Design and analysis of multi-hospital kidney exchange mechanisms using random graphs ☆

Panos Toulis [a,*], David C. Parkes [b]

[a] *Department of Statistics, Harvard University, 1 Oxford St, Cambridge, MA 02138, USA*
[b] *SEAS, Harvard University, 33 Oxford St, Cambridge, MA 02138, USA*

A B S T R A C T

Kidney exchanges enable transplants when a pair of a patient and an incompatible donor is matched with other similar pairs. In multi-hospital kidney exchanges pairs are pooled from multiple hospitals, and each hospital is able to decide which pairs to report and which to hide and match locally. Modeling the problem as a maximum matching on a random graph, we first establish that the expected benefit from pooling scales as the square-root of the number of pairs in each hospital. We design the xCM mechanism, which achieves efficiency and incentivizes hospitals of moderate-to-large size to fully report their pairs. Reciprocal pairs are crucial in the design, with the probabilistic uniform rule used to ensure incentive alignment. By grouping certain pair types into so-called virtual-reciprocal pairs, xCM extends to handle 3-cycles. We validate the performance of xCM in simulation, demonstrating its efficiency and incentive advantages over the Bonus mechanism (Ashlagi and Roth, 2014).

1. Introduction

The scarcity of cadaver kidneys coupled with the significant medical benefits from live kidney transplants has prompted the advance of kidney-paired donation (KPD), also referred to as *kidney exchanges*, in recent years (Montgomery et al., 2005). In KPD, a patient with an incompatible donor can form a patient–donor pair and, by entering a kidney exchange program, match with one or more other pairs, so that the patient receives a compatible kidney and the donor donates a kidney to some other patient. In the simplest case this match occurs through a 2-cycle (or swap) involving the transplant of two kidneys. Longer cycles, and in particular 3-cycles, are also practical but beyond 3-cycles the logistics of the simultaneous operations become difficult.[1]

[1] Simultaneous operations are required for ethical reasons because, for instance, it is possible that a donor could otherwise give up a kidney without the paired patient receiving a kidney. Since 2007, a growing practice is to adopt *non-simultaneous altruistic donor chains* in which a donor with no designated recipient can initiate a chain of transplants. Chains can be longer than cycles because transplants typically take place in an asynchronous manner without the

Kidney exchange programs exist around the world (Saidman et al., 2006), although their large-scale expansion has been hindered by ethical, logistical and incentive issues (Ross and Woodle, 2000; Delmonico et al., 2002). In the U.S., there is a growing number of multi-hospital exchanges which seek to pool the patient–donor lists from multiple hospitals in order to facilitate thicker markets and identify additional transplants (Rees et al., 2008), providing patients with access to larger pools of paired donations.[2] Each patient–donor pair is associated with a hospital, and hospitals can choose to share lists of pairs with a multi-hospital exchange.

The matching problem in a kidney exchange can be modeled as a *compatibility graph*, with a vertex corresponding to a patient–donor pair and edges representing compatibility between pairs. For 2-cycles, the edges are undirected and indicate mutual compatibility between pairs. We take a welfare maximizing outcome to be a maximum cardinality matching.[3] If 3-cycles are also possible, a directed graph is adopted, with a directed edge from pair u to v indicating that the donor in pair u is compatible with the patient in pair v. In this setting, we take a maximum cardinality, vertex-disjoint set of directed 2- or 3-cycles to be welfare maximizing.

We model the matching problem as one of finding maximum matchings on a random graph. This random graph arises from the blood-type and tissue type compatibility relationships of patient–donor pairs sampled from a population (see also Ashlagi and Roth, 2014). Characterizing the structural properties of maximum matchings on this random graph provides the probabilistic foundation for the analysis of multi-hospital exchange mechanisms. We first quantify the benefit of multi-hospital exchanges through a *square-root law*. This expresses the expected increase in the number of transplants that arises from pooling patient–donor lists in terms of the square-root of the size of each hospital's list. This analytical result is established for 2-cycles and a uniform tissue-type compatibility model (the same tissue-type compatibility model is also adopted in Abraham et al., 2007; Roth et al., 2007; Ünver, 2010). The square-root law explains simulation results in earlier papers (Abraham et al., 2007; Roth et al., 2007).

A multi-hospital exchange mechanism receives hospitals' reports of lists of patient–donor pairs, and determines a matching on reported pairs. An efficient mechanism achieves a maximum cardinality matching, either restricted to 2-cycles or 3-cycles depending on design constraints. A known challenge is to align incentives, so that each hospital reports its complete list of patient–donor pairs. We assume that each hospital seeks to maximize the number of pairs on its own list that are matched. In a poorly designed mechanism, a hospital may benefit by revealing only hard-to-match pairs to the exchange and free-riding on the reports of others, while completing other matches on pairs that it chooses not to report (Roth et al., 2005; Ashlagi and Roth, 2014).

We design an efficient and *ex post* incentive compatible multi-hospital exchange mechanism which we term xCM. The xCM mechanism achieves maximum cardinality matchings by allowing 2-cycles on pairs pooled from multiple hospitals where the pairs are specific donor–patient pair types known as *reciprocal*, and uses the *probabilistic uniform allocation rule* (Sasaki, 1997; Sprumont, 1991; Moulin, 2002; Ehlers and Klaus, 2003) to align incentives and ensure that no hospital can gain benefit from misreports. In order to enable maximum cardinality matchings in an exchange with 3-cycles, xCM allows 3-cycles on pairs pooled from multiple hospitals that consist of a reciprocal pair from a first hospital and a *virtual-reciprocal* pair from a second hospital, this virtual-reciprocal pair consisting of two pairs from the second hospital.

The patient–donor list is modeled as private information to a hospital. Efficiency and incentive compatibility are established under technical properties on the maximum matchings associated with patient–donor lists. Given this, it is an *ex post* Nash equilibrium for every hospital to truthfully report its complete patient–donor list, whatever the exact lists of each hospital. This *ex post* incentive compatibility does not imply that the xCM mechanism is strategy-proof (with truthful reporting a dominant strategy), because truthful reporting is only a best-response when the reports of other hospitals satisfy these technical properties. Truthful reports will satisfy the properties (by assumption), but the properties need not hold for arbitrary reports.

In simulation, we demonstrate that the aforementioned technical properties on hospital donor–patient lists hold with high probability for a standard model of uniform tissue-type sensitivity (see Section 2) as long as each hospital's list is moderately sized with at least 30 pairs. We validate the incentive properties of xCM in simulation, and demonstrate that xCM also has robust incentive properties for hospitals with smaller lists.

In addition, we compare xCM with the Bonus mechanism of Ashlagi and Roth (2014) in an extensive simulation. In settings with 4 to 12 hospitals, we see a significant benefit to a hospital in Bonus to first compute a maximum matching on its pairs and report only pairs that it cannot match locally. Although this incentive for strategic under-reporting is less useful for systems with larger numbers of hospitals (consistent with the theoretical results in Ashlagi and Roth, 2014), it continues to be present for up to 30 hospitals. A theoretical analysis of the incentive properties of Bonus that explains the aforementioned simulation results is provided in Section 6.

---

need to coordinate multiple simultaneous operations (Rees et al., 2009; Dickerson et al., 2012b). We focus on transplants that occur through simultaneous cycles rather than chains. Extending our results to chains is left as an open problem for future work.

  [2] The largest pool in the U.S. is the National Kidney Register (NKR), which has hundreds of active donors and has facilitated more than 1000 kidney transplants. Large exchanges are operated by the Alliance for Paired Donation (APD) and the United Network for Organ Sharing (UNOS).

  [3] In practice, kidney exchange programs select weighted maximum matchings with the weights set by medical professionals taking into account factors such as age, gender, health status, geography and patient preferences (Bertsimas et al., 2013). Other notable approaches include dynamic matching that take into account the randomness of arrival and expiration times in the transplant waiting list (Ünver, 2010; Dickerson et al., 2012a).

To compare xCM and Bonus, we evaluate the cardinality of matchings in both mechanisms and across various environments. Adopting truthful reporting in xCM and under-reporting in Bonus, we find that Bonus is less efficient than xCM. For xCM, the average number of pairs matched is in the range 93–99% and 88–97% of the cardinality of the possible maximum matching, for uniform and non-uniform tissue-type sensitivity, respectively. This performance range considers settings that allow for 3-cycles, and settings that insist on only 2-cycles. The comparable ranges of cardinality of matchings in Bonus are 82–95% and 66–86%.

