The Need for (long) Chains in Kidney Exchange (Preliminary version for NBER conference)

Itai Ashlagi David Gamarnik Alvin E. Roth*

October 13,2011

Abstract

It has been previously shown both analytically and numerically that for sufficiently large pools of patient-donor pairs, (almost) efficient kidney exchange can be achieved by using at most 3-way exchanges, i.e. by using exchanges among no more than 3 patient-donor pairs. However, as kidney exchange has grown in practice, exchanges among n>3 pairs have proved useful, and long chains initiated by non-directed, altruistic donors have proven to be very effective. We explore why this is the case, both empirically and theoretically.

We provide an analytical model of exchange when there are many highly sensitized patients, and show that large cycles of exchange or long chains can significantly increase efficiency. As large cycles of exchange cannot be used in practice, long non-simultaneous chains initiated by altruistic donors significantly increase efficiency in patient pools of the size and composition that presently exist.

1 Introduction

Various clearinghouses for kidney exchange today make a thicker market for patients and their incompatible donors who wish to participate in an exchange with other such pairs. Such clearinghouses are an attempt to organize a barter economy on a large scale, with the aid of a computer-assisted clearinghouse.[1]

*Ashlagi: iashlagi@mit.edu. Gamarnik: gamarnik@mit.edu. Roth: al_roth@harvard.edu. We have had valuable conversations about this paper with Jacob Leshno and Mike Rees.

[1]Recall that Jevons (1876) proposed that money was invented as a market design solution to overcome precisely the difficulties of organizing barter economies–in particular, the low volume of trade that could be expected through barter because of the difficulty of satisfying the "double coincidence" of wants involved in simultaneous 2-way exchange.

Roth et al. (2004) proposed how to organize kidney exchange, integrating exchanges (closed cycles of trade) and chains (which need not be closed cycles). Due to incentive and logistical constraints, exchanges need to be of small size. Initially clearinghouses organized exchanges involving no more than 2 patient-donor pairs (Roth et al. (2005b)), and even today exchanges often involve no more than 3 pairs (since a 3-way exchange involves the simultaneous coordination of 6 operating rooms and surgical teams).

Clearinghouses for kidney exchange first conducted only simultaneous exchanges and an important task was to identify efficient outcomes (see Saidman et al. (2006) and Roth et al. (2007)). Roth et al. (2006) proposed that the growing number of non-directed donors would allow the simultaneity constraint to be relaxed. (When a chain is initiated by a non-directed donor, i.e. a donor who does not come with a particular patient who he wishes will receive a kidney, it can be organized so that the non-directed donor makes the initial donation, and no patient-donor pair has to donate a kidney before they have received one. Since this reduces the cost of a broken link, such chains can be organized non-simultaneously, which allows them to be longer, as more operating rooms and surgical teams can be assembled over a longer time.) Particularly after the publication of Rees et al. (2009), which reported the initial ten transplants from the first nonsimultaneous chain, a growing number of long nonsimultaneous chains have been conducted. However there is an ongoing debate about whether long chains increase efficiency beyond simultaneous small exchanges.

The empirical part of this paper examines the patient-donor pairs enrolled in several kidney exchange clearinghouses and shows that allowing long chains significantly increases efficiency. The key for this finding is the high percentage of "highly sensitized" patients in the data. Patients are highly sensitized if (for reasons other than blood type incompatibility) the probability is low that any particular donor's kidney will be compatible for them; i.e. highly sensitized patients are those for whom finding a transplantable kidney will be difficult, even from a donor with the same blood type.

We further provide an analytical framework for studying such highly sensitized pools.

Efficiency has been studied previously quite extensively. Roth et al. (2007) assume compatibility between a patient and a donor of two distinct incompatible pairs in the exchange pool is determined merely according to blood types and show that 4-way exchanges are sufficient for efficiency. Ünver (2010) showed a similar result in a dynamic setting. In practice to receive a donor's kidney a patient needs to also pass a tissue type test (crossmatch test). The more sensitized a patient, the less likely he or she will match with a random donor. However, simulations conducted by Roth et al. (2007) (without this assumption), promisingly show that even on small exchange pool sizes, of 25–100 patient-donor pairs, even using exchanges of at most size 3 hardly results in any efficiency loss. Ashlagi and Roth (2011) show this analytically in large random exchange pools. Although these papers do not study

non-directed donors, one can derive from their models of very large exchange pools that even if each non-directed donor is limited to a chain of length 2, and exchanges to size at most 3, there will be (almost) no efficiency loss.

These results seem to settle the efficiency problem, at least for exchange pools that grow without limit.

Existing clearinghouses, however, operate with relatively small pools of patient-donor pairs. Examining three data sets from three different exchange pools, we find that limiting the maximum size of exchanges to 3-way and even 4-way exchanges results in a significant loss of transplants. Alternatively, when an altruistic donor is in the pool, finding long chains together with exchanges of at most size 3 can produce over 30% more transplants than using only exchanges limited to size 3 or 4.

The data reveal that the empirical percentage of highly sensitized patients in the exchange pool is significantly higher than the percentage of highly sensitized patients found numerically in pools representative of the general patient population (for example in simulations used by Roth et al. (2007)), and even higher than the percentage of highly sensitized patients reported by the hospitals that enroll the incompatible pairs in the kidney exchange pool. The empirical efficiency findings are a direct result of highly sensitized pools – *small* exchanges that include a highly sensitized patient are rare, because highly sensitized patients can receive kidneys from only a few donors, which makes the double coincidence needed for e.g. a 2-way exchange unlikely. In a related paper, Ashlagi et al. (2011a) use simulations, also based on clinical data, to show that long chains increase efficiency in a dynamic setting (see also Ashlagi et al. (2011b)).[2][3]

Previous analytical models that have been adopted to study efficiency either assumed no tissue type incompatibilities or derived it from the limiting assumptions of the model. The relations between various groups of incompatible pairs distinguished by their blood types is now well understood. In this paper we take an opposite approach; essentially we focus on immunological (tissue-type) incompatibilities. For clarity, in section 4 we will model patient-donor pairs of a single blood type, to show how chains and large exchanges become important even in the absence of blood type incompatibility, when patients are highly sensitized. Our results shed light on why chains are proving to be so useful, and how much they add to efficiency.

[2]Understanding the causes of the empirical distribution of sensitivity of patients in the exchange pools is an important topic for future research. At least one cause is that hospitals do not fully participate in kidney exchange programs, for example by not revealing easy to match pairs that they can transplant through internal (to the hospital) exchanges (see Ashlagi and Roth (2011), Ashlagi et al. (2010) and Toulis and Parkes (2010) who study how to prevent this problem by redesigning the mechanism).

[3]Dickerson et al. (2011) further confirm the results by Ashlagi et al. (2011a) that long chains increase efficiency in dynamic settings.

2 A Model for Kidney Exchange

In an exchange pool there are kidney patients, each paired with an incompatible living donor.[4] An exchange pool induces a **compatibility graph** D_V which captures the compatibilities between donors and patients in the pool. The set of incompatible pairs are the nodes of the graph, denoted by V. For any two incompatible pairs $v, u \in V$, there is a directed edge from u to v if and only if the donor of node u is compatible with the patient of node v.

