended version of the algorithm due to Abraham et al. (2007), built by us, is used in the United Network for Organ Sharing (UNOS) 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 match cardinality.

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, Johnson, Nemhauser, Savelsbergh, & Vance, 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. Our solver that is used by UNOS—as

well as in this paper—uses several additional techniques to make kidney exchange clearing scalable for memory and time (Abraham et al., 2007). It uses empirically and theoretically motivated heuristics to seed the initial cycle (i.e., decision 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.

4.4.1 A Liver-Specific Cycle Seeding Heuristic

The selection of the initial seed columns—representing individual cycles or chains—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 negligible compared to the NP-hard clearing problem.

5. 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 expire while 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 would be by running two separate nationwide liver and kidney exchanges.

5.1 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 8 shows timing results on liver

exchange graphs of various sizes $|V|$ and exogenous incompatibility rates f drawn from the DENSE distribution. Intuitively, when f is low (or zero), the optimizer must consider many more edges than when f is high, resulting in longer run times for denser graphs. As expected, the computation time increases drastically with graph size—although our solver is still able to solve large problems to optimality.
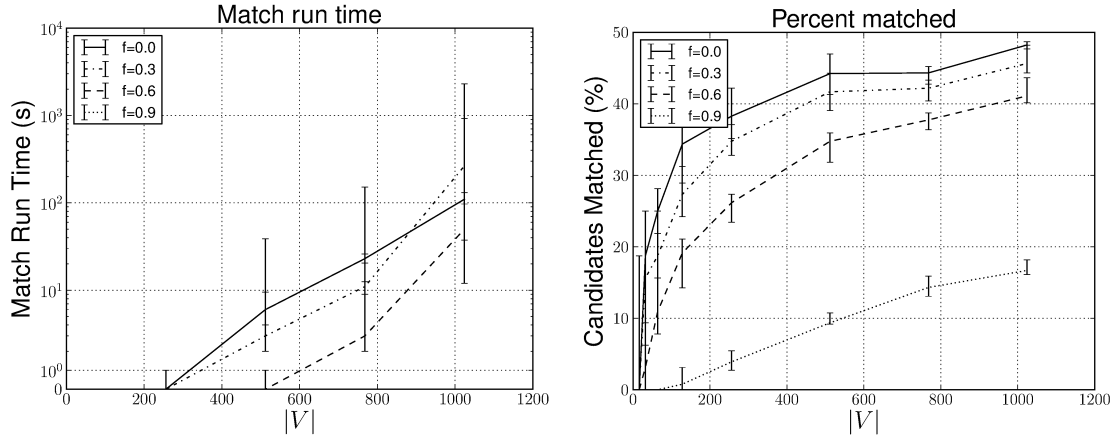

Figure 8: Median match run time (left) and median percentage of candidates matched (right), varying incompatibility rate f and graph size $|V|$, with first and third quartile error bars, for DENSE compatibility graphs.

Figure 8 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 incompatibility rate f. Intuitively, when f is held low, the percentage of candidates matched is higher than when the incompatibility 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 (nationwide or international) liver exchange.

5.1.1 ADDRESSING THE NEEDS OF SOCIETY

The estimated steady-state size of the nationwide kidney exchange is 10,000 candidate-donor pairs (Abraham et al., 2007). The rate of live liver donation is 1/8th of the rate of live kidney donation—5% of all liver transplants in the US 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 as kidney donors in steady-state. With 101,257 candidates currently waiting for a kidney and 15,268 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 15,268 / 101,257 ≈ 7.5% of 10,000, or roughly 750 candidates. So, our clearing algorithm should be able to handle batch runs of a nationwide liver exchange.

5.2 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 UNOS nationwide kidney exchange and others currently operate 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 & Sandholm, 2009; Ünver, 2010; Dickerson, Procaccia, & Sandholm, 2012a; Ashlagi et al., 2013; Anshelevich, Chhabra, Das, & Gerrior, 2013; Anderson, Ashlagi, Gamarnik, & Kanoria, 2015a; Akbarpour, Li, & Gharan, 2014; Dickerson & Sandholm, 2015). Regardless of the optimization method used, organ exchange is inherently dynamic, with candidates and donors arriving and departing over time; we work in such a setting here.