1.1. Outline

Section 2 defines the random graph model of kidney exchanges and the relevant technical properties of maximum matchings on such random graphs, also stating a theoretical result in regard to existence and structure of certain maximum matchings. This section also defines the probabilistic uniform allocation rule, and reviews related work including the Bonus mechanism. Section 3 gives the square-root law. Section 4 defines the xCM mechanism for 2-cycles (only), and states the first main result in regard to the efficiency and incentive compatibility properties of xCM. Section 5 defines the xCM mechanism for 3-cycles, generalizing the technical properties and stating the second main result in regard to efficiency and incentive compatibility. Section 6 develops a theoretical analysis to isolate some concerns in regard to the incentive properties of Bonus (Ashlagi and Roth, 2014). Section 7 presents the simulation results in regard to efficiency and robustness of incentive compatibility of xCM. Most proofs are presented in Appendix A, and some additional analysis is deferred to supplementary material.

2. Preliminaries

For a kidney transplant to be possible, the donor and patient must be both blood-type and tissue-type compatible. Human blood-type is one of O, A, B and AB, and indicates the presence of proteins A and B (e.g., O-type has neither, AB-type has both). A patient–donor pair is blood-type compatible if the patient's blood includes every protein that is present in the blood of the donor. For example, a donor with blood-type A is blood-type compatible with a patient with type A or AB but not with a patient of type O or B.

By convention, a patient–donor *pair* is denoted by X–Y, where X is the blood-type of the patient and Y is the blood-type of the donor. A patient–donor list of a hospital consists of multiple such pairs, and it is these pairs that are matched on cycles. Following Ünver (2010), we associate pairs with one of four *pair-types*: *under-demanded* (UD), *over-demanded* (OD), *reciprocal* (R) and *self-demanded* (S), based on the blood-types of the pair. Intuitively, OD pairs such as A–O are relatively easy to match (hence "over-demanded") because the donor's blood-type contains fewer proteins than the patient's blood-type.

In regard to tissue-type compatibility, this requires that a patient and donor share as many *human leukocyte antigens* (HLAs) as possible to prevent a positive cross-match and organ rejection. The probability of a positive cross-match (and thus tissue-type incompatibility) depends on the tissue-type sensitivity of the patient. The *Panel Reactive Antibody* (PRA) sensitivity of a patient measures the percentage of the population with whom the patient will be tissue-type incompatible. Higher PRA values result in a higher probability of incompatibility. Following Zenios et al. (2001), we consider two models for PRA sensitivity. In the *uniform PRA model*, the sensitivity is assumed to be the same for all patients. In the *non-uniform PRA model*, a patient can be in one of three sensitivity groups, namely *low, medium* and *high*, each associated with a different probability of a positive cross-match.

2.1. The random graph model

Given a set of pairs and tissue-type and blood-type incompatibility relationships, this defines a *compatibility graph*. Focusing first on exchanges that involve only 2-cycles, each patient–donor pair is associated with a vertex, and an undirected edge between two vertices u, v indicates that the donor in u is compatible with the patient in v and vice versa. A matching on an undirected graph is a set of edges with no common vertices. The size (or cardinality) of a matching is the number of vertices incident to edges in the matching. A maximum matching is a matching of maximum cardinality. A matching is *perfect* if it includes every vertex. A matching is *almost-perfect* if it includes all but one vertex.

A distribution on blood-types and tissue-type compatibility relationships defines a distribution on compatibility graphs. Let $\widetilde{G}_n$ denote a random, undirected compatibility graph with n pairs. Let V denote the set of vertices. For pair $u \in V$, we adopt u_p and u_d to denote the patient and donor in the pair, respectively. Let predicate $ABO(x, y)$ take on value *true* when patient x and donor y are blood-type compatible, and *false* otherwise. Let $p_u \in [0, 1]$ denote the PRA sensitivity of the patient in pair u. A random graph $\widetilde{G}_n$ is constructed through a generative process. First, vertices are introduced by:

- independently sampling $u_p, u_d \sim F_{ABO}$, where F_{ABO} is the blood-type distribution in the population,
- sampling $p_u \sim F_{PRA}$ of the patient, where F_{PRA} is the PRA distribution in the population, and
- introducing pair $u = (u_p, u_d)$ if (i) not $ABO(u_p, u_d)$, or (ii) with probability p_u if $ABO(u_p, u_d)$.

Once n vertices have been introduced, then for every pair of vertices $u, v \in V$, an edge is introduced if the pairs are mutually compatible, which requires (i) $ABO(v_p, u_d)$ and $ABO(u_p, v_d)$, and (ii) tissue-type compatibility, which requires that
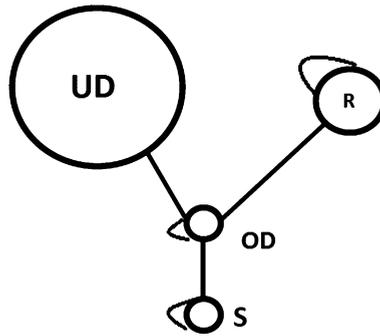
Fig. 1. A high-level depiction of the compatibility graph for different pair-types. An edge indicates the possibility of mutual, blood-type compatibility between two pair-types using only 2-cycles.

Table 1
Patient–donor pairs grouped by pair-type into under-demanded (UD), over-demanded (OD), reciprocal (R) and self-demanded (S) pairs.

Patient	Donor			
	O	A	B	AB
O	S	UD	UD	UD
A	OD	S	R	UD
B	OD	R	S	UD
AB	OD	OD	OD	S

a Bernoulli trial with probability $(1 - p_u)(1 - p_v)$ is a success, so that each patient is tissue-type compatible with the other patient's donor.

In the non-uniform PRA model, the pairs that enter tend to include patients that are *highly-sensitized* and hard to match. This leads to a more sparse compatibility graph. Following Zenios et al. (2001), we take $p_u = 0.2$ in the uniform PRA model. In the non-uniform model, we associate PRA sensitivity of 0.05, 0.45 and 0.9 with low, medium and high PRA sensitivity, and adopt probability 0.7, 0.2 and 0.1 for a patient to belong to the low, medium or high PRA group respectively.

Let q_x denote the probability that a random pair in $\widetilde{G_n}$ is of pair-type x, where $x \in \{OD, UD, S, R\}$. Assuming the uniform PRA model, and with a blood-type distribution representative of the worldwide population and tissue-type compatibility statistics (Zenios et al., 2001), simple algebra yields the following proportions[4]:

$$q_{UD} = 0.56, \qquad q_{OD} = 0.11, \qquad q_S = 0.15, \qquad q_R = 0.18. \tag{1}$$

Fig. 1 illustrates the high-level structure of a compatibility graph, with each circle corresponding to a pair-type and drawn in proportion to the expected frequency of different types in the exchange.[5] An edge between two pair-types or a self-edge indicates that a 2-cycle is possible between the pair-types incident to the edge based on blood-type compatibility. For example, an OD pair may be mutually compatible with any other pair-type while UD pairs cannot match with each other.[6]

For matchings that can include 3-cycles, a *directed compatibility graph* is used to model the matching problem. An edge from pair u to pair v indicates that the donor in pair u is blood-type compatible with the patient in pair v. Let $\widetilde{G_n^3}$ denote a random, directed compatibility graph with n pairs. The vertices are generated in the same way as for the undirected graph $\widetilde{G_n}$. For every pair of vertices $u, v \in V$, a directed edge from u to v requires (i) $ABO(v_p, u_d)$, and (ii) directed tissue-type compatibility, which requires the success of a Bernoulli trial with probability $(1 - p_v)$ so that the patient in pair v is tissue-type compatible with the donor in pair u.

---

[4] The worldwide blood-type distribution is roughly O (50%), A (30%), B (15%) and AB (5%) (Wikipedia, 2010). Let f_x denote the probability of pair-type x in the population. We would expect $f_{OD} = f_{UD} = 0.2725$, $f_S = 0.365$, and $f_R = 0.09$. With $p_c = 0.2$ to denote the PRA sensitivity, we have $q_{OD} = \frac{f_{OD} \cdot p_c}{Z}$, $q_S = \frac{f_S \cdot p_c}{Z}$, $q_{UD} = \frac{f_{UD}}{Z}$, and $q_R = \frac{f_R}{Z}$, where $f_{OD} = f_O(f_A + f_B) + f_{AB}(f_O + f_A + f_B)$, $f_{UD} = f_{OD}$, $f_S = f_O^2 + f_A^2 + f_B^2 + f_{AB}^2$, $f_R = 2f_A \cdot f_B$, and denominator $Z = f_{OD} \cdot p_c + f_S \cdot p_c + f_{UD} + f_R$ provides normalization.

[5] We see that UD pair-types are much more abundant than other types, and that OD pair-types are approximately $1/p_c = 5$ times less frequent than UD pair-types. This is because, an OD pair-type "enters" $\widetilde{G_n}$ only because of tissue-type incompatibility, which happens with probability p_c (under the uniform PRA model).