A k-way **exchange** ($k > 0$) involves k incompatible pairs $v_1, v_2, \ldots, v_k$ such that for every $i, 1 \leq i < k$, the donor of v_i is compatible with the patient of v_{i+1}) and v_k is compatible with the patient of v_1.

In practice the size of an exchange is limited (mostly due to logistical constraints). We assume there is an exogenous maximum size limit $k > 0$ for any exchange. Thus if $k = 3$ only exchanges of size 2 and 3 can be conducted.[5]

Although exchanges of size more than 3 are not used in practice due to simultaneity constraints, *chains*, initiated by an altruistic donor can be conducted non-simultaneously and thus can be of any size. A k-way **chain** ($k > 0$) involves an altruistic donor and k incompatible pairs $v_1, v_2, \ldots, v_k$ such that for every $i, 1 \leq i < k$ donor of v_i is compatible with the patient of v_{i+1}).[6]

An allocation is a set of distinct exchanges. An allocation M is called (k, l)-**efficient** if it matches the maximum number of transplants possible using only exchanges of size no more than k and chains of size no more than l. It is called k-**efficient** if it it is $(k, 0)$-efficient. It is called **efficient** if it is (k, l)-efficient for unbounded k and l.

Roth et al. (2007) showed that in an exchange pool without altruistic donors and with no tissue type incompatibilities across incompatible pairs, there is an efficient allocation which uses exchanges no more than size 4 (assuming no tissue type incompatibilities implicitly assumes a large pool). Ashlagi and Roth (2011) relaxed this assumption and showed that in a large random exchange pool, based on a Bayesian distribution consistent with practice, exchanges of size no more than 3 are needed for efficiency.[7] Both these papers assume the exchange pool is large. However, Roth et al. (2007) also support these results using

[4]Pairs that are compatible would presently go directly to transplantation and not join the exchange pool (but see e.g. Roth et al. (2005a) and Sönmez and Ünver (2011) on the advantages of changing this policy).

[5]In the APD and NEPKE k was originally set to 2, was increased to 3, and now optimization is conducted over even larger exchanges and chains, and the pilot national program considers exchanges up to size 3. Exchanges are generally conducted simultaneously, so an exchange of size k requires 2k operating rooms and surgical teams for the k nephrectomies (kidney removals) and k transplants.

[6]One can model chains begun by nondirected altruistic donors as (cyclic) exchanges by adding the altruistic donors to the pool and associate each of them with a virtual patient that can receive a kidney from any other donor in the pool.

[7]See also Toulis and Parkes (2010) for 2-way efficient allocations.

simulations on small exchange pools of sizes 25 to 100.

Although these studies do not consider altruistic donors, one can show that in the limit as the pool grows large, each such donor will contribute only two more transplants to the efficient outcome (see Section 4.1). These results seem to provide a good understanding about efficiency in kidney exchange even with the presence of altruistic donors . We shall see however in the next section that in real data 3 and even 4-way exchanges are not sufficient to provide efficiency, and thus chains will be needed.

3 Empirical Efficiency

The empirical data we present here is from 3 kidney exchange clearinghouses. Each patient has a level of percentage reactive antibodies (PRA) which determines the likelihood that the patient will be compatible with a random donor. The lower the PRA of a patient, the more likely the patient is compatible with a random donor. In their simulations, Roth et al. (2007) use a PRA distribution which is taken from the UNOS data (see the first column in Table 1). Three intervals are described, low (0-9), medium (10-79) and high (80-100).

In all data sets we can estimate the PRA by observing for each patient the number of blood type compatible donors it is also tissue-type compatible with. In two of the data sets we also have the reported PRA by the hospitals that submit the medical data about their patients. Table 1 provides the estimated and reported (in parenthesis) distributions. Only in the 1st data set we also have historical incompatible pairs (i.e. the set of all pairs enrolled in a clearinghouse, as opposed to a snapshot of the pairs still enrolled at a given time), thus the two columns for this data set. The size of the sets are: 361,131,251 and 74 respectively.[8]

	Unos	Historical 1st Set	1st Set	2nd Set	3rd Set
Size		497	131	251	74
Low PRA	70.19	24.93 (36.6)	21.37	6.77	0 (24.32)
Med PRA	20	23.26 (27.2)	12.97	13.14	38.38 (14.86)
High PRA	9.81	51.80 (36.2)	65.64	80.07	71.62 (60.810)

Table 1: PRA distributions in various real data sets.

Naturally the PRA distributions in the kidney exchange data sets will differ from the PRA distribution in UNOS since low PRA patients are easier to match and therefore less likely to

[8]Dickerson et al. (2011) also mention that almost all 77 patients in the the UNOS program which launched in 2010 are highly sensitized.

join the pool. One way to generate a PRA distribution that is more realistic is by generating incompatible pairs as follows: first randomize a patient, and then randomize a number of potential donors for this patient. If either one of the donors matches the patient this patient is assumed to find a donor and is not part of an incompatible pair. Similar processes were used in Roth et al. (2007), Ashlagi et al. (2011a) and Ashlagi and Roth (2011). In Table 2 we report the PRA distributions using this process for various distributions of number of donors for each patent. We also report in parenthesis the PRA distribution in the population of the compatible donors that is induced by this process. The blood type of a patient as well PRA of a patient in this process are similar to the ones used in Roth et al. (2007).

Note that uniformly generating 1-5 donors for each patient produces a PRA distribution that is "close" to the reported distribution in 1st data set. Furthermore, the distribution of compatible pairs generated by these processes doesn't change substantially. Note that only when the generating process of incompatible pairs is very conservative (many potential donors for each patient), does the percentage of high PRA patients gets close to the percentage of high PRA patients in the the empirical 1st data set. Only when we get up to the (unrealistically high) level of 5 donors for each patient does the percentage of highly sensitized patients reach the level we observe in the 1st data set. This suggests that the high levels of highly sensitized patients in the data is at least in part due to hospitals withholding the more easily matched pairs.

	$1-3$ donors	$1-5$ donors	$1-13$ donors	5 donors	6 donors
Low PRA	47.65 (74.26)	39.7 (73.9)	28.73 (73.77)	17.04 (73.62)	12.2 (73.75)
Med PRA	28.21 (14.08)	27.67 (15.7)	22.12 (18.99)	24.41 (18.51)	21.88 (19.18)
High PRA	24.14 (1.62)	32.63 (2.1)	49.15 (3.56)	58.55 (2.83)	65.82 (3.16)

Table 2: PRA distributions from generated incompatible pairs processes. All but the last column provides the distribution generated using a different uniform distribution of random potential donors for each patient. The last column is the PRA distribution among all patients who received transplants since 2008 according to UNOS (in parenthesizes are the PRA distributions of compatible pairs generated from the process.

The very high percentage of high PRA patients (even reported PRA) in the exchange pool suggests that compatibility graphs are sparse, and, so exchanges and short chains might not be sufficient. We looked at the allocations in each of these sets allowing various maximum exchange sizes. The results, reported in Table 3, show that by limiting only to 3 and even 4-way exchanges there is a substantial efficiency loss.