We start with a pool of $|V| = 400$ pairs assumed to contain highly-sensitized patients who built up in the system over time. These are matched myopically at each time period, including the final time period. Given a matched cycle by the algorithm, we then simulate that transplant actually succeeding in real life via an exogenous parameter set to 0.7. This post-match, pre-transplant failure probability is drawn from real data, as motivated by Dickerson et al. (2018). If any edge in a cycle fails, that entire cycle fails, and all candidates are returned to the pool (with the failed edge or edges removed). We simulate candidates leaving the pool (either through finding a transplant or dying). While 12% of patients in need of a kidney will be alive after 10 years via dialysis while waiting for a kidney (HHS/HRSA/HSB/DOT, 2011), no such treatment exists for livers; thus, life expectancy drops to 1–2 years, which we simulate. In expectation $|V_{new}| = 233$ new candidates arrive in the pool per month, and the algorithm continues. We test over 24 months.
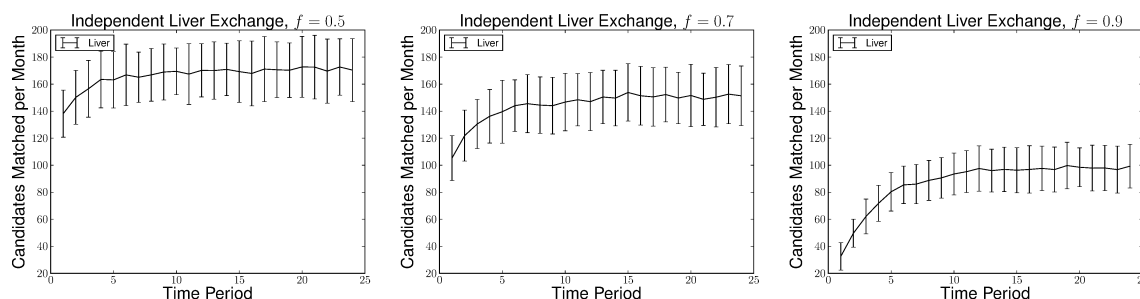

Figure 9: Number of candidates matched per time period in a dynamic setting over $T = 24$ months, for exogeneous incompatibility rates $f \in \{0.5, 0.7, 0.9\}$, for compatibility graphs drawn from the DENSE distribution.

Figure 9 shows the number of candidates matched by the clearing engine at each time period for dynamic graphs drawn from the DENSE distribution with exogeneous incompatibility rate $f \in \{0.5, 0.7, 0.9\}$. Shown in the figure is the number of candidates matched *by the algorithm*, but before the virtual post-match failures are taken into account. Initially, there is a period of a few months during which the dynamic pool builds density as more candidate-donor pairs enter, followed by a relatively constant steady state. Intuitively, those simulated graphs with lower incompatibility rates f result in a larger number of matches

per time period and overall, which qualitatively aligns with the static results of Section 5.1. Clearing times ranged between two and three minutes per time period.

5.3 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). Ongoing work by Ergin, Sönmez, and Ünver (2014) is attempting to exchange *lungs*, another related but different organ exchange problem (e.g., typically two donors are required per candidate); it is likely that the first such hand-organized exchange will take place in Japan. 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 Propositions 1, 2, and 3.

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 OPTN waitlist data. We mimic the experiments in the previous section, with a starting pool size of $|V| = 400$ 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 post-match failure rate (0.7) as in the previous section, and simulate candidate-donor pairs entering and exiting the pool in a similar fashion. For DENSE experiments, 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. For UNOS experiments, we draw from the UNOS generator as described in Section 4.3. We test over 24 months.

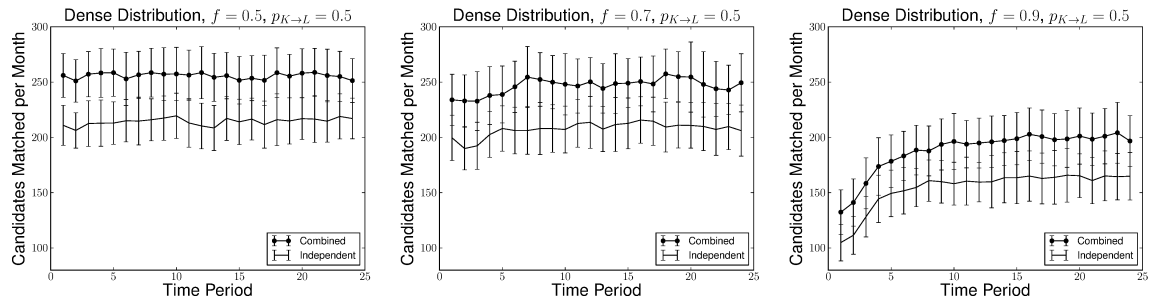Figure 10: Number of matches per time period in independent liver and kidney exchanges and a combined multi-organ exchange in a dynamic setting over $T = 24$ months, for graphs drawn from the DENSE distribution with $f \in \{0.5, 0.7, 0.9\}$ and $p_{K \to L} = 0.5$