[6] Similarly, an S pair can only be mutually compatible with another S pair or with an OD pair. The absence of an edge between UD and R shows that there can be no edges between vertices that correspond to these two pair-types. These relationships can be checked from the requirement that every protein in a donor's blood is present in a patient's blood, and considering the mapping from blood-types into pair-types in Table 1.

2.2. Structural properties

Given an undirected compatibility graph, we can define a subgraph for each pair-type (OD, UD, S and R), consisting of the vertices of this pair-type and edges between these vertices. The S-subgraph has four components, comprised by T–T pairs for each blood-type T ∈ {O, A, B, AB}. The R-subgraph is bipartite, with one part consisting of A–B pairs and one of B–A pairs. The *long-side* of the R-subgraph corresponds to the pair-type with more pairs (breaking ties at random); the other part is the *short-side*. The R-subgraph is *balanced* if each part has the same cardinality, and *almost-balanced* if the two sides differ in cardinality by one. The UD-subgraph has no internal edges.

Definition 1 *(Perfect-Matching (PM) properties).* An undirected compatibility graph, G, is:

- *S-perfect* if, for each blood-type T, there is a perfect matching for any graph G' formed by retaining all T–T pairs of G when there is an even number, and dropping any single T–T pair when there is an odd number.
- *R-perfect* if there is a perfect matching for any graph G' formed by retaining all the R pairs on the short-side of the R-subgraph of G along with any subset of R pairs on the long-side of cardinality equal to that of the short-side.
- *OD/UD-perfect* if there exists a matching on G in which every OD pair is matched with some UD pair.

A *matching* on G is S-perfect, R-perfect, or OD/UD-perfect when the matching satisfies the associated perfect matching property.

Say that an edge is *internal* to a subgraph if it is incident to two vertices in the same subgraph.

Definition 2 *(Regular matching).* A *regular matching* on an undirected compatibility graph G is a matching that is OD/UD-perfect (matches every OD pair with some UD pair) and maximizes, amongst all matchings,
 (i) the number of matched edges internal to the S-subgraph, and
 (ii) the number of matched edges internal to the R-subgraph.

A matching is *almost-regular* if it is maximum, and maximizes the number of OD–UD pairs that match across all maximum matchings on G.

The simple structure of regular matchings is helpful for the analysis of the incentive properties of xCM, and because of existence and optimality:

Theorem 1. *A regular matching on an undirected compatibility graph G is maximum amongst all matchings that use only 2-cycles, and exists in random graph $\widetilde{G}_n$ w.h.p. for sufficiently large n.*

In regard to maximum cardinality, a similar result is stated as Proposition 1 in Roth et al. (2007). Existence of a regular matching follows as a special case of the more general Theorem 5 in a context with 3-cycles, and proved as Proposition 5.1 in Ashlagi and Roth (2014).

We provide a simple proof sketch of maximality. Let N_S, N_R, N_{OD}, N_{UD} denote the number of pairs matched in a regular matching (thus $N_{OD} = N_{UD}$). Any non-regular matching would match $k > 0$ OD pairs with either S or R pairs, and since UD pairs can be matched only to OD pairs, k fewer UD pairs would be matched. At best, the total pairs matched in that non-regular matching would be $(N_S + N_R + k) + N_{OD} + (N_{UD} - k)$, which is no better than the number of pairs matched in the regular matching.

For existence, let $\widetilde{G}_{m \times m, p}$ denote a random bipartite graph with m nodes in each part, and edges between nodes in different parts with uniform probability p. Erdos and Rényi (1964) proved the following:

Lemma 1. *(See Erdos and Rényi, 1964.) A random bipartite graph $\widetilde{G}_{m \times m, p}$, with any constant $p > 0$, assumes a perfect matching w.h.p. for sufficiently large m.*

Consider some UD pair with blood-type X–Y (patient X, donor Y). The symmetric OD pair is Y–X; e.g., consider UD pair O–A and OD pair A–O. Because OD pairs only enter when the patient is tissue-type incompatible with the donor while UD pairs always enter, OD pairs are less abundant than UD pairs, w.h.p. for large n. Based on this, we can appeal to Lemma 1 on the balanced subgraph formed for X–Y and Y–X pairs.

The following lemma confirms the existence in the limit of S-perfect and R-perfect matchings:

Lemma 2. *A random compatibility graph $\widetilde{G}_n$ on n vertices satisfies the S-perfect and R-perfect properties w.h.p. for sufficiently large n.*

The result follows by appeal to Lemma 1. For S-perfect, this follows by noting that the T–T components of the S-subgraph, for T ∈ {O, A, B, AB}, can be reduced to an almost-balanced bipartite graph by placing pairs into two, almost-balanced sets and dropping between-set edges. For R-perfect, this follows by considering a random, balanced subgraph.

2.3. Probabilistic uniform rule

The xCM mechanism uses the *probabilistic uniform rule* (Sasaki, 1997) as a subroutine. The probabilistic uniform rule generalizes the uniform rule for divisible goods (Sprumont, 1991) to provide a strategy-proof (i.e., dominant-strategy incentive compatible), Pareto efficient method to allocate indivisible, identical items.

In defining this rule, let $x > 0$ denote the number of identical, indivisible units of a resource to allocate. Consider agents $j \in \{1, \dots, m\}$, each of which demands some quantity $z_j \geq 0$ of items. Each agent strictly prefers more items to less up to z_j items.

Definition 3 *(Probabilistic uniform rule).* The *probabilistic uniform rule* takes demand $\vec{z} = (z_1, \dots, z_m)$ and supply x of identical indivisible items and returns an allocation $\vec{y} = (y_1, \dots, y_m) = share(z, x)$ as follows.

First, initialize $J := \{j : z_j > 0\}$, remaining supply $x_{rem} := x$ and allocation $y_j := 0$ for all $j \in \{1, \dots, m\}$. Then:

While $(x_{rem} > 0)$ *and* $(J \neq \emptyset)$:

- *If* $x_{rem} \geq |J|$, *increment* y_j *by one for each* $j \in J$, *decrement* x_{rem} *by* $|J|$, *and drop any* j *from* J *for which* $y_j \geq z_j$.
- *Else: select a random subset of size* x_{rem} *of agents in* J, *and increment* y_j *by one for each selected agent. Set* $x_{rem} := 0$.

Lemma 3. *Given supply* $x > 0$ *and demand* $z = (z_1, \dots, z_m)$, *the probabilistic uniform rule* $share(z, x)$ *allocates expected quantity* $y_h = \min(z_h, g_{\mathrm{uniform}}(z_{-h}, x)) \geq 0$ *to agent* h, *where* $g_{\mathrm{uniform}}(z_{-h}, x) \geq 0$ *is a quantity that depends on demands* $z_{-h} = (z_1, \dots, z_{h-1}, z_{h+1}, \dots, z_m)$ *of others and supply* x, *and satisfies properties:*

(1) $g_{\mathrm{uniform}}(z_{-h}, x + c) \leq g_{\mathrm{uniform}}(z_{-h}, x) + c$ *for all* z_{-h}, x *and* $c \geq 0$; *and*
(2) $g_{\mathrm{uniform}}(z_{-h}, x) \geq x - \sum_{j \neq h} z_j$.

Proof. Let $d(k) = \sum_{j \neq h} \min(z_j, k) + k$. This is the total demand when it is capped at k for all agents, and assuming agent h's demand is at least k. Let k^* denote the largest k for which $d(k) \leq x$. Let m' denote the number of agents $j \neq h$ for which $z_j > k^*$. Note that the number of unallocated items once up to k^* have been allocated is $x - d(k^*) < m' + 1$. All agents with demand $\leq k^*$ are allocated in full. Agents with demand $> k^*$ also receive one additional item with probability $\frac{x - d(k^*)}{m' + 1}$, since the subset allocated one of the remaining items is selected uniformly at random. Therefore, the expected quantity allocated to agent h is equal to $\min(z_h, g_{\mathrm{uniform}}(z_{-h}, x))$, and $g_{\mathrm{uniform}}(z_{-h}, x) = k^* + \frac{x - d(k^*)}{m' + 1}$. Property (1) holds because the allocation to agents $j \neq h$ cannot decrease when more units are introduced. Property (2) holds because no agent j is allocated more than z_j. □

2.4. Related work

Ashlagi and Roth (2014) also study the design of multiple-hospital exchanges under a random graph model. But while their analysis insists on a large number of hospitals, each of which has a small patient–donor list, our analysis insists that each hospital has a moderately-sized or large patient–donor list. In this sense the two designs complement each other.[7]

The xCM and Bonus mechanisms also differ in their design focus. xCM emphasizes the use of pooling to match R pairs, while Bonus emphasizes the use of pooling to match OD and UD pairs. We identify an incentive concern in regard to the use of a lottery in Bonus for matching OD and UD pairs. Although Ashlagi and Roth (2014) argue approximate Bayes–Nash incentive compatibility in the limit of a large number of hospitals each with a finite size list, our analysis suggests that Bonus remains susceptible to holding back OD pairs and sending UD pairs even with as many as 30 hospitals (see Section 6).[8,9] The xCM and Bonus mechanisms also differ fundamentally in the way that R pairs are matched, with xCM using the probabilistic uniform allocation rule.