To summarize, we observed that large exchanges (and thus long chains) are needed for efficiency in exchange pools of the size and composition we are seeing in clinical settings.

	$k=2$	$k=3$	$k=4$	$k=5$	$k=\infty$	$(3,\infty)$	(∞,∞)	$(3,5)$
Historical 1st Set	118	169	189	191	193	–	–	
Current 1st Set	6	9	11	13	16	–	16	–
2nd Set	22	42	45	47	51	54.5	54.5	47.6
3rd Set	0	3	4	4	4	7.5	7.5	5.8

Table 3: Cycles in Data. In the first 5 columns, we report the size of the k-efficient allocation where ∞ stands for unrestricted k. In the last 3 columns a single random altruistic donor (also from the data) is added. In those columns the size of efficient allocation is reported given (k,l) where k is the size of the maximum exchange and l is the size of the maximum length of a chain.

This follows from the high percentage of high PRA patients, which makes the compatibility graphs sparse. An important question for future research and design is to understand why is there such a big gap between the predicted PRA distribution and the empirical PRA distribution. Alternatively, what causes the big difference between the efficiency results obtained by Roth et al. (2007) and Ashlagi and Roth (2011), and the empirical efficiency results we show here? At least part of the answer has to do with hospitals not revealing their incompatible pairs to the exchange pools (see Ashlagi and Roth (2011)).

An additional factor that adds to this gap is the accumulation of very highly sensitized patients over time due to the dynamic nature of exchange programs; consider for example a very highly sensitized patient p in the pool who can receive a kidney from a donor d who is also currently in the the pool. If p and d cannot be part of a 2-way or a 3-way exchange, this donor might be part of a different exchange making the highly sensitized patient p wait a long time for a different compatible donor.

In the next section we adopt a stylized model based on Erdos-Renyi random graphs, to show that exchanges of size more than 3 (and therefore chains) are often needed for efficiency. In the following section we will see that the absent of a linear fraction of unsensitized patients leads to a linear efficiency loss (linear in the size of the pool). That is, chains are needed in pools with many highly sensitized patients, and the presence of unsensitized patients facilitates finding matches for the sensitized patients. (So when hospitals withhold their easy to match pairs, and the resulting pool is highly sensitized, many fewer matches can be made.)

4 Random Compatibility Graphs with Many Highly Sensitized Patients

Ashlagi and Roth (2011) analyzed random compatibility graphs that are generated from a blood-type and tissue type distributions consistent with practice. In this section we will simplify the distribution from which we generate the graph and assume all patients and donors have the same blood type. In particular edges only depend on whether a donor and a patient are tissue type compatible. This is equivalent to focusing on a subgraph of the real compatibility graph. Roth et al. (2007) and Ashlagi and Roth (2011) showed that, as the number of patient-donor pairs goes to infinity, efficient allocations match *all* the pairs in those subgraphs to each other in exchanges of size at most 3. We will show here that under more careful modeling of highly sensitized patients, even on those subgraphs, exchanges longer than 3 and even 4-way are often needed for efficiency. [9]

In a random compatibility graph there are n pairs. Recall that the PRA of a patient is interpreted as the probability in which a patient will match a random donor. We simplify here the PRA characteristics and assume there exist two levels of PRA, L and H. The probability that a patient with PRA L (H) and a random donor are tissue type compatible is given by p_L (p_H). We assume $p_L > p_H$. Furthermore, the probability that a random patient has PRA L is given by $\upsilon \geq 0$. A random compatibility graph is denoted by $D(n, p_L, p_H, \upsilon)$.

Before we proceed to our results it will be useful to note that our model is closely related to a directed Erdos-Renyi random graph. A directed (Erdos-Renyi) random graph $D(n, p(n))$ is a graph with n nodes and for every two distinct nodes v, u there is an edge from u to v with probability $p(n)$. Thus, if all donors have the same PRA, the graph is exactly a directed random graph. The efficiency result by Ashlagi and Roth (2011) was derived from a theorem of Erdos and Renyi which characterize the existence of a perfect matching in large random graphs (in graph theory terminology, a perfect matching is an allocation using only exchanges of size 2):

Erdos-Renyi Theorem: Let $p(n) = (1+\tau)(\frac{\ln n}{n})^{\frac{1}{2}}$ where $\tau \neq 0$. In almost every large graph $D(n, p(n))$ there exists a perfect matching if and only if $\tau > 0$.

In practice the probability that a patient is compatible with a random donor does not depend on the size of the population. From the Erdos-Renyi theorem one can show if γ_L and γ_H are constants, an efficient allocation using exchanges of size at most 3 exists in $D(n, \gamma_L, \gamma_H, \upsilon)$ as n goes to infinity.

To model highly sensitized patients and the sparse compatibility graphs they generate we

[9] Our results can be generalized to other subgraphs in the original random exchange pool analyzed in those papers, e.g. the subgraph induced by only incompatible pairs with an A patient and B donor and vice versa.

will neglect the fact that compatibility between a random patient and a random donor doesn't depend on the size of the population and assume that γ_H is a decreasing function of n.[10] In particular we will analyze sparse large graphs (instead of directly analyzing finite sparse graphs which make the problem intractable). For "medium" size pools such probabilities make a good approximation since highly sensitize patients have very low probability to be matched. For any k, we will denote by W_k the size of a $k-$efficient allocation when $D(n, p_L, p_H, \upsilon)$ n is infinitely large.

In the next section we look more carefully into specific sets in the "real" compatibility graphs and show that reinvestigating the structure of efficient allocations is needed. In the subsequent sections we analyze two different settings that will shed light on why long exchanges are useful for efficiency in sparse graphs.

4.1 Towards a (new) efficient allocation - more empirical findings

Before we continue with our model, let us recall the following result with blood types:

Theorem 4.1 (Ashlagi and Roth (2011))**.** *Under any [sufficiently regular] blood type distribution, in almost every large (limit) graph without altruistic donors there exist an efficient allocation with exchanges of size at most 3 whose structure is as described in Figure 3.*

Theorem 4.1 implies not only that exchanges of more than 3 are not needed but also not "many" exchanges of size 3 are needed. The longer exchanges are (even if feasible) the harder they are to implement, thus such a structure is important for practice. The Theorem also implies that each altruistic donor can add at most 2 transplants. Under an efficient allocation as described in Figure 3, an altruistic donor with blood type A,B or O can start a chain with two underdemanded pairs (e.g. an A altruistic donor can give to an O-A pair who can give to an A-AB pair).

As we observed in Section 3, long cycles and chains are needed for efficiency in real data. We take here a closer look at the set of incompatible pairs in which both the patient and the donor have blood type O (as an example). As described in Table 4, exchanges of size even 3 are not sufficient to match all O-O pairs and are not efficient within this set. Importantly, highly sensitized patients significantly benefit from longer exchanges. Observe that almost all low PRA patients are matched already for $k = 3$ and significantly more high PRA patients are matched when longer than 3-way exchanges are allowed. For example, in the entire set,

[10] One way to motivate the assumption that the probability that a patient will be compatible declines as the number of patient-donor pairs grows is if the population grows over time, so that there is some probability that a patient may no longer be healthy enough for a transplant from donors who enter the pool further in the future. We don't formally take this approach here.