Figure 10 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. We set $p_{K \to L} = 0.5$, but relax this assumption later. Clearly evident is the loss of life resulting from keeping both the liver and kidney pools independent, with the bi-organ exchange matching roughly 40 more candidates per month, depending on exogeneous incompatibility rate f, 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 the experiments of Figure 10 with $p_{K \to L} = 0.5$, the combined bi-organ pool produced roughly 20% more matches than the sum of the two independent organ pools—specifically, 19.3%, 18.8%, and 21.8% for each of $f = 0.5$, $f = 0.7$, and $f = 0.9$, respectively. Independent samples t-tests 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 ($t(73) = 44.141$, $t(42) = 38.872$, and $t(81) = 41.651$ for each of $f = 0.5$, $f = 0.7$, and $f = 0.9$, respectively, with two-tailed $p \ll 0.0001$ for each). Qualitatively, this behavior is repeated for other values of $p_{K \to L} \in (0, 1.0]$, which we explore next.
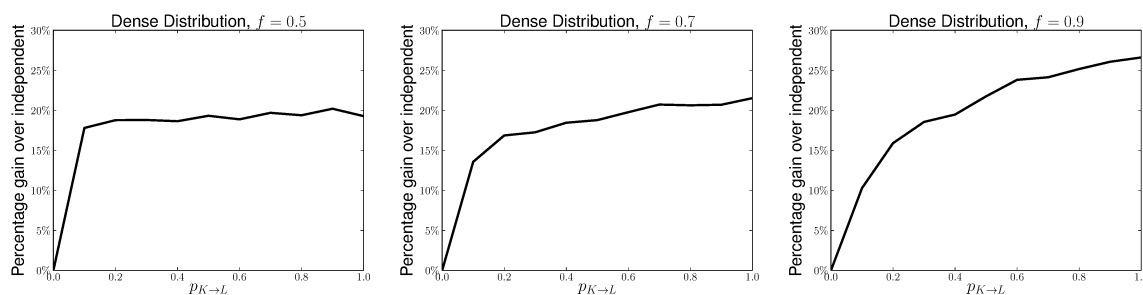

Figure 11: Percentage gain in number of matches over two independent exchanges for a combined exchange as $p_{K \to L}$ increases, for generated graphs from the DENSE distribution and $f \in \{0.5, 0.7, 0.9\}$.

Figure 11 shows the percentage gain in total number of matches of a combined exchange over two independent exchanges as $p_{K \to L}$, the probability of a kidney-paired donor being willing to donate a liver lobe, increases. As expected, higher rates of $p_{K \to L}$ result in better aggregate matching performance due to chains triggered by kidney-yielding altruists having more options to thread into the liver pool, and as longer (i.e., more valuable) chains. The effect of the exogeneous incompatibility rate f on the immediacy of this increase in match performance is noticeable; the relatively dense $f = 0.5$ DENSE graphs appear to maximize the use of chains for very low values of $p_{K \to L} = 0.1$, while the sparser compatibility graphs see an increase across the entire range of values for $p_{K \to L} > 0$.

Finally, Figure 12 shows similar results on dynamic graphs taken from the UNOS distribution of graphs. While the absolute efficiency gains are less than in the DENSE experiments, gains of up to 7.6% are witnessed. This lower gain may be due in part to other intricacies specific to the UNOS generator, which mimics some of the operational constraints of the fielded UNOS exchange. For example, pairs have the ability to specify a maximum cycle
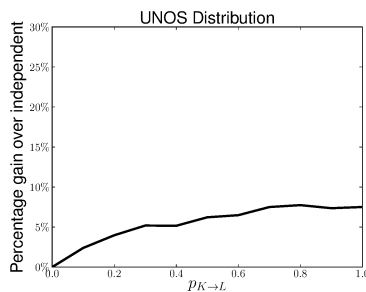
Figure 12: Percentage gain in number of matches over two independent exchanges for a combined exchange as $p_{K \to L}$ increases, for generated graphs from the UNOS distribution.

or chain length in which they would be willing to participate; indeed, some pairs express a preference *not* to participate in chains at all, which we honor in these simulations. If pairs in the liver pool choose not to participate in chains, then the gains seen from combining pools will be lower than those that would be seen if all pairs chose to participate in all potential matching structures. Donors also are associated with specific transplant centers and can specify maximum travel distance or a list of (un)acceptable transplant centers where a donation can take place; many donors in the UNOS pool take advantage of this expressiveness, which further constrains the set of possible matches in graphs drawn from the UNOS distribution.

Additional experimental results and statistical significance testing for the data in Figures 10, 11 and 12 are presented in Appendix B.

The curves in Figures 11 and 12 are in part a function of the number of altruistic donors available in the pool. A greater number of altruists would result in lower necessary values of $p_{K \to L}$. Participation rates of altruists in kidney exchange are still in flux, the value of $p_{K \to L}$ is not yet known, and the level and type of sparsity of a real-world multi-organ exchange cannot truly be determined until one has been fielded; that said, the results of Figure 11 and 12 conclusively support the gain in number of matches from combining exchanges across a large variety of values for these unknown parameters.