Multi-hospital kidney exchanges have also been studied in a worst-case framework by Ashlagi et al. (2013), who consider a setting where matchings are restricted to 2-cycles, and prove that no deterministic strategy-proof mechanism can ensure an approximation ratio less than two (relative to the cardinality of the maximum matching) on every possible input. In addition, no randomized strategy-proof mechanism can provide an approximation ratio better than 8/7; see also Roth et al. (2005) and Ashlagi and Roth (2014) for more discussion. Ashlagi et al. (2013) also develop the MIX-AND-MATCH mechanism, which is a randomized, strategy-proof 2-approximation mechanism.

---

[7] The *strong regularity* assumption in Ashlagi and Roth (2014) is shown in their simulations to require each hospital to have no more than 30 pairs. The technical properties (perfect matching) that we require for the analysis of xCM are shown in simulation to hold for 4 or more hospitals, each of size 30 or larger (2-cycles or 3-cycles), under the uniform PRA model.

[8] Our simulation results for Bonus disagree with the simulation results of Ashlagi and Roth (2014), in that we find benefits from deviation in environments where they find none. We have been unable to explain this discrepancy because the code used by Ashlagi and Roth is not available (personal communication). On the other hand, our simulation results are still consistent with their broader theory, in that the benefit from deviation becomes less severe as the numbers of hospitals increases.

[9] In comparison, xCM only pools OD and UD pairs for the purpose of enabling 3-cycles that involve R-pairs, and in a last step on unmatched pairs.

3. A square-root law for the benefit from pooling

In this section, we quantify the welfare benefit from pooling patient–donor lists. Given random compatibility graph $\widetilde{G}_n$, let G_R and G_S denote the R- and S-subgraphs, respectively. We first quantify the expected cardinality of a maximum matching on these subgraphs.

Lemma 4. *Assuming the S-perfect property, the expected cardinality of a maximum matching M_S using 2-cycles on subgraph G_S of fixed size in compatibility graph $\widetilde{G}_n$, is given by*

$$E\big[|M_S|\big] = |G_S| - 2, \tag{2}$$

where $|G_S|$ is the fixed number of vertices in subgraph G_S, and $|M_S|$ is the size of matching M_S.

An informal argument is that the marginal distribution on the parity of the number of pairs in any T–T component (for blood-type T) of the S-subgraph places equal probability on odd and even parity. Based on this, the expected number of unmatched pairs per component is $0.5(0) + 0.5(1) = 0.5$, and with four components the expected number of unmatched pairs is $4(0.5) = 2$. A full proof must condition on the parity of $|G_S|$ and handle dependence between the parity of the components of G_S; see Lemma 4 in the supplementary material.

Lemma 5. *Assuming the R-perfect property, the expected cardinality of a maximum matching M_R using 2-cycles on subgraph G_R of fixed size in random compatibility graph $\widetilde{G}_n$, is given by*

$$E\big[|M_R|\big] = |G_R| - \sqrt{\frac{2|G_R|}{\pi}}, \tag{3}$$

where $|G_R|$ is the fixed number of vertices in subgraph G_R, and $|M_R|$ the size of matching M_R.

In effect, Lemma 5 quantifies the size of the expected imbalance between the long-side and short-side of the R-subgraph. A formal proof appears in the supplementary material.[10]

Theorem 2. *Assuming the S-perfect, R-perfect and OD/UD-perfect properties, the expected cardinality of a maximum matching M_n using 2-cycles on random compatibility graph $\widetilde{G}_n$ is,*

$$E\big[|M_n|\big] = \gamma \cdot n - r_n \sqrt{\frac{2q_R}{\pi} n} - 2, \quad \gamma = 1 - q_{UD} + q_{OD},$$

where q_R, q_{UD} and q_{OD} denote the probability of reciprocal, under-demanded and over-demanded pair-types in the exchange respectively, and $r_n \in (1 - O(1/n), 1]$.

The full proof is given in the supplementary material. To gain additional simplification, let's assume $r_n = 1$. Substituting values for q_{UD}, q_{OD} and q_R under the uniform PRA model (1), we obtain the following approximate expression for the expected cardinality of the maximum matching:

$$E\big[|M_n|\big] \approx 0.556n - 0.338\sqrt{n} - 2. \tag{4}$$

This analysis is accurate enough to explain many simulation results in previous work, and provides an initial validation of the PM properties.[11]

Furthermore, we can quantify the increase in the expected number of transplants from pooling the patient–donor pairs from multiple hospitals. For this, let $\mu(n)$ denote the expected cardinality of a maximum matching on a random graph $\widetilde{G}_n$. Let $W(m,n)$ denote the expected increase in number of transplants from pooling together m hospitals each of size n compared to matching each hospital individually. By symmetry, the expected gain to any one hospital h is simply $W_h(m,n) = \frac{1}{m} W(m,n)$.

---

[10] For some intuition, consider the simpler task of calculating the *square* of the imbalance between the count of heads and tails after k tosses of a fair coin. This imbalance models the unmatched side (A–B or B–A) in a maximum matching within the bipartite R-subgraph. Let D_k denote this random variable. Feynman (1964, Chapter 6) argues that $E[D_k^2] \approx k$ for large k, since $D_k = D_{k-1} \pm 1$ and so $D_k^2 = D_{k-1}^2 \pm 2D_{k-1} + 1$, and the expected value of D_{k-1} is converging to 0 as k grows. By Jensen's inequality, $E(D_k^2) \geq E(D_k)^2$ and thus $E(D_k) = O(\sqrt{k})$.

[11] For example, the cardinality of the matchings in Roth et al. (2007) (see Propositions 1 and 2 and Table 2), are explained; e.g., for $n = 100$, expression (4) yields an expected number of matched vertices of $55.6 - 3.38 - 2 = 50.22$, and close to the reported value of 49.7. Similarly, the results in Table 1 of Abraham et al. (2007) can be explained.

Theorem 3 *(Square-root law). Assuming the S-perfect, R-perfect and OD/UD-perfect properties for the pooled compatibility graph of m hospitals, each with patient–donor lists of size n, the expected benefit to an individual hospital h from pooling, when using 2-cycles, satisfies*

$$W_h(m, n) \geq \left(1 - \frac{1}{\sqrt{m}}\right)\sqrt{\frac{2q_R}{\pi}n} + 2\left(1 - \frac{1}{m}\right), \tag{5}$$

where q_R is the fraction of R pair-types in the exchange.

Proof. The expected number of matches for each hospital individually is at most $\mu(n)$, since this is the expected number in a regular match, which is maximum when it exists. By the PM assumptions, the expected number of matches in the combined graph is $\mu(mn)$. Therefore, $W(m, n) \geq \mu(mn) - m\mu(n)$, and the result follows from Theorem 2 and simple algebra. □

When every compatibility graph satisfies the S-perfect, R-perfect and OD/UD-perfect properties, the expression holds with equality. Also, as the number of hospitals $m \to \infty$, the individual expected benefit from pooling satisfies $\lim_{m \to \infty} W_h(m, n) \geq \sqrt{\frac{2q_R}{\pi}n} + 2$. This quantifies the benefits that are available from solving the incentive challenges in multi-hospital exchanges, and also pinpoints the R-subgraph as being of central importance for achieving efficiency gains. The design of xCM promotes pooling of reciprocal-type pairs.

4. The xCM mechanism with 2-cycles

Assume m hospitals each of size n pairs, let $h \in \{1, \ldots, m\}$ denote a hospital, and let G_h denote the compatibility graph for hospital h. This graph corresponds to the list of patient–donor pairs of a hospital, and is private information to a hospital (analogous to *type* in mechanism design.) Define a *marginal economy* to be any set of hospitals that includes every hospital except one.

Let s_h denote the *strategy* of hospital h, and determine its reported list of pairs. Equivalently, we can consider that a hospital reports a compatibility graph and adopt $G'_h = s_h(G_h)$ to denote the reported graph of hospital h. We assume that a hospital can hide pairs but can neither introduce false pairs or false information about the compatibility between pairs.[12] Let $\vec{G}' = (G'_1, \ldots, G'_m)$ denote the vector of reports, and let $G'_\oplus = \bigoplus_h G'_h$ denote the compatibility graph obtained by combining the reports in the natural way.[13]

A *multi-hospital kidney exchange mechanism* Γ takes reported graphs $\vec{G}'$ and computes a matching, denoted by $\Gamma(G'_1, \ldots, G'_m)$. We adopt $\Gamma_h(G'_1, \ldots, G'_m)$ to denote the pairs of hospital h that are matched. Let $s = (s_1, \ldots, s_m)$ denote the strategy profile, and adopt $s(\vec{G}) = (s_1(G_1), \ldots, s_m(G_m))$ to denote the vector of reported graphs.