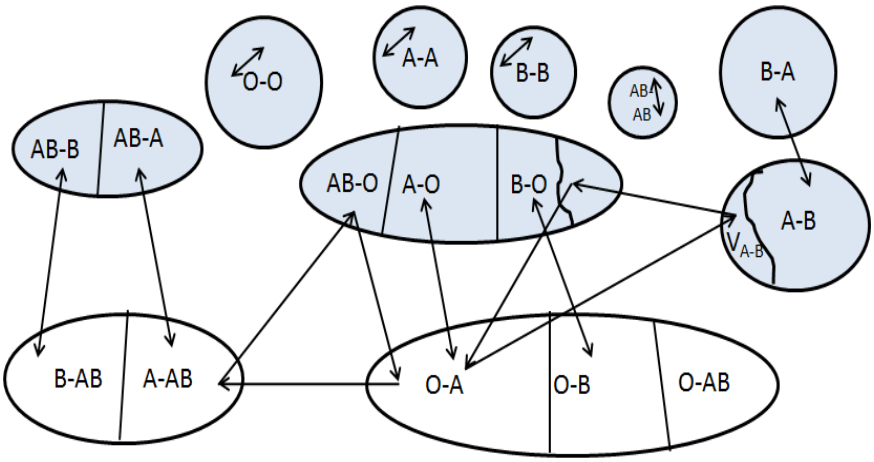

Figure 1: The structure of an efficient allocation without altruistic donors. All selfdemanded pairs are matched to each other. (excluding all selfdemanded pairs). All B-A pairs are matched to A-B (assuming there are more B-A than A-B), the remainder of the A-B pairs are matched in 3-way exchanges using O-A's and B-O's. AB-O are matched in 3-ways each using two overdemanded pairs, and every other overdemanded pair is matched to a corresponding underdemanded pair.

all 26 low PRA patients are matched already for $k = 3$ but $k = 4$ yields 6 more high PRA matches than $k = 3$.

Table 5 is similar to Table 4 only the high PRA is set to 90.

Finally, we plot in Figure 2 a real graph induced by the incompatible pairs in which both the donor and the patient have blood type A. Observe that many patients can get none or one kidney from other donors, while their incompatible donors are compatible with other patients. Furthermore, blue edges are part of cycles and blue nodes are incompatible pairs with low PRA patients. Observe that there are many edges and paths that are not part of cycles, which should be interpreted as potential chains.

4.2 All patients are highly sensitized

We assume in this section that all patients are highly sensitized. Our intention here is to show that whether or not long exchanges are needed depends on the sparsity of the graph.

Technically, the random compatibility graph is just $D(n, p(n))$, where $p(n) = p_H(n)$. If $p(n)$ is "very small", exchanges do not exist at all. The Erdos-Renyi Theorem implies that for large enough $p(n)$, allocations with only 2-way exchanges can cover all but at most one pair, i.e. $W_2 + 1 \geq W_k$ for all $k > 2$. The next result shows that for every k, when the graph is sufficiently sparse, $k+1$-efficient allocations provide more transplants than k-efficient allocations.

Size		$k=2$	$k=3$	$k=4$	$k=5$
30 (11.54,18.46)	matches	12.52	18.91	20.57	21.6
	exchanges	21.9	122.47	565.22	2098.53
	matched (L,H)	(8.14, 4.38)	(11.42, 7.49)	(11.29, 9.28)	(11.21, 10.39)
45 (18.07, 26.93)	matches	22.16	32.56	36.22	36.96
	exchanges	50.35	391.47	2637.34	19641.47
	matched (L,H)	(13.29, 8.87)	(17.41, 15.15)	(17.29, 18.93)	(17.46, 19.5)
65 (26, 39)	matches	36	52	58	58
	exchanges	107	1180	11932	123963
	matched (L,H)	(21, 15)	(26, 26)	(26, 32)	(26, 32)

Table 4: Cycles in selfdemanded pairs. High PRA is considered to be 80. The pools size is given in the first column where in parenthesizes are the number of low and high PRA patients respectively. The results in the last block refer to the entire pool we draw from. The rows in each block describe: (i) number of matches obtained, (ii) number of all exchanges (not necessarily disjoint), and (iii) number of matched low and high PRA patients restrictively.

Theorem 4.2. *For every $k \geq 2$, there exists $p_k(n)$ such that $W_k = o(W_{k+1})$ in almost every graph $D = D(n, p_k(n))$. That is $\frac{W_k}{W_{k+1}}$ tends to zero as n tends to infinity.*

The proof 4.2 will follow directly from the next lemma which states in more detail the size of exchanges for various k's. The idea of the proof is to show that in sparse enough graphs the number of exchanges of size exactly $k+1$ in the pool is approximately the same number of disjoint exchanges of size $k+1$. Furthermore the number of exchanges of size at most k is smaller than the number of exchanges of size $k+1$.

Lemma 4.3. *Let $D(n, p(n))$ be a random directed graph with $p(n) = n^{-1+\epsilon}$ for some real $\epsilon > 0$.*

1. *If $0 < \epsilon < \frac{1}{k+1}$, $W_k = \Theta(n^{k\epsilon})$ in almost every $D(n, p(n))$. Furthermore the size of the k-efficient allocation is $\Theta\left(n^{k\epsilon}\right)$.*[11]
2. *If $\epsilon < 0$, $W_k = o(1)$ in almost every $D(n, p(n))$.*
3. *If $\epsilon > \frac{1}{k}$, $W_k = n$ in almost every $D(n, p(n))$.*

In the next section we analyze allocations in compatibility graphs that have a mixture of low and high sensitized patients.

[11] For two functions $f(n)$ and $g(n)$ we write $f = \Theta(g)$ if there exist constants $c_1 < c_2$ such that $c_1 f(n) \leq f(n) \leq c_2 g(n)$ for every sufficiently large m.

Size		$k=2$	$k=3$	$k=4$	$k=5$
30 (17.64, 12.36)	matches	12.1	18.8	20.43	21.25
	exchanges	21.5	124.61	490.46	2594.77
	matched (L,H)	(10.3, 1.8)	(15.44, 3.36)	(16.33, 4.1)	(16.7, 4.55)
45 (26.83, 18.17)	matches	21.62	32.36	36.03	36.87
	exchanges	48.39	376.08	2668.54	18624.77
	matched (L,H)	(17.91, 3.71)	(25.75, 6.61)	(26.44, 9.59)	(26.09, 10.78)
65 (39, 26)	matches	36	52	58	58
	exchanges	107	1180	11932	123963
	matched (L,H)	(29, 7)	(39, 13)	(39, 19)	(39, 19)

Table 5: Cycles in selfdemanded pairs. High PRA is set to 90.