6. 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, presently fielded kidney exchange. On demographically accurate data, vetted kidney exchange clearing algorithms (with a small change) 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. We 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. Recent and ongoing work by Ergin et al. (2014) is attempting to do this for a *lung exchange*, another related but different organ exchange problem (e.g., typically two donors are required per candidate), with the first trial likely to occur in Japan. Recent work by Luo and Tang (2015) approaches lung exchange from a game-theoretic point of view. Another direction is to develop scalable computational methods for the dynamic problem; even for kidneys, the best current techniques are for simplified models (Ünver, 2010; Ashlagi et al., 2013; Anshelevich et al., 2013; Anderson et al., 2015a; Akbarpour et al., 2014) or face computational challenges (Awasthi & Sandholm, 2009; Dickerson et al., 2012a; Dickerson & Sandholm, 2015). Indeed, due to the shorter expected maximum waiting time for candidates in need of a liver, such dynamic models may result in greater relative gains in number of matches compared to static models.

Even for the static problem, scalability problems tend to get worse with the inclusion of donation chains started by altruistic donors. The cycle 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 et al., 2012a, 2012b; Gentry & Segev, 2011; Gentry, Montgomery, Swihart, & Segev, 2009; Anderson, 2014; Glorie, van de Klundert, & Wagelmans, 2014; Ding, Ge, He, & Ryan, 2015; Dickerson et al., 2016). 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, Daller, Aeder, Shapiro, Sandholm, Casingal, Goldfarb, Lewis, Goebel, & Siegler, 2010). However, that remains to be seen. In any case, one could include chains started by kidney-donating altruists into a bi-organ exchange—as we do in this paper—if the scalability and ethical challenges of such multi-organ 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, Gupta, Procaccia, & Sharma, 2013; Anderson, 2014; Anderson, Ashlagi, Gamarnik, & Roth, 2015b; Blum, Dickerson, Haghtalab, Procaccia, Sandholm, & Sharma, 2015; Dickerson et al., 2016, 2018) to both theoretically and empirically maximize the expected number of actual transplants (possibly with respect to some fairness constraints (Dickerson et al., 2014b; Li, Liu, Huang, & Tang, 2014; Sönmez & Ünver, 2015; McElfresh & Dickerson, 2018) that could try to balance factors including the increased risk of liver versus kidney donation) stemming from an algorithmic match. Recent work by Glorie (2012) is an initial foray into learning a better estimate of the probability of a transplant failure between a patient and a donor, but much is left to be done.

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.

Acknowledgments

Appendix A. A Parameterized, Realistic Compatibility Graph Generator

In this section, we provide a more in-depth enumeration of the steps taken to generate realistic liver or multi-organ exchange compatibility graphs; a shorter explanation was given in the main paper in Section 4.1. Section A.1 describes the process of drawing data from reliable sources (here, specific to the US), while Section A.2 shows how we feed this generated data 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. As noted in the main paper, in the large and 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., gender, age, weight, blood type distributions) into the generator algorithm would provide realistic generation of non-US compatibility graphs.) Our generator is a generalization of (i.e., more powerful than) the current standard generator proposed by Saidman et al. (2006).

A.1 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.[10] All OPTN data is current as of November 11, 2011.

---

10. http://optn.transplant.hrsa.gov/data/

A.1.1 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.[11] Table 1 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

Table 1: Distribution of (liver) candidate and donor genders, drawn from OPTN and 2010 US Census data, respectively.

A.1.2 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.[12] 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. Table 2 shows the exact distribution and the ABO-compatibility matrix, with percentages shown for liver-needing candidates.

A.1.3 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, Table 3 does not separate the population into one-year segments as rows, while our generator does. In our generator we also take

---

11. http://www.census.gov/compendia/statab/cats/population.html
12. http://bloodcenter.stanford.edu/about_blood/blood_types.html

Donor	Candidate			
ABO	O	A	B	AB
O	✓	✓	✓	✓
A	✗	✓	✗	✓
B	✗	✗	✓	✓
AB	✗	✗	✗	✓

	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

Table 2: Left: ABO blood type compatibility matrix. Marks indicate a donor (row) as ABO-compatible with a candidate (column). Right: ABO percentages for candidates and donors.

into account the constraint that organ donors must be 18 years old, and we normalize the distributions accordingly.

	Male		Female	
Age	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

Table 3: Probability distribution of ages, respective of candidate and donor gender.

A.1.4 WEIGHT

Unlike in kidney exchange, the physical weight of both the candidate and donor play an enormous role in the feasibility of liver transplantation.[13] 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 (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

---

13. 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.

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.

A.1.5 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 databases[14]), we exclude %PRA in our liver experiments. However, %PRA is included in our multi-organ experiments for kidney candidates.

A.2 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 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.