By assumption, each hospital wants to maximize the number of its own pairs that are matched. In addition to the pairs matched by the mechanism, a hospital has a *recourse* action in which it can match pairs that are either returned unmatched by the mechanism or were not reported. Define the *utility* $u_h(s, \vec{G})$ to hospital h, given strategy profile s as the total number of pairs of the hospital that are either matched by the mechanism or matched during recourse. The utility is,

$$u_h(s, \vec{G}) = \left|\Gamma_h(s(\vec{G}))\right| + \left|M(G_h \setminus \Gamma_h(s(\vec{G})))\right|, \tag{6}$$

where $M(G)$ is a maximum matching on graph G, and $G_h \setminus G'$ denotes the graph after removing the vertices G' from graph G_h.

Definition 4 *(EPIC). A matching mechanism Γ is ex post incentive compatible (EPIC) for property P if, for any vector of graphs $\vec{G} = (G_1, \ldots, G_m)$ that satisfy property P, it is a best-response for every hospital to report its true patient–donor list when other hospitals are truthful.*

For randomized mechanisms, EPIC for property P requires that each hospital maximizes its expected number of matches with respect to the random draws of the mechanism by reporting truthfully. For xCM, property P is defined to require that S-perfect and R-perfect holds for the compatibility graph that corresponds to the marginal economy without any hospital.[14]

---

[12] This is consistent with previous work. Ashlagi and Roth (2014) note that the information contained in blood tests is becoming standardized, sometimes coming from a centralized testing facility. In that respect, a hospital's strategy amounts to reporting a subset of its patient–donor pairs (vertices in the hospital graph), as the edges can be recovered given a known set of vertices.

[13] Include all vertices and edges reported by hospitals, and introduce edges between mutually-compatible pairs from distinct hospitals according to the random process defined in Section 2.1.

[14] EPIC with property P does not imply that a mechanism is dominant-strategy incentive compatible. This is because property P may not hold for arbitrary reports from other hospitals, while to establish EPIC it suffices to show that truthful reporting is optimal for a hospital given that reports of others satisfy property P.

Definition 5 *(k-Way efficient).* A matching mechanism Γ is *k-way efficient* if the matching $\Gamma(\vec{G})$ is maximum for any profile of graphs $\vec{G} = (G_1, \ldots, G_m)$ amongst matchings that use cycles of length at most k.

4.1. The xCM *mechanism*

In the following, let α_h, β_h and τ_h denote the number of A–B, B–A and T–T pair-types (for some blood-type T) in the patient–donor list of hospital h. Let $\alpha_\oplus$, $\beta_\oplus$ and $\tau_\oplus$ denote the corresponding counts when summed over all hospitals. Notation α'_h, β'_h and so forth indicates a count based on reported lists.

Definition 6 *(xCM mechanism).* The xCM mechanism receives a vector of graphs $\vec{G}' = (G'_1, \ldots, G'_m)$. Let $G'_\oplus$ denote the combined graph. Initialize matching μ_{xCM} to the empty matching. Matching μ_{xCM} is output at the end of Step 3.

Step 1. (Match S pairs) *Repeat for each self-demanded pair-type T–T, where* T $\in$ {O, A, B, AB}:
  (i) For each hospital h, compute a maximum matching $M'_{h,T}$ on the subgraph of G'_h induced by T–T pairs.
  (ii) Let G'_T denote the subgraph of $G'_\oplus$ induced by T–T pairs. Select matching μ_T uniformly at random from the set of matchings on G'_T that are maximum amongst those that satisfy,

$$\big\{\mu : N_T(\mu, h) \geq \big|M'_{h,T}\big|, \ \forall h\big\}, \tag{7}$$

  where $N_T(\mu, h)$ is the number of T–T pairs of hospital h in μ.
  (iii) $\mu_{\text{xCM}} := \mu_{\text{xCM}} \cup \mu_T$.
Step 2. (Match R pairs)
  (i) For each hospital h, let $m^*_h = \min(\alpha'_h, \beta'_h)$, and $z_{h,\text{AB}} = \alpha'_h - m^*_h$, $z_{h,\text{BA}} = \beta'_h - m^*_h$ denote the excess A–B and B–A pairs, respectively. Let $x_{\text{AB}} = \sum_j z_{j,\text{AB}}$ and $x_{\text{BA}} = \sum_j z_{j,\text{BA}}$.
  (ii) If $(\alpha'_\oplus \geq \beta'_\oplus)$ then: $y_{\text{AB}} : share(z_{\text{AB}}, x_{\text{BA}})$, $y_{\text{BA}} := 0$. Else $y_{\text{AB}} := 0$, $y_{\text{BA}} := share(z_{\text{BA}}, x_{\text{AB}})$.
  (iii) For each hospital h, compute a maximum matching on $G'_{h,R}$, and let m_h denote the number of A–B pairs (and also B–A pairs) matched. Let $\delta_h = m^*_h - m_h$. Let G'_R denote the subgraph of $G'_\oplus$ induced by R pairs. Let K_q, for $q = 0, 1, \ldots$, denote the set of matchings on G'_R that are maximum amongst those that satisfy,

$$\big\{\mu : N_{\text{AB}}(\mu, h) \geq m_h + \max(0, y_{h,\text{AB}} + \delta_h - q), \ \forall h, \text{ and}$$
$$N_{\text{BA}}(\mu, h) \geq m_h + \max(0, y_{h,\text{BA}} + \delta_h - q), \ \forall h\big\}, \tag{8}$$

  where $N_{\text{AB}}(\mu, h)$ and $N_{\text{BA}}(\mu, h)$ are the number of A–B and B–A pairs of hospital h in matching μ. Let q^* denote the smallest q for which K_q is non-empty. Select matching μ_R uniformly at random from K_{q^*}.
  (iv) $\mu_{\text{xCM}} := \mu_{\text{xCM}} \cup \mu_R$.
Step 3. (Match remaining same-hospital pairs) *Repeat for each hospital* h:
  (i) Compute an almost-regular matching μ'_h on the graph G''_h formed by removing the pairs in $\mu_{\text{xCM}} \cap G'_h$ from G'_h.
  (ii) $\mu_{\text{xCM}} := \mu_{\text{xCM}} \cup \mu'_h$.
Step 4. (Match remaining pairs) Compute a random, almost-regular matching on the combined graph formed from any remaining unmatched pairs. Add this matching μ_{xCM}.[15]

In Step 1, the xCM mechanism computes a maximum matching on the pooled S-subgraph, making sure that each hospital does as well as it would if matching internally. In Step 2, the uniform probabilistic rule is used to determine a matching on the pooled R-subgraph. In Step 3, this matching is augmented with an almost-regular maximum matching for each hospital on any unmatched, reported pairs (without any additional pooling). In Step 4, this matching is augmented with an almost-regular maximum matching that allows for pooling of unmatched pairs across hospitals.

The quantity m^*_h represents the maximum number of A–B pairs h could possibly expect to match on its own R-subgraph. If the pool is long A–B, then xCM assigns a target number of A–B pairs (equal to $m_h + y_{h,\text{AB}} + \delta_h = m^*_h + y_{h,\text{AB}}$) to each hospital h. xCM relaxes this target if it is not attainable, but always insists that each hospital gains at least the minimal number of matches it would achieve internally.[16]

Theorem 4. *The* xCM *mechanism is EPIC and 2-way efficient for properties (i) S-perfect and R-perfect on compatibility graphs the size of every marginal economy and larger, and (ii) OD/UD-perfect on every hospital's compatibility graph.*

The proof of EPIC and efficiency for xCM under properties S-perfect, R-perfect and OD/UD-perfect is given in Appendix A. An important step in the proof of EPIC is to use the properties of the probabilistic uniform allocation rule to establish that

---

[15] This step has no effect when the PM and regularity assumptions hold, but improves welfare when xCM is used in an environment where these properties fail to hold.
[16] In Section 7 we explain how this procedure can be efficiently implemented.

a hospital cannot do better in expectation by misreporting its full list of reciprocal patient–donor pairs. Once EPIC for these PM properties is established, then efficiency follows because the overall matching is regular, since it is maximum on the combined S- and R-subgraphs, and matches every OD pair with a UD pair.