4.3 Many (but not all) highly sensitized patients $0 < \upsilon \leq \frac{1}{2}$

In this section we will show that even with a substantial fraction of low sensitized patients in the graph, considering only 2 and 3-way exchanges is not sufficient for efficiency. For simplicity we will say a node/pair is an H (L) node if its patient has H (L) PRA. An assumption we make through out this section is that the chance that a patient has Low PRA is smaller than 1/2, i.e. $\upsilon \leq \frac{1}{2}$. This assumption is consistent with the the empirical findings.[12] As before $D = D(n, p_L(n), p_H(n), \upsilon)$ denotes a random compatibility graph. In this section we assume $p_L = p_L(n)$ is a constant and $p_H(n) = \frac{c}{n}$ for some constant $c > 0$. This reflect the assumption that incoming edges to H pairs have a small probability.

For our main result we need the following lemma about matchings in bipartite non-directed random graphs. Given $\upsilon \in (0,1)$ and c, let $G(n, \upsilon, c)$ be an undirected bi-partite random graph with υn and $(1-\upsilon)n$ nodes on either sides and independent edge probability $\frac{c}{n}$. Let $M(n, \upsilon, c)$ denote the size of a largest matching in this graph.

Lemma 4.4. *The limit $\alpha(\upsilon, c) \triangleq \lim_n M(n, \upsilon, c)/n$ exists and satisfies the following properties:*

1. $0 < \alpha(\upsilon, c) < \min(\upsilon, 1-\upsilon)$,
2. *$\alpha(\upsilon, c)$ is a strictly increasing function in c, when υ is fixed.*

In our main result in this section we show that 4-efficient allocations yield significantly more transplants than 2-efficient allocations. Furthermore, there is a threshold of low PRA patients above which 4-efficient allocations yield significantly more transplants than

[12] One can derive similar results for the reversed inequality.

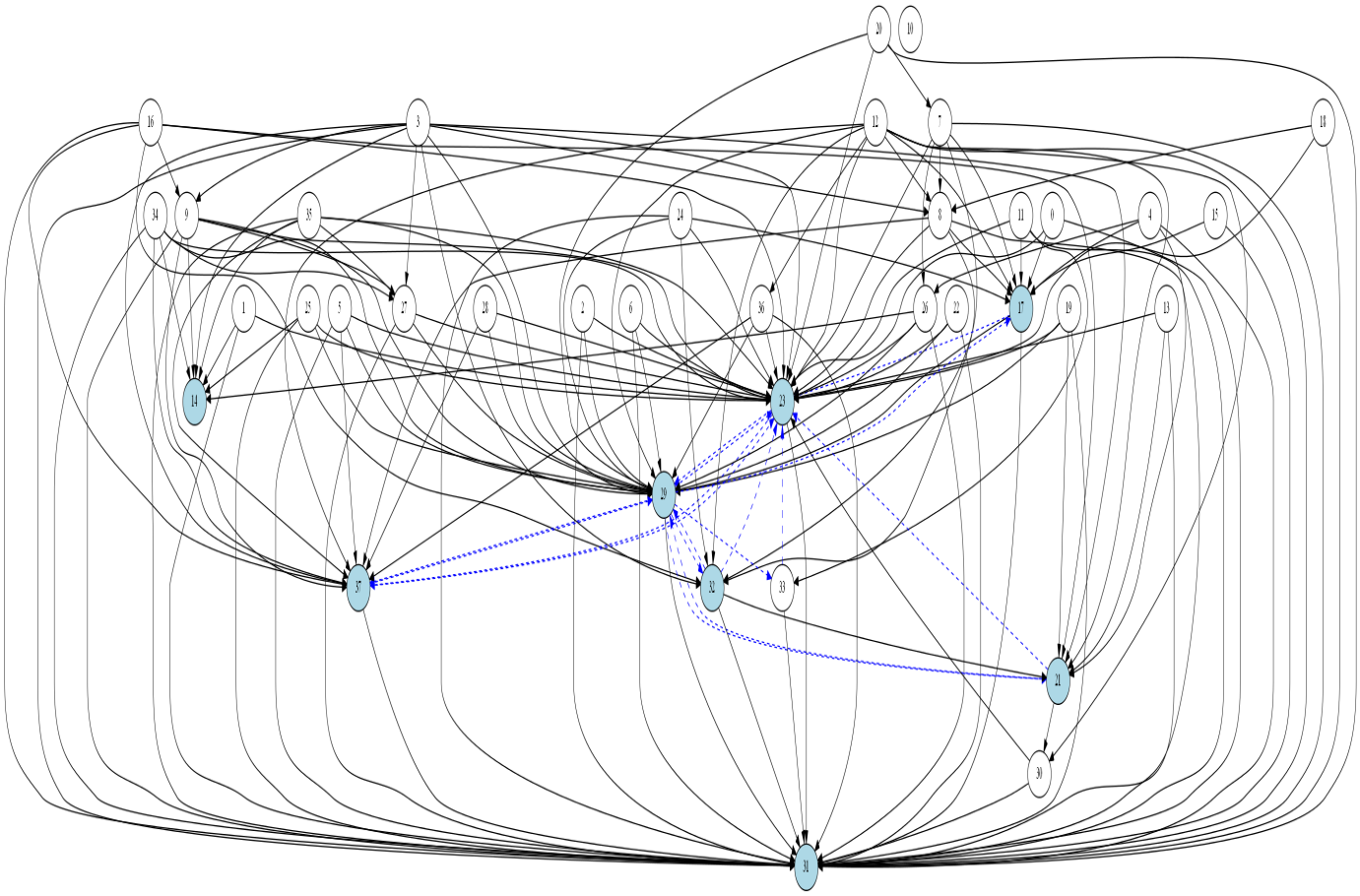

Figure 2: The graph induced by A-A incompatible pairs in the historical data set. Blue dashed edges are part of exchanges (cycles) and blue filled nodes are incompatible pairs whose patient has a low PRA.

3-efficient allocations. Finally, for all $k \geq 2$, every k-efficient allocation matches all low PRA patients, which means that high PRA patients are the ones that benefit from 4-way exchanges (or chains). Note that this is consistent with the empirical findings in Section 4.1.[13] We define $\alpha(\cdot,\cdot)$ as in Lemma 4.4.

Theorem 4.5. *With probability approaching one the the following holds*

1. $W_4 = \alpha(\upsilon, c)n + \upsilon n + o(n)$, *and* $W_4 = W_2 + \rho n + o(n)$, *for some* $\rho > 0$. *In particular, 2-efficient allocations are smaller than the efficient allocation by at least a linear in* n *value.*

2. *If* $\alpha(\upsilon, c(1-p_L)) < \upsilon - \alpha(\upsilon, cp_L)$, *then* $W_4 = W_3 + o(n)$, *and otherwise* $W_4 = W_3 + \rho n + o(n)$, *for some* $\rho > 0$. *In particular, the 3-efficient allocations are nearly efficient*

[13] For $k = 2$ not all low PRA are matched in the data since the sample size is quite small.

in the former case, but are smaller than the efficient allocation by at least a linear in n value.