---

14. The relationship (e.g., parent-child, spousal) between candidate and donor can yield information on HLA compatibility (e.g., due to inheritance of HLA from each parent or changes in HLA antibodies due to pregnancy), and is supported by the generator of Saidman et al. (2006) and our generator.

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 = candidate, d = donor
        $c.drawOrganType()$
        $\{c, d\}.drawGender()$
        $\{c, d\}.drawBlood(gender)$
        $\{c, d\}.drawAge(gender)$
        $\{c, d\}.drawTissueType(gender)$
        $\{c, d\}.drawWeight(gender, age)$
        if $\neg isCompatible(c, d)$ **then**
            $V = V \cup \{v^{c,d}\}$

    for $v_i, v_j \in V$ s.t. $V_i \neq V_j$ **do**
        if $isCompatible(v_j^c, v_i^d)$ and $x \in U[0, 1] > f$ **then**
            if $isWilling(v_i, v_j)$ **then**
                $E = E \cup \{(v_i, v_j)\}$

    return *directed compatibility graph G*

When we ran the liver and multi-organ experiments in Section 5, the kidney waitlist was 5.84 times longer than the liver waitlist, which was reflected in this algorithm. (When this paper was submitted, the kidney waitlist was 6.50 times longer than the liver waitlist.)

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. Recent kidney exchange research suggests that incentivizing even compatible pairs to join a nationwide exchange could result in better matchings (Rees et al., 2009; Ashlagi & Roth, 2014). 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).

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 "incompatibility 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

$f = 0.5$		Independent		Combined				t-test			Mann-Whitney	
$p_{K \to L}$	n	Avg. #	Stdev	Avg. #	Stdev	% Gain		t	p		U	p
0.0	90	5143.7	(129.9)	–	–	–		–	–			
0.1	41	–	–	6059.4	(155.0)	17.80%		34.884	≪ 0.001		0.0	≪ 0.001
0.2	53	–	–	6109.3	(153.3)	18.77%		39.825	≪ 0.001		0.0	≪ 0.001
0.3	63	–	–	6110.4	(149.0)	18.79%		42.332	≪ 0.001		0.0	≪ 0.001
0.4	79	–	–	6102.4	(143.9)	18.64%		45.240	≪ 0.001		0.0	≪ 0.001
0.5	73	–	–	6137.5	(155.7)	19.32%		44.141	≪ 0.001		0.0	≪ 0.001
0.6	83	–	–	6114.5	(126.2)	18.87%		49.491	≪ 0.001		0.0	≪ 0.001
0.7	81	–	–	6156.4	(153.2)	19.69%		46.472	≪ 0.001		0.0	≪ 0.001
0.8	77	–	–	6140.8	(140.3)	19.38%		47.364	≪ 0.001		0.0	≪ 0.001
0.9	79	–	–	6182.7	(143.7)	20.20%		49.060	≪ 0.001		0.0	≪ 0.001
1.0	81	–	–	6135.3	(133.2)	19.28%		48.953	≪ 0.001		0.0	≪ 0.001

Table 4: Statistical significance testing for DENSE distribution graphs with $f = 0.5$.

new compatibility aspects into account. The algorithm also calls a function *isWilling(v_i, v_j)*, which returns true if the donor at v_i is willing to give an organ of the type needed by the patient in v_j. This corresponds to, e.g., the probabilities $p_{K \to L}$ and $p_{L \to K}$ used in this paper's theoretical and experimental sections.

Appendix B. Additional Experimental Results

In this section, we provide statistical significance testing for the dynamic bi-organ experiments of Section 5.3. The tables are organized as follows. Each table corresponds to a different distribution of compatibility graphs. Tables 4, 5, and 6 give results for DENSE graphs with exogeneous incompatibility rates $f = 0.5$, $f = 0.7$, and $f = 0.9$, respectively. These tables support Figures 10 and 11 in the body of the paper. Table 7 gives results for the UNOS family of graphs. This table supports Figure 12 in the body of the paper.

Each row in a table corresponds to a different value of $p_{K \to L}$; the value of $p_{K \to L}$ is specified in the first column of the table. From left to right, the columns represent: n, the number of independent runs used to support the results in this row; the average number of patients matched in total for an independent liver and independent kidney exchange; the standard deviation of the previous; the average number of patients matched in total for a combined bi-organ exchange; the standard deviation of the previous; the percentage gain in number of matched patients achieved by combining exchanges; t-statistic from an independent samples t-test; the associated two-tailed p-value for the previous; U-statistic from a Mann-Whitney U test (roughly, a non-parametric version of the independent samples t-test); and the associated one-tailed p-value for the previous.