5. Generalizing xCM to allow for 3-cycles

In this section, we generalize the xCM mechanism to allow for 3-cycles. The generalized mechanism remains EPIC for generalized PM properties to allow for 3-cycles, and is 3-way efficient. The increase in welfare from using 3-cycles rather than 2-cycles relies on 3-cycles involving OD, UD and R pairs because these can address the imbalance between the two sides of the R-subgraph.

Definition 7 *(Virtual-reciprocal pair).* Pairs (B–O, O–A) for which the donor of the OD pair is tissue-type compatible with the patient of the UD pair forms a *virtual-B–A* pair. Pairs (A–O, O–B) for which the donor of the OD pair is tissue-type compatible with the patient of the UD pair forms a *virtual-A–B* pair.

A virtual-reciprocal pair can form a cycle with a compatible reciprocal pair: B–O → O–A → A–B and A–O → O–B → B–A.

Define the *number of virtual-B–A pairs* (denoted v_β) as the number of B–O pairs matched in a maximum matching on a particular undirected bipartite graph. The bipartite graph is formed from a graph G by B–O pairs in one part and O–A pairs in the other, and an edge between a B–O and O–A pair when the O donor and O patient are tissue-type compatible. The *number of virtual-A–B pairs* (denoted v_α) is defined in an analogous way through a maximum matching on a graph with A–O and O–B pairs.

Other 3-cycles of interest for maximum cardinality matchings involve the OD pair with blood-types AB–O, since this can form the following 3-cycles with two UD pairs: AB–O → O–A → A–AB, and AB–O → O–B → B–AB. Another 3-cycle of interest involve S pairs; for example, A–A → A–A → A–A.

Definition 8 *(Generalized Perfect-Matching (PM) properties).* Consider a directed compatibility graph G. Let n_s and n_l denote the number of R pairs on the short-side and long-side of the R-subgraph of G respectively, and v_s the number of virtual-R pairs on the short-side (i.e., virtual-A–B if A–B pairs are on the short-side). Compatibility graph G is:

- *3S-perfect* if, for each blood-type T, there is a perfect matching (allowing for 2-cycles and 3-cycles) on graph G' formed by retaining all T–T pairs of graph G.
- *3R-perfect* if (i) when $n_s + v_s \geq n_l$, there is a matching (allowing for 2-cycles and 3-cycles) that matches every R pair on graph G' that is formed by retaining all R pairs and virtual-R pairs of graph G, or
  (ii) when $n_s + v_s < n_l$, there is a matching (allowing for 2-cycles and 3-cycles) on graph G' that matches every R pair, where G' is formed from G by retaining all virtual-R pairs, all R pairs on the short-side, and any subset of cardinality $n_s + v_s$ of R pairs on the long-side.
- *3OD/UD-perfect* if (i) there is a matching on G in which every AB–O pair matches with two UD pairs using 3-cycles, and (ii) having removed any number of virtual-R pairs and all 3-cycles involving AB–O pairs and two UD pairs from G, there exists a matching using 2-cycles where every remaining OD pair matches with a UD pair.

A graph is *strong 3R-perfect* if the graph meets case (i) of 3R-perfect, and matches every R pair. A matching is 3S-perfect if the matching attains the structure of the associated graph property (similarly for 3R-perfect). A matching is 3OD/UD-perfect if every AB–O pair matches on a 3-cycle with two UD pairs, and every other OD pair not used in a virtual-R pair matches in a 2-cycle with a UD pair.

Definition 9 *(Extended-R subgraph).* The extended-R subgraph of a directed compatibility graph G includes:

1. all A–B, B–A, B–O, O–A, A–O and O–B pairs,
2. all edges that may be used when matching virtual-R pairs in 3-cycles with R-pairs and matching R-pairs in 2-cycles with R-pairs.

For example, the extended R-subgraph includes compatibility edges from B–O to O–A pairs when they exist, and from O–A to A–B pairs when they exist.

Definition 10 *(3-Way regular matching).* A *3-way regular matching* on a directed compatibility graph G is a matching that uses 2-cycles and 3-cycles, is 3OD/UD-perfect, and maximizes, amongst all matchings,
   (i) the number of matched edges internal to the S-subgraph, and
   (ii) the number of matched edges internal to the extended-R subgraph.

A *strong 3-way regular* matching is a 3-way regular matching in which every R pair is matched. A matching is *3-way almost-regular* if it is maximum and maximizes the number of OD–UD 2-cycles matched across all maximum matchings.

Theorem 5. *(See Ashlagi and Roth, 2014; Toulis and Parkes, 2011.) A strong 3-way regular matching on a directed compatibility graph is 3-way efficient, and exists in $\widetilde{G}_n^3$ w.h.p. for sufficiently large n.*

The existence and efficiency of strong 3-way regular matchings was proved by Ashlagi and Roth (2014) (Proposition 5.1). The most important change in the design of xCM in allowing for 3-cycles is that virtual-R pairs can be matched in Step 2. But this is restricted to virtual-R pairs in which the OD and UD pairs belong to the same hospital.

Given this, we define the *multi-hospital* extended-R subgraph in the natural way on the combined graph from all hospitals, so that it only includes edges B–O to O–A and A–O to O–B when these pairs belong to the same hospital. Similarly, *multi-hospital 3R-perfect* (and *multi-hospital strong 3R-perfect*) is defined in the natural way on the combined graph, with virtual-R pairs (and the count of these pairs) restricted to insist on the OD and UD pairs being from the same hospital.

Following our convention, let $v\alpha_h$, $v\beta_h$, $v\alpha_\oplus$, $v\beta_\oplus$ denote the associated counts of same-hospital virtual-A–B and virtual-B–A pairs for truthful reports, with $v\alpha'_h$, $v\alpha'_\oplus$ and so forth to denote the associated counts after misreports.

Definition 11 *(3-xCM mechanism).* The 3-xCM mechanism receives graphs $\vec{G}' = (G'_1, \dots G'_m)$ from each hospital. Let $G'_\oplus$ denote the combined graph. Initialize μ_{xCM} to the empty matching. Matching μ_{xCM} is output at the end of Step 3.

Step 0. (Match AB–O pairs)
  (i) For each hospital h, compute a maximum matching $\mu_{h,0}$ of G'_h restricted to the AB–O pairs and all UD pairs, and only using OD-UD-UD 3-cycles involving AB–O pairs.
  (ii) Update $\mu_{xCM} := \mu_{xCM} \cup \mu_{h,0}$, and remove any matched OD and UD pairs from $G'_\oplus$ and G'_h.
Step 1. (Match S pairs) Same as Step 1 in xCM, except allowing for 2-cycles and 3-cycles.
Step 2. (Match R pairs)
  (i) For each hospital h, let $m^*_{h,AB} = \min(\alpha'_h, \beta'_h + v\beta'_h)$ and $m^*_{h,BA} = \min(\beta'_h, \alpha'_h + v\alpha'_h)$. Let $z_{h,AB} = \alpha'_h - m^*_{h,AB}$ and $z_{h,BA} = \beta'_h - m^*_{h,BA}$ denote the number of excess A–B and B–A pairs, respectively. Let $z_{h,vAB} = v\alpha'_j - \max(0, m^*_{j,BA} - m^*_{j,AB})$ and $z_{h,vBA} = v\beta'_j - \max(0, m^*_{j,AB} - m^*_{j,BA})$ denote the number of excess virtual-AB and virtual-BA pairs, respectively.
  Let $x_{AB} = \sum_j(z_{j,AB} + z_{j,vAB})$ and $x_{BA} = \sum_j(z_{j,BA} + z_{j,vBA})$.
  (ii) If $(\alpha'_\oplus \geq \beta'_\oplus)$, then: $y_{AB} := share(z_{AB}, x_{AB})$, $y_{BA} := 0$. Else: $y_{AB} := 0$, $y_{BA} := share(z_{BA}, x_{AB})$.
  (iii) For each hospital h, compute a maximum matching on the extended-R subgraph in G'_h, maximizing the number of R pairs matched (breaking ties to minimize the number of 3-cycles). Let $m_{h,AB}$ and $m_{h,BA}$ denote the number of A–B and B–A pairs matched. Let $\delta_{h,AB} = m^*_{h,AB} - m_{h,AB}$ and $\delta_{h,BA} = m^*_{h,BA} - m_{h,BA}$.
  Let G'_R denote the multi-hospital extended-R subgraph of $G'_\oplus$. Let K_q, for $q = 0, 1, \dots$, denote the set of matchings on G'_R that maximize the number of R pairs matched (breaking ties to minimize the number of 3-cycles) amongst those that satisfy,

$$\{\mu : N_{AB}(\mu, h) \geq m_{h,AB} + \max(0, y_{h,AB} + \delta_{h,AB} - q), \ \forall h, \text{ and}$$
$$N_{BA}(\mu, h) \geq m_{h,BA} + \max(0, y_{h,BA} + \delta_{h,BA} - q), \ \forall h\}, \tag{9}$$

  where $N_{AB}(\mu, h)$ and $N_{BA}(\mu, h)$ are the number of A–B and B–A pairs of hospital h in matching μ. Let q^* denote the smallest q for which K_q is non-empty. Select matching μ_R uniformly at random from K_q.
  (iv) $\mu_{xCM} := \mu_{xCM} \cup \mu_R$.
Step 3. (Match remaining same-hospital pairs) *Repeat for each hospital h*:
  (i) Compute a 3-way almost-regular matching μ'_h on the graph G''_h formed by removing the pairs in $\mu_{xCM} \cap G'_h$ from G'_h.
  (ii) $\mu_{xCM} := \mu_{xCM} \cup \mu'_h$.
Step 4. (Match remaining pairs) Compute a random, 3-way, almost-regular matching on the combined graph formed from any remaining unmatched pairs. Add this matching μ_{xCM}.