The proof is given in the appendix and we provide here the intuition for the first part (the intuition is similar proof for the second part).

Let us begin with two observations that hold for both 2-efficient and 4-efficient allocations.

(i) The number of matches that can be achieved using exchanges that involve only H pairs is $o(n)$ (by Lemma 4.3), in particular it will be negligible in comparison to the number of matches achieved at these efficient allocations.

(ii) Any linear fraction of L pairs can all (but at most one) be matched in a 2-way allocation.

We next construct a 2-way allocation that is approximately 2-efficient.[14] Consider the maximum 2-way allocation M_2 that involves only exchanges with one H node and one L node (see Figure 3). Applying Lemma 4.4 shows that the number of exchanges in M_2 is $\alpha(\upsilon, cp_L)n$. Adding to M_2 an allocation that matches all remaining L nodes (to each other) using 2-way exchanges (which is feasible due to the second observation) results by the first observation in the desired allocation.

Consider next a maximum set of disjoint (one-way) directed edges from L nodes to H nodes, M_4. By Lemma 4.4 the cardinality of such a set is $\alpha(\upsilon, c)n > \alpha(\upsilon, cp_L)n$. To prove the result it remains to show that all nodes in the graph induced only by nodes in M_4 can be matched to each other in 4-way exchanges, since again all remaining L nodes can be matched to each other using 2-way allocations.

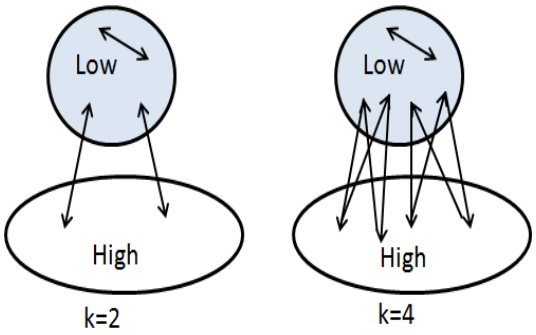

Figure 3: Construction of allocations with $k = 2$ and $k = 4$ in a graph in which all pairs are blood type compatible and at least half the pairs are highly sensitized (for example the graph induced by O-O pairs).

[14] Up to $o(n)$ from the maximum size.

The idea follows from the fact that "returning" to L nodes has a larger probability. Let Q be the set of L nodes in M_4. One can construct an artificial graph $\tilde{G}$ with set of nodes Q, and an edge between every two nodes $v, w \in Q$ exists if there is a 4-way exchange in the graph induced by M_4 involving two H nodes, v and w (see Figure 3). The key is to notice that an edge in $\tilde{G}$ exists with constant probability, p^2. Applying the Erdos-Renyi Theorem to the $\tilde{G}$ completes the argument.

Importantly, Theorem and its proof imply that allowing exchanges longer than 4 will benefit high PRA patients.

In our last result we will show that there are still many pairs that can be matched through potential chains: allowing for longer chains will add significantly more H matched pairs. In particular, beyond the 4-efficient allocation constructed in the previous theorem, there are still many more potential chains that involve high PRA patients.

The following definition is needed. A (k, l)-quasi-allocation is a set of disjoint exchanges of size at most k and paths (chains not starting with an altruistic donor) of size l. The random compatibility graph does not include altruistic donors and therefore we are interested in quasi-allocations (to which a non-directed altruistic donor could be appended to initiate a chain). We denote by S_l the size of a maximum $(4, l)$-quasi-allocation.

Theorem 4.6. *Fix $l \geq 0$. There exists $\rho_l > 0$ such that $S_{l+1} \geq S_k + \rho_l n$. Furthermore, an efficient $(4, l+1)$-quasi-allocation yields $\rho_l n$ more high PRA matches than $(4, l)$-quasi-allocation.*

5 Experimental Results

In Table 6 we provide simulation results using real data (not limited to specific types of pairs) that show the benefit of chains. For each scenario we ran 300 trials drawing pairs and altruistic donors from the entire data with re-sampling.

Observe that long chains allow more than 10% more transplants than when chains are bounded to length 3 when the pool size is more than 50. Note that the additional high PRA patients who are matched often substantially exceeds 10%. Finally, allowing for unbounded chains and 3-way exchanges obtains almost the same number of matches that can be achieved when both exchanges and chains can be unbounded in size, so long chains substitute well for large exchanges when there are some non-directed donors available.

6 Conclusion

There are a number of ways in which barter may be inefficient. Jevons (1876) famously pointed to the double coincidences needed to make pairwise exchanges. Theory based on large graphs that become very well connected in the limit predicted that also allowing 3-way exchanges and short chains would overcome this difficulty. However the compatibility graphs so far seen in practice are sparse. We show here empirically the effectiveness of long non-simultaneous chains. In particular long chains are needed for efficiency due to the sparsity of compatibility graphs that arise when many patients are very highly sensitized. We further provide theoretical foundations to better understand the structure of efficient allocations in sparse comparability graphs. Our findings suggest that tissue type incompatibilities are of first order importance in understanding efficient kidney exchange, and that as long as we see such a high percentage of highly sensitized patients, long chains will help achieve more transplants.

References

I. Ashlagi and A. E. Roth. Free riding and participation in large scale, multi-hospital kidney exchange. Working paper, 2011.

I. Ashlagi, F. Fischer, I. Kash, and A.D. Procaccia. Mix and Match. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC), 2010*, pages 295–304, 2010.

I. Ashlagi, D. S. Gilchrist, A. E. Roth, and M. A. Rees. Nonsimultaneous Chains and Dominos in Kidney Paired Donation - Revisited. *American Journal of Transplantation*, 11(5):984–994, 2011a.

I. Ashlagi, D. S. Gilchrist, A. E. Roth, and M. A. Rees. Nonsimultaneous Chains in Transplantation. *American Journal of Transplantation*, 2011b. forthcoming.

T. Bohman and A. Frieze. Karp-sipser on random graphs with a fixed degree sequence. *Combinatorics, Probability and Computing*, 20(5):721–741, 2011.

J. Dickerson, A. D. Procaccia, and T. Sandholm. Optimizing Kidney Exchange with Transplant Chains: Theory and Reality. Unpublished, 2011.

W. S. Jevons. *Money and the Mechanism of Exchange*. New York: D. Appleton and Company., 1876. http://www.econlib.org/library/ YPDBooks/Jevons/jvnMME.html.

M. A. Rees, J. E. Kopke, R. P. Pelletier, D. L. Segev, M. E. Rutter, A. J. Fabrega, J. Rogers, O. G. Pankewycz, J. Hiller, A. E. Roth, T. Sandholm, M. U. Ünver, and R. A. Montgomery. A non-simultaneous extended altruistic donor chain. *New England Journal of Medecine*, 360:1096–1101, 2009.

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

A. E. Roth, T. Sönmez, and M. U. Ünver. A kidney exchange clearinghouse in New England. *American Economic Review Papers and Proceedings*, 95(2):376–380, 2005a.