$f = 0.7$		Independent		Combined				t-test		Mann-Whitney	
$p_{K \to L}$	n	Avg. #	Stdev	Avg. #	Stdev	% Gain		t	p	U	p
0.0	99	4979.6	(127.6)	–	–	–		–	–		
0.1	90	–	–	5655.3	(146.6)	13.57%		33.685	$\ll 0.001$	0.0	$\ll 0.001$
0.2	69	–	–	5819.3	(116.5)	16.86%		43.215	$\ll 0.001$	0.0	$\ll 0.001$
0.3	50	–	–	5838.9	(134.3)	17.25%		37.874	$\ll 0.001$	0.0	$\ll 0.001$
0.4	35	–	–	5898.4	(143.0)	18.45%		35.188	$\ll 0.001$	0.0	$\ll 0.001$
0.5	42	–	–	5914.6	(134.5)	18.78%		38.872	$\ll 0.001$	0.0	$\ll 0.001$
0.6	32	–	–	5964.0	(119.9)	19.77%		38.196	$\ll 0.001$	0.0	$\ll 0.001$
0.7	33	–	–	6011.4	(175.6)	20.72%		36.094	$\ll 0.001$	0.0	$\ll 0.001$
0.8	36	–	–	6006.9	(134.4)	20.63%		40.476	$\ll 0.001$	0.0	$\ll 0.001$
0.9	40	–	–	6010.4	(152.2)	20.70%		40.419	$\ll 0.001$	0.0	$\ll 0.001$
1.0	36	–	–	6051.4	(156.8)	21.52%		40.192	$\ll 0.001$	0.0	$\ll 0.001$

Table 5: Statistical significance testing for DENSE distribution graphs with $f = 0.7$.

$f = 0.9$		Independent		Combined				t-test		Mann-Whitney	
$p_{K \to L}$	n	Avg. #	Stdev	Avg. #	Stdev	% Gain		t	p	U	p
0.0	79	3708.4	(113.2)	–	–	–		–	–		
0.1	79	–	–	4089.6	(113.3)	10.28%		21.018	$\ll 0.001$	19.5	$\ll 0.001$
0.2	81	–	–	4298.0	(108.0)	15.90%		33.501	$\ll 0.001$	0.0	$\ll 0.001$
0.3	82	–	–	4396.4	(121.8)	18.55%		36.866	$\ll 0.001$	0.0	$\ll 0.001$
0.4	81	–	–	4430.6	(133.8)	19.47%		36.577	$\ll 0.001$	0.0	$\ll 0.001$
0.5	81	–	–	4514.9	(129.4)	21.75%		41.651	$\ll 0.001$	0.0	$\ll 0.001$
0.6	81	–	–	4591.5	(139.0)	23.81%		43.721	$\ll 0.001$	0.0	$\ll 0.001$
0.7	78	–	–	4603.2	(133.8)	24.13%		44.977	$\ll 0.001$	0.0	$\ll 0.001$
0.8	78	–	–	4641.6	(153.0)	25.16%		43.200	$\ll 0.001$	0.0	$\ll 0.001$
0.9	78	–	–	4675.2	(112.6)	26.07%		53.307	$\ll 0.001$	0.0	$\ll 0.001$
1.0	79	–	–	4695.1	(121.2)	26.61%		52.553	$\ll 0.001$	0.0	$\ll 0.001$

Table 6: Statistical significance testing for DENSE distribution graphs with $f = 0.9$.

UNOS $p_{K \to L}$	Independent			Combined			t-test		Mann-Whitney	
	n	Avg. #	Stdev	Avg. #	Stdev	% Gain	t	p	U	p
0.0	82	4003.8	(108.3)	–	–	–	–	–		
0.1	86	–	–	4099.5	(108.3)	2.39%	5.689	$\ll 0.001$	1959.5	$\ll 0.001$
0.2	84	–	–	4162.9	(124.6)	3.97%	8.719	$\ll 0.001$	1154.5	$\ll 0.001$
0.3	80	–	–	4211.5	(119.3)	5.19%	11.538	$\ll 0.001$	617.5	$\ll 0.001$
0.4	79	–	–	4210.4	(109.2)	5.16%	11.978	$\ll 0.001$	561.0	$\ll 0.001$
0.5	74	–	–	4252.8	(103.6)	6.22%	14.539	$\ll 0.001$	306.5	$\ll 0.001$
0.6	76	–	–	4263.4	(115.3)	6.48%	14.501	$\ll 0.001$	303.0	$\ll 0.001$
0.7	76	–	–	4304.1	(112.8)	7.50%	16.961	$\ll 0.001$	185.5	$\ll 0.001$
0.8	62	–	–	4313.6	(124.6)	7.74%	15.813	$\ll 0.001$	143.0	$\ll 0.001$
0.9	67	–	–	4298.3	(121.2)	7.36%	15.544	$\ll 0.001$	181.0	$\ll 0.001$
1.0	68	–	–	4304.3	(120.7)	7.51%	15.949	$\ll 0.001$	163.0	$\ll 0.001$

Table 7: Statistical significance testing for UNOS distribution graphs.