Step 0 matches as many AB–O pairs on 3-cycles with UD pairs as possible, but without pooling pairs across hospitals. Step 1 matches S pairs, allowing for pooling across hospitals and using both 2- and 3-cycles. Step 2 generalizes Step 2 of xCM to allow for same-hospital virtual-R pairs. Steps 3 and 4 follow the same logic as Steps 3 and 4 for xCM.

The quantity $m^*_{h,AB}$ represents the maximum number of A–B pairs that hospital h can match on its own extended-R subgraph. If the pool is long A–B, then 3-xCM assigns a target number of A–B pairs (equal to $m_{h,AB} + y_{h,AB} + \delta_{h,AB} = m^*_{h,AB} + y_{h,AB}$) to each hospital h. Mechanism 3-xCM relaxes this target if it is not attainable, but always insists that each hospital gains at least the minimal number of matches it can achieve locally.

Theorem 6. *The 3-xCM mechanism is EPIC for properties (i) 3S-perfect and multi-hospital 3R-perfect on compatibility graphs of every marginal economy and larger, and (ii) 3OD/UD-perfect for individual hospitals. The 3-xCM mechanism is 3-way efficient if the combined graph is multi-hospital strong-3R perfect.*

The theoretical analysis of 3-xCM follows the same outline as the analysis for xCM, and is provided in the supplementary material. A part of the analysis is to establish, for an arbitrary strategy of hospital h and the generalized PM assumptions, that the mechanism matches every S pair, and with respect to the reported R pairs, as many R pairs on the long-side of the pooled R-subgraph as enabled by the total number of R-pairs and virtual-R pairs on the short-side. Efficiency follows by showing that the matching computed in 3-xCM is strong 3-way regular.

6. Incentive analysis of the Bonus mechanism

In this section, we examine the incentive properties of Bonus. We first provide a formal definition.

Definition 12 *(Bonus mechanism).* (See Ashlagi and Roth, 2014.) The Bonus mechanism receives graphs $\vec{G} = (G'_1, \ldots, G'_m)$ from each hospital. Let $G'_\oplus$ denote the combined graph. Assume an even-number of hospitals, and perform the following steps:

Step 1: (Match S pairs) Compute μ_S, the maximum matching of the S-subgraph of $G'_\oplus$.
Step 2: (Match R pairs) For each hospital h compute $\mu_{h,R}$, a maximum matching of the R-subgraph of G'_h. Compute μ_R, a maximum matching of the R-subgraph of $G'_\oplus$, breaking ties to maximize the intersection with $\bigcup_h \mu_{h,R}$.
Step 3: (Match OD-UD pairs) Initialize $\mu_{UO} := \emptyset$ and split the hospitals at random into two equal-sized groups L_1 and L_2. For each under-demanded type X–Y, do the following:
  (i) Select UD pairs from L_1 to match with OD pairs from L_2 as follows:
    (a) For each hospital $h \in L_1$, let $V_{h,\text{X–Y}}$ denote the under-demanded X–Y pairs reported by h. Let $Q_{h,\text{X–Y}} \subseteq V_{h,\text{X–Y}}$ denote a set of X–Y pairs that can be maximally matched in G'_h. Calculate $q_1 \triangleq \sum_{h \in L_1} |Q_{h,\text{X–Y}}|$.
    (b) For each hospital $h \in L_2$, let $V_{h,\text{Y–X}}$ denote the reported over-demanded Y–X pairs. Calculate $q_2 := \sum_{h \in L_2} |V_{h,\text{Y–X}}|$.
    (c) If $q_2 > q_1$: Run a lottery to pick additional X–Y pairs from L_1. Proceed in rounds:
      – In each round, select a hospital $h \in L_1$ with probability proportional to $|V_{h,\text{X–Y}}|$. Pick a pair at random from $V_{h,\text{X–Y}} \setminus Q_{h,\text{X–Y}}$, and move it from $V_{h,\text{X–Y}}$ to $Q_{h,\text{X–Y}}$. If there is no new pair to pick, remove h from the lottery process.
      – Repeat until $q_2 = \sum_{h \in L_1} |Q_{h,\text{X–Y}}|$, or there are no hospitals left in L_1.
    (d) Let $Q_{\text{X–Y}} \triangleq \bigcup_{h \in L_1} Q_{h,\text{X–Y}}$ and $V_{\text{Y–X}} \triangleq \bigcup_{h \in L_2} V_{h,\text{Y–X}}$. Compute a random maximum matching on the subgraph induced by pairs $Q_{\text{X–Y}} \cup V_{\text{Y–X}}$, and add this matching to μ_{UO}.
  (ii) Select UD pairs from L_2 to match with OD pairs from L_1 similar to Step 3(i).
Step 4: Return $\mu_S \cup \mu_R \cup \mu_{UO}$.

In Step 1, S pairs are matched without considering pooling across hospitals. In Step 2, a maximum matching is selected on the pooled R pairs, breaking ties to maximize the overlap with the number of pairs that each hospital would match in a maximum matching on its own R subgraph.

In Step 3, the hospitals are split into two groups, with the OD pairs in one group matched with the UD pairs in the other. Consider the process of lotterying hospitals in group L_1 to match with OD pairs reported by hospitals in group L_2 for a particular X–Y (UD) and Y–X (OD) pair. Step 3(i)(a) determines the UD pairs that each hospital would match just considering its own X–Y and Y–X pairs. The tally q_1 is the total count of pairs matched in this way, and represents the UD pairs that are preselected to match with some of the q_2 OD pairs that are available from L_2 (calculated in Step 3(i)(b)). If $q_2 > q_1$, and there are more OD pairs to allocate, then Step 3(i)(c) lotteries access to the remaining OD pairs, giving preference according to the number of UD pairs reported by each hospital. Once there are no UD pairs left to select, or the number of UD pairs selected is equal to the number of available OD pairs, Step 3(i)(d) finds a random maximum matching between the selected UD pairs and the OD pairs.

The 2-cycle and 3-cycle versions of Bonus are similar. The main difference is that the 3-cycle version can find 3-cycles in Step 1 when matching on the S-subgraph. 3-cycles can also be used when determining the internal matchings in Steps 2 and 3, but these are only used to control the primary matching process.[17,18]

6.1. Weakness of Bonus to canonical deviations

The lottery procedure in Bonus is designed to provide incentives for reporting UD and OD pairs. The idea is to give a hospital more chances to be selected to match with OD pairs reported by others if it reports more UD pairs. However, a numerical example taken from Ashlagi and Roth (2014) (see their Example 1) illustrates a weakness in this argument.

---

[17] The part of the matching generated by Bonus in Step 2 is limited to 2-cycles because it is restricted to the R-subgraph. The part of the matching generated in Step 3(i)(d) is limited to 2-cycles because it is restricted to a subgraph including one particular X–Y UD pair-type and the symmetric Y–X OD pair-type.
[18] xCM makes more use of 3-cycles, especially in regard to matching R pairs (see Section 6.2). This leads to better efficiency, as demonstrated in Section 7.

Reporting both UD and OD pairs can reduce the effective number of OD pairs available to compete for in the lottery stage (Step 3(i)(c)). The overall effect is that it can be useful to hide some UD and OD pairs.

In Example 1, we consider a *canonical deviation*, in which a hospital finds a maximum matching on its own pairs and reports the remainder. We assume for simplicity that there are at least as many UD pairs as OD pairs in the lottery, and that the matching in Step 3(i)(d) is perfect, so that every selected UD pair is matched.[19]

Example 1. Fix the two groups, L_1 and L_2, and consider a particular UD and OD pair-type (e.g., O–A and A–O). For group L_1, assume hospital $h \in L_1$ has 1 OD pair and 2 UD pairs, and can match its OD pair with one of its UD pairs. Assume that 4 UD pairs are reported in total by the other hospitals in L_1, and that no other hospital in L_1 can match a reported UD pair with one of its own reported OD pairs. The number of OD pairs reported by other hospitals in L_1 is immaterial. For group L_2, assume there are two OD pairs reported in total (so that $q_2 = 2$ in Step 3(i)(b)), and that there are plenty of reported UD pairs.