A. E. Roth, T. Sönmez, and M. U. Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125:151–188, 2005b.

A. E. Roth, T. Sönmez, and M. U. Ünver. Efficient kidney exchange: coincidence of wants in markets with compatibility-based preferences. *American Economic Review*, 97:828–851, 2007.

S. L. Saidman, A. E. Roth, T. Sönmez, M. U. Ünver, and F. L. Delmonico. Increasing the Opportunity of Live Kidney Donation by Matching for Two and Three Way Exchanges. *Transplantation*, 81:773–782, 2006.

T. Sönmez and M. U. Ünver. Altruistic Kidney Exchange. Unpublished, 2011.

P. Toulis and David Parkes. A Random Graph Model of Kidney Exchanges : Optimality and Incentives. Working paper, 2010.

M. U. Ünver. Dynamic Kidney Exchange. *Review of Economic Studies*, 77(1):372–414, 2010.

7 Appendix

Proof of Theorem 4.3. Denote by S_k the number of k-way exchanges in $D(n,p)$. Thus for some constant $c > 0$

$$ES_k = c\binom{n}{k}p^k \approx (np)^k = \Theta\left(n^{k\epsilon}\right) \tag{1}$$

Since k is a constant $E\left(\sum_{l=2}^{k} S_l\right) = \Theta\left(n^{k\epsilon}\right)$. Since the number of disjoint cycles is at most the number of cycles the second part follows.

We proceed with the first part. Let $0 < \epsilon \leq \frac{1}{k}$. Denote by Q_k the number of disjoint exchanges of size k. By (1) $E\left(S_l\right) = o(S_{k+1})$ for any $2 \leq l \leq k$. It is sufficient to show that $Q_k = \Theta(S_k)$. We will show this for $k = 3$ (for larger k's the proof is similar).

Denote by Y_k the number of non-disjoint k-way exchanges.

Let $\epsilon < \frac{1}{3}$. To see that $Q_3 = \Theta(S_3)$ it is sufficient to show that $Y_3 = o(S_3)$. First we show that $EY_3 = o(n^{3\epsilon})$. Observe that two different 3-way exchanges in $D(m, p(m))$ can overlap by either having exactly one edge in common or having exactly one node in common. Therefore

$$EY_3 \approx n^4 p^5 + n^5 p^6 = n^{-1+5\epsilon} + n^{-1+6\epsilon} = O\left(n^{-1+6\epsilon}\right) = o\left(n^{3\epsilon}\right) \tag{2}$$

where the last inequality follows since $\epsilon < \frac{1}{3}$.

Let $0 < \delta < 1 - 3\epsilon$. By Markov's inequality

$$\Pr\left(Y_3 \geq n^{-1+6\epsilon+\delta}\right) \leq \frac{n^{-1+6\epsilon}}{n^{-1+6\epsilon+\delta}}. \tag{3}$$

Therefore $Y_3 = o(ES_3)$.

We next show that S_3 is concentrated around its mean. The covariance between any two potential 3-way cycles is 0 unless the have exactly one edge in common. Therefore

$$Var(S_3) = 2\binom{n}{4} p^5 \approx 2n^{-5+4\epsilon} \approx n^{5\epsilon-1}.$$

By Chevyshev's inequality for small enough $\delta > 0$

$$\Pr\left(|S_3 - ES_3| \geq n^{3\epsilon-\delta}\right) \leq \frac{2n^{5\epsilon-1}}{n^{6\epsilon-2\delta}}, \tag{4}$$

implying that $S_3 = O(ES_3)$ and therefore $Y_3 = o(S_3)$. ∎

Proof of Lemma 4.4. The existence and the actual value of the limit $\alpha(\upsilon, c)$ is a delicate issue, but can be addressed using the Karp-Sipser algorithm technique (see Bohman and Frieze (2011)).

Now let us established the claimed properties. First, since each side contains a linear number of isolated nodes, then $\alpha(\upsilon, c) < \min(\upsilon, 1 - \upsilon)$. Similarly, there exists a linear set of isolated edges, and thus $\alpha(\upsilon, c) > 0$. To prove strict monotonicity, fix $c < c'$ and consider graphs $G(n, \upsilon, c)$ and $G(n, \upsilon, c')$ coupled on the same probability space by adding edges to $G(n, \upsilon, c)$ with probability x such that $c/n + (1 - c/n)x = c'/n$. The newly obtained graph is exactly $G(n, \upsilon, c')$. Since the extra edges create a linear size matching between isolated nodes in $G(n, \upsilon, c)$, we see that $M(n, \upsilon, c')$ is larger than $M(n, \upsilon, c)$ by a linear term and the proof is complete. ∎

Proof of Theorem 4.3. We will first show how to construct an (approximately) 2-efficient allocation and then show how to improve the number of matches using also 4-way exchanges. Let $\mathcal{L}$ and $\mathcal{H}$ denote the set of L and H nodes in the graph $D = D(n, p_L, c/n, \upsilon)$ respectively. Let $\mathcal{W}_k$ denote (any) allocation achieving W_k.

First we observe that by Theorem 4.3 and a simple monotonicity of edge likelihoods, for any $k \geq 2$, the size of the k-efficient allocation in the graph induced by $\mathcal{H}$ is $o(n)$. We therefore assume without loss of generality from now one that there are no edges between any two H pairs.

Claim 7.1. *With high probability,* $W_2 = \alpha(\upsilon, cp_L)n + \upsilon n + o(n)$.

Proof. Consider an undirected bi-partite graph between nodes $\mathcal{L}$ and $\mathcal{H}$, where an edge between nodes $u \in \mathcal{L}$ and $v \in \mathcal{H}$ exists if and only if directed edges (u, v) and (v, u) both exist in D. Since each such edge exists with probability $(c/n)p_L$ independently for all edges, then this graph is precisely $G(n, \upsilon, cp_L)$. Now observe that two way exchanges involving nodes in $\mathcal{H}$ constitute a matching between $\mathcal{H}$ and $\mathcal{L}$ in $G(n, \upsilon, cp_L)$. Thus the portion of $\mathcal{W}_2$ involving $\mathcal{H}$ is upper bounded by $\alpha(\upsilon, cp_L)n + o(n)$, and, as result W_2 is at most the claimed size. On the other hand, let M_2 be the any optimal size matching in $G(n, \upsilon, cp_L)$ between nodes $\mathcal{L}$ and $\mathcal{H}$. Let $\mathcal{L}_2$ be the nodes in $\mathcal{L}$ not involved in M_2. Consider an undirected graph on $\mathcal{L}_2$, where nodes u and v are connected if directed edges (u, v) and (v, u) exist in D. Since this defines a dense random graph on $\mathcal{L}_2$ with edge probability p_L^2, there exists a perfect (or nearly perfect) matching between these nodes. This matching together with M_2 gives a two-way exchange with size $\alpha(\upsilon, c)n + |\mathcal{L}| + o(n) = \alpha(\upsilon, c)n + \upsilon n + o(n)$, as claimed. ∎

Claim 7.2. *With high probability,* $W_4 \geq \alpha(\upsilon, c)n + \upsilon n + o(n)$.