References

Abraham, D., Blum, A., & Sandholm, T. (2007). Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pp. 295–304.

Akbarpour, M., Li, S., & Gharan, S. O. (2014). Dynamic matching market design. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, p. 355.

Anderson, R. (2014). *Stochastic models and data driven simulations for healthcare operations*. Ph.D. thesis, Massachusetts Institute of Technology.

Anderson, R., Ashlagi, I., Gamarnik, D., & Kanoria, Y. (2015a). A dynamic model of barter exchange. In *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1925–1933.

Anderson, R., Ashlagi, I., Gamarnik, D., & Roth, A. E. (2015b). Finding long chains in kidney exchange using the traveling salesman problem. *Proceedings of the National Academy of Sciences*, *112*(3), 663–668.

Anshelevich, E., Chhabra, M., Das, S., & Gerrior, M. (2013). On the social welfare of mechanisms for repeated batch matching. In *AAAI Conference on Artificial Intelligence (AAAI)*, pp. 60–66.

Ashlagi, I., Fischer, F., Kash, I. A., & Procaccia, A. D. (2015). Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, *91*, 284–296.

Ashlagi, I., Gamarnik, D., Rees, M., & Roth, A. E. (2012). The need for (long) chains in kidney exchange. NBER Working Paper No. 18202.

Ashlagi, I., Gilchrist, D. S., Roth, A. E., & Rees, M. (2011). Nonsimultaneous chains and dominos in kidney-paired donation—revisited. *American Journal of Transplantation*, *11*(5), 984–994.

Ashlagi, I., Jaillet, P., & Manshadi, V. H. (2013). Kidney exchange in dynamic sparse heterogenous pools. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pp. 25–26.

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

Awasthi, P., & 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)*, pp. 405–411.

Barnhart, C., Johnson, E., Nemhauser, G., Savelsbergh, M., & Vance, P. (1998). Branch-and-price: column generation for solving huge integer programs. *Operations Research*, *46*, 316–329.

Blum, A., Dickerson, J. P., Haghtalab, N., Procaccia, A. D., Sandholm, T., & Sharma, A. (2015). Ignorance is almost bliss: Near-optimal stochastic matching with few queries. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, pp. 325–342.

Blum, A., Gupta, A., Procaccia, A. D., & Sharma, A. (2013). Harnessing the power of two crossmatches. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pp. 123–140.

Brown, R. S. (2008). Live donors in liver transplantation. *Gastroenterology*, *134*(6), 1802–1813.

Caragiannis, I., Filos-Ratsikas, A., & Procaccia, A. D. (2015). An improved 2-agent kidney exchange mechanism. *Theoretical Computer Science*, *589*, 53–60.

Chan, S. C., Lo, C. M., Yong, B. H., Tsui, W. J., Ng, K. K., & Fan, S. T. (2010). Paired donor interchange to avoid ABO-incompatible living donor liver transplantation. *Liver Transplantation*, *16*(4), 478–481.

Cheah, Y. L., Simpson, M. A., Pomposelli, J. J., & 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.

Constantino, M., Klimentova, X., Viana, A., & Rais, A. (2013). New insights on integer-programming models for the kidney exchange problem. *European Journal of Operational Research*, *231*(1), 57–68.

Dickerson, J. P., Manlove, D., Plaut, B., Sandholm, T., & Trimble, J. (2016). Position-indexed formulations for kidney exchange. In *Proceedings of the ACM Conference on Economics and Computation (EC)*.

Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2012a). Dynamic matching via weighted myopia with application to kidney exchange. In *AAAI Conference on Artificial Intelligence (AAAI)*, pp. 1340–1346.

Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2012b). Optimizing kidney exchange with transplant chains: Theory and reality. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 711–718.

Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2014a). Empirical price of fairness in failure-aware kidney exchange. In *Towards Better and more Affordable Healthcare: Incentives, Game Theory, and Artificial Intelligence (HCAGT) workshop at AAMAS-2014*.

Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2014b). Price of fairness in kidney exchange. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 1013–1020.

Dickerson, J. P., Procaccia, A. D., & Sandholm, T. (2018). Failure-aware kidney exchange. *Management Science*. To appear. Earlier version appeared in ACM EC-13.

Dickerson, J. P., & Sandholm, T. (2015). FutureMatch: Combining human value judgments and machine learning to match in dynamic environments. In *AAAI Conference on Artificial Intelligence (AAAI)*, pp. 622–628.

Ding, Y., Ge, D., He, S., & Ryan, C. (2015). A non-asymptotic approach to analyzing kidney exchange graphs. In *Proceedings of the ACM Conference on Economics and Computation (EC)*, pp. 257–258.