Notice the claim coupled with Claim 7.1 and $p_L < 1$ implies that $W_4 = W_2 + \rho n + o(n)$ for some $\rho > 0$, namely the second claim of our theorem.

Proof. Consider an undirected bi-partite graph between nodes $\mathcal{L}$ and $\mathcal{H}$, where an edge between nodes $u \in \mathcal{L}$ and $v \in \mathcal{H}$ exists if and only if the directed edges (u, v) exists in D (but no requirement on the opposite edge (v, u) is placed). Since each such edge exists with probability (c/n) independently for all edges, then this graph is precisely $G(n, \upsilon, c)$. Let $\hat{M}_2$ be any optimal size matching in this graph. Let $\mathcal{H}(M_2)$ and $\mathcal{L}(M_2)$ be the H and L nodes in this matching. For every two nodes $u_1, u_2 \in \mathcal{L}(M_2)$ introduce an undirected edge between u_1 and u_2 if the edges (v_1, u_2) and (v_2, u_1) exist in the original graph D, where v_1 and v_2 are the matches of u_1 and u_2 with respect to the matching M_2. Each such edge (u_1, u_2) exists with probability p_L^2, and thus there exists a perfect or near-perfect matching between the nodes $\mathcal{L}(M_2)$. For each edge (u_1, u_2) in this matching the 4-way exchange $(u_1, v_1, u_2, v_2, u_1)$ exists. Thus we can create a 4-way exchanges which cover nodes $\mathcal{H}(M_2)$ and $\mathcal{L}(M_2)$. Consider the remaining nodes $\mathcal{L} \setminus \mathcal{L}(M_2)$. As before, for every two nodes u_1, u_2 there introduce an undirected edge if both of the directed edges (u_1, u_2) and (u_2, u_1) exist in the original graph D. Since the edge density in this random graph is p_L^2, there exists a nearly perfect matching and we can create two-way exchanges covering all or all but one nodes among $\mathcal{L} \setminus \mathcal{L}(M_2)$.

We have obtained a combination of two and four-way exchanges which cover $|\mathcal{H}(M_2)| + |\mathcal{L}|$ nodes. Since $|M_2| = \alpha(\upsilon, c)n + \upsilon n + o(n)$, we obtained the claimed inequality. ∎

Claims 7.1 and 7.2 together establish the first part of Theorem 4.5.

We now prove the second part of the theorem. As above, let M_2 be any maximal size matching between $\mathcal{H}$ and $\mathcal{L}$, where an edge between $u \in \mathcal{L}$ and $v \in \mathcal{L}$ exists if there is a directed edge (u, v) in the original graph D. As before, the size of this matching is $\alpha(\upsilon, c)n + o(n)$ with high probability. Let $\mathcal{L}_0$ be the set of L nodes matched. Consider the partition of this set $\mathcal{L}_0 = \mathcal{L}_1 \cup \mathcal{L}_2$, where $\mathcal{L}_1$ is the set of nodes u such that the reverse edge (v, u) exists in the original graph D, where v is the match of u according to M_2. Let $\mathcal{L}_2$ be the complement of $\mathcal{L}_1$ in $\mathcal{L}_0$. The size of $\mathcal{L}_1$ is $\alpha(\upsilon, cp_L)n + o(n)$ with high probability. For every $u \in \mathcal{L}_2$ and its match $v \in \mathcal{H}$, it has to be the case that v is a part of a 3-way exchange $(u, v, w), w \in \mathcal{L} \setminus \mathcal{L}_0$, if v is to be a part of the exchange at all. Thus there must exist a matching between $\mathcal{L}_2$ and a subset of nodes $\mathcal{L} \setminus \mathcal{L}_0$. Such matching cannot exist if the size of $\mathcal{L}_2$ which is asymptotically $\alpha(\upsilon, c(1 - p_L)n + o(n)$ is larger than the size of $\mathcal{L} \setminus \mathcal{L}_0$ which is $\upsilon n - \alpha(\upsilon, c)n$. This establishes the second part of the claim. For the first part we observe that if the situation is reversed than a perfect or nearly matching exists, since the likelihood of a cycle (u, v, w) is p_L^2 and we again invoke the perfect matching existence property of random dense graphs. ∎

Size	Alts	$(2,3)$	$(3,3)$	$(3,\infty)$	(∞,∞)
50	0	8.24 (1.73)	11.31 (2.77)	11.31 (2.77)	13.85
	1	9.78 (1.87)	12.56 (2.81)	14.17 (3.31)	14.88
	2	11.83 (2.27)	14.61 (3.22)	16.53 (3.83)	16.69
	3	(13.16 (2.73)	15.55 (3.32)	17.36 (3.98)	17.47
	4	14.72 (2.85)	16.57 (3.67)	17.99 (4.19)	18.03)
	5	15.89 (3.35)	17.67 (4.05)	18.89 (4.4)	18.91
100	0	22 (5)	28.8 (7)	28.8 (7)	33.6
	1	26.8 (7)	37.8 (10.4)	42.2 (11.8)	42.6
	2	25.6 (6.6)	33.2 (8.8)	37.6 (10.6)	38.6
	3	25.4 (3.6)	33.6 (7.8)	39 (10.8)	39
	4	30.8 (5.6)	38.4 (9)	43.4 (11.4)	43.4
	5	31.8 (8.4)	40.2 (12)	44 (13.6)	44
150	0	35.34 (8.59)	50.39 (14.88)	50.39 (14.88)	58.65
	1	36.22 (8.29)	51.56 (14.68)	58.88 (18)	60.08
	2	39.36 (9.3)	54.64 (15.95)	62.46 (19.65)	62.72
	3	40.83 (9.43)	55.54 (15.68)	63.24 (19.32)	63.34
	4	44 (10.03)	58.93 (16.4)	66.04 (19.97)	66.09
	5	45.48 (10.58)	59.24 (16.48)	66 (19.92)	66.02
200	0	48.3 (11.9)	70.32 (21.48)	70.32 (21.48)	81.7
	1	50.84 (12.09)	73.34 (22.01)	83.4 (26.88)	84.67
	2	54.28 (12.79)	76.03 (22.46)	86.66 (27.84)	87.02
	3	55.48 (13.02)	76.88 (22.61)	87.54 (27.94)	87.85
	4	58.6 (13.94)	80.27 (23.93)	90.38 (29.1)	90.49
	5	60.66 (14.49)	81.83 (24.14)	91.52 (29.5)	91.58
250	0	63.56 (16.24)	93 (29.3)	93 (29.3)	107.2
	1	66.97 (16.4)	96.32 (29.48)	107.84 (35.28)	110.36
	2	68.02 (16.77)	96.34 (29.72)	109.78 (36.82)	110.53
	3	71.6 (17.48)	100.96 (30.88)	113.88 (37.84)	114.2
	4	73.72 (17.36)	102.22 (30.65)	114.94 (37.16)	115.09
	5	73.97 (17.84)	101.54 (30.9)	113.96 (37.15)	114.06

Table 6: Size of allocations with different number of altruistic donors (Alts) and different size pools. The number of matched high PRA patients is given in parenthesis.