Egawa, H., Oike, F., Buhler, L., Shapiro, A., Minamiguchi, S., Haga, H., Uryuhara, K., Kiuchi, T., Kaihara, S., & Tanaka, K. (2004). Impact of recipient age on outcome of ABO-incompatible living-donor liver transplantation. *Transplantation, 77*(3), 403.

Ergin, H., Sönmez, T., & Ünver, M. U. (2014). Lung exchange. Working paper.

Gentry, S. E., Montgomery, R. A., Swihart, B. J., & Segev, D. L. (2009). The roles of dominos and nonsimultaneous chains in kidney paired donation. *American Journal of Transplantation, 9*(6), 1330–1336.

Gentry, S. E., & Segev, D. L. (2011). The honeymoon phase and studies of nonsimultaneous chains in kidney-paired donation. *American Journal of Transplantation, 11*(12), 2778–2779.

Glorie, K. M. (2012). Estimating the probability of positive crossmatch after negative virtual crossmatch. Tech. rep., Erasmus School of Economics.

Glorie, K. M., van de Klundert, J. J., & Wagelmans, A. P. M. (2014). Kidney exchange with long chains: An efficient pricing algorithm for clearing barter exchanges with branch-and-price. *Manufacturing & Service Operations Management (MSOM), 16*(4), 498–512.

HHS/HRSA/HSB/DOT (2011). OPTN/SRTR annual data report.

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., & 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., & Rucinski, A. (2011). *Random Graphs*. Wiley Series in Discrete Mathematics and Optimization. Wiley.

Klimentova, X., Alvelos, F., & Viana, A. (2014). A new branch-and-price approach for the kidney exchange problem. In *Computational Science and Its Applications (ICCSA-2014)*, pp. 237–252. Springer.

Krivelevich, M., Lubetzky, E., & Sudakov, B. (2013). Longest cycles in sparse random digraphs. *Random Struct. Algorithms*, *43*(1), 1–15.

Li, J., Liu, Y., Huang, L., & Tang, P. (2014). Egalitarian pairwise kidney exchange: Fast algorithms via linear programming and parametric flow. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pp. 445–452.

Li, Y., Kalbfleisch, J., Song, P. X., Zhou, Y., Leichtman, A., & Rees, M. (2011). Optimization and simulation of an evolving kidney paired donation (KPD) program. Department of biostatistics working paper 90, University of Michigan.

Luo, S., & Tang, P. (2015). Mechanism design and implementation for lung exchange. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 209–215.

McDowell, M., Fryar, C., Ogden, C., & Flegal, K. (2008). Anthropometric reference data for children and adults: United States, 2003–2006. *Nutrition*, *10*(10), 1–45.

McElfresh, D., & Dickerson, J. P. (2018). Balancing lexicographic fairness and a utilitarian objective with application to kidney exchange. In *AAAI Conference on Artificial Intelligence (AAAI)*.

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

Rees, M., Kopke, J., Pelletier, R., Segev, D., Rutter, M., Fabrega, A., Rogers, J., Pankewycz, O., Hiller, J., Roth, A., Sandholm, T., Ünver, U., & Montgomery, R. (2009). A nonsimultaneous, extended, altruistic-donor chain. *New England Journal of Medicine*, *360*(11), 1096–1101.

Roth, A., Sönmez, T., & Ünver, U. (2004). Kidney exchange. *Quarterly Journal of Economics*, *119*(2), 457–488.

Roth, A., Sönmez, T., & Ünver, U. (2005a). A kidney exchange clearinghouse in New England. *American Economic Review*, *95*(2), 376–380.

Roth, A., Sönmez, T., & Ünver, U. (2005b). Pairwise kidney exchange. *Journal of Economic Theory*, *125*(2), 151–188.

Roth, A., Sönmez, T., & 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., & Delmonico, F. (2006). Increasing the opportunity of live kidney donation by matching for two and three way exchanges. *Transplantation*, *81*(5), 773–782.

Segev, D. L., & Montgomery, R. A. (2010). The application of paired donation to live donor liver transplantation. *Liver Transplantation*, *16*(4), 423–425.

Sönmez, T., & Ünver, U. (2015). Enhancing the efficiency of and equity in transplant organ allocation via incentivized exchange. Working paper.

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.

Toulis, P., & Parkes, D. C. (2015). Design and analysis of multi-hospital kidney exchange mechanisms using random graphs. *Games and Economic Behavior*, *91*(0), 360–382.

Ünver, U. (2010). Dynamic kidney exchange. *Review of Economic Studies*, *77*(1), 372–414.

Woodle, S., Daller, J., Aeder, M., Shapiro, R., Sandholm, T., Casingal, V., Goldfarb, D., Lewis, R., Goebel, J., & Siegler, M. (2010). Ethical considerations for participation of nondirected living donors in kidney exchange programs. *American Journal of Transplantation*, *10*, 1460–1467.