If hospital h is truthful, then its OD pair will be placed into a lottery and matched with a UD pair reported by group L_2. In regard to its UD pairs, one of its UD pairs (but none of the UD pairs of other hospitals in L_1) will be preselected, and so $q_1 = 1$. The lottery in Step 3(i)(c) involves 6 UD pairs (2 of them from h) competing to be the second UD pair selected from L_1 to match with the OD pairs provided by L_2. Another way to say this is that there is a lottery for the 1 remaining OD pair in the supply from L_2 that is not already assigned. Hospital h will be selected in the lottery with probability $\frac{2}{6}$. Therefore its expected number of pairs matched is $1 + 1 + \frac{2}{6} = \frac{35}{15}$, representing the OD pair that it matches, its preselected UD pair, and the chance of matching a second UD pair.

If hospital h follows a canonical deviation, then it matches a UD pair against an OD pair and sends just a single UD pair. Now none of the UD pairs in L_1 are preselected by the $\texttt{Bonus}$ lottery. The lottery in Step 3(i)(c) involves 5 UD pairs (1 of them from h) competing to be part of the 2 UD pairs selected from L_1 to match with the OD pairs provided by L_2. Another way to say this is that there is a lottery for the 2 OD pairs in the supply from L_2, now 2 rather than 1 because no pairs are already assigned. Hospital h will be selected in the lottery with probability $\frac{1}{5} + \frac{4}{5} \cdot \frac{1}{4} = \frac{6}{15}$, representing the probability that its UD pair is selected in the first or second round. Therefore its expected number of pairs matched is $2 + \frac{6}{15} = \frac{36}{15} > \frac{35}{15}$, considering also the matching that it completes based on the pairs it does not report. Thus, $\texttt{Bonus}$ does not align incentives with truthful reporting in this example.

More generally, consider a UD pair-type with patient–donor blood-type X–Y and its symmetric OD pair-type with patient–donor blood-type Y–X. Consider some hospital h in L_1, with x_h pairs of type X–Y, y_h pairs of type Y–X, and that can maximally match m_h pairs of type X–Y, internally. Let $q_2 = \sum_{i \in L_2} y_i$ denote the supply of Y–X pairs from L_2 and $q_1 = \sum_{i \in L_1} m_i$ denote the initial demand for these pairs based on internal matches in L_1. Also, define x_{-h} so that $x_{-h} = \sum_{i \in L_1, i \neq h} x_i$.

In order to simplify the analysis, we assume that all hospitals in L_1 remain interested in competing for OD pairs through all rounds of the lottery.[20] Because of this, the expected utility (= # matches) of hospital h, if it is truthful, is given by

$$U^t_{\texttt{Bonus}} = m_h + y_h + \min\left\{x_h - m_h, (q_2 - q_1)^+ \frac{x_h}{x_h + x_{-h}}\right\}, \tag{10}$$

where $x^+ = \max\{0, x\}$. Eq. (10) holds because h will match all m_h UD pairs, and all y_h OD pairs will be matched by the aforementioned perfect-matching assumption of the lottery. Furthermore, no more than $x_h - m_h$ pairs beyond m_h can be selected in the lottery, and with the effect of its x_h lottery tickets, hospital h expects to receive a proportion of $\frac{x_h}{x_h + x_{-h}}$ of the $(q_2 - q_1)^+$ OD pairs that will be randomly allocated in the lottery.

Now, let us consider that hospital h follows a canonical deviation and hides k X–Y pairs, where these pairs can be matched to k of its own Y–X pairs. Then its report will be $x_h - k$ UD pairs and $y_h - k$ OD pairs, and the excess will be $q_2 - (q_1 - k)$. By recording the benefit of $2k$ and adjusting terms in the previous expression, the expected utility of the deviating agent, is:

$$U^d_{\texttt{Bonus}} = 2k + (m_h - k) + (y_h - k) + \min\left\{x_h - m_h, (q_2 - q_1 + k)^+ \frac{x_h - k}{x_h - k + x_{-h}}\right\}$$

$$= m_h + y_h + \min\left\{x_h - m_h, (q_2 - q_1 + k)^+ \frac{x_h - k}{x_h - k + x_{-h}}\right\}. \tag{11}$$

By deviating, hospital h gets a smaller fraction from the lottery since $\frac{x_h - k}{x_h - k + x_{-h}} < \frac{x_h}{x_h + x_{-h}}$, for $k > 0$. However, the lottery is selecting an additional k UD pairs from the pool (equivalently it is allocating k additional OD pairs). By matching its own

---

[19] These assumptions correspond to asymptotic properties established by Ashlagi and Roth (2014), see Claim 10.15. They can be justified for a large number of hospitals each with small lists.

[20] In the case of dropouts, (10) changes slightly by considering instead a fraction $x_h/(x_h + x_{-h} + z_{-h})$, where z_{-h} is an adjustment made to allow for the possibility that other hospitals will drop out of the lottery when their remaining demand to match OD pairs is satiated (see Step 3(i)(c)).

pairs before sending, the hospital guarantees the use of its own OD pairs and still gets to compete for the full pool of OD pairs provided by others to the lottery.[21]

Through simple algebra, a sufficient condition for the deviation to be useful is $x_h - k > q_2 - q_1$. This reveals the concern. If the report after deviation ($x_h - k$) is still larger than the excess to allocate in the lottery ($q_2 - q_1$), then the deviating hospital will be better off. Under the OD/UD-perfect assumption, hospital h will hide every OD pair and a corresponding number of UD pairs, and the quantity $x_h - k$ is $\propto n$; i.e., the individual hospital's graph size. However, ($q_2 - q_1$) is a quantity that is proportional to the excess of the long-side in a bipartite graph pooled from m hospitals; by Lemma 5 it follows that $|q_2 - q_1| \propto m\sqrt{n}$ in expectation.

Thus, opportunities to manipulate the Bonus mechanism exist for a broad range of settings where the size of an individual hospital n is not fixed, and the number of hospitals m increases at a rate $m = O(\sqrt{n})$ (and thus not too quickly). This analysis also suggests that as the number of hospitals m grows while the size of each hospital remains fixed, this deviation becomes less useful. This in agreement with the theory of Ashlagi and Roth (2014), as well as our simulation results (see Section 7.4).[22]

6.2. *Bonus vs. xCM on the R-subgraph*

The two mechanisms also differ in regard to how they match R pairs. Bonus finds, amongst the maximum matchings on the pooled R-subgraph, one that maximizes the intersection with maximum matchings on the R-subgraphs restricted to each hospital. In this sense, Bonus treats individual rationality as a soft constraint. In contrast, xCM uses the probabilistic uniform rule to determine how many R pairs a hospital can match above those it would match by itself, and ensures that each hospital matches at least as many as it would by itself. The supplementary material illustrates the difference between the two designs on a simple example. In particular, we show that Bonus does not provide *ex post* incentive-compatibility in regard to truthful reporting of R-pairs. The Bonus and xCM mechanisms also differ in that xCM but not Bonus allows R-pairs to match on 3-cycles.

7. Experimental results

In this section, we report results from simulations on mechanisms xCM and Bonus, and compare them with a baseline mechanism, rCM, which returns a random maximum matching on the combined graph.[23] We test the various PM properties, validate the square-root law, and explore the welfare and incentive properties, considering both uniform and non-uniform PRA.[24] All results are presented for the blood-type frequencies and tissue-type compatibility models introduced in Section 2, and averaged over 1000 trials except where otherwise stated.

7.1. *Perfect-matching properties and square-root law*

First, we investigate the validity of the S-perfect, R-perfect and OD/UD-perfect properties. For 2-cycles, we consider graphs $\widetilde{G_n}$ with sizes n from 30 to 130, and for 3-cycles we consider graphs $\widetilde{G_n^3}$ with sizes n from 30 to 90. The tests compute the difference in number of pairs matched on the relevant subgraph and the idealized number that would match under the various perfect properties; see Appendix B for details. All counts are normalized by the cardinality of the maximum matching on the graph, and averaged over 500 trials.

The results for 2-cycles are shown in Table 2, where idealized performance would be a 0% violation. For uniform PRA, the OD/UD-perfect property is satisfied with 1% violation for graphs of size 30 or larger. The R-perfect and S-perfect properties are satisfied with 1% violation for graphs of size at least 70 and 130 pairs, respectively. For Theorem 6, OD/UD-perfect should hold for individual hospitals, while R-perfect and S-perfect for combined lists of all but one hospital. Allowing for a 2% violation, the assumptions required for Theorem 6 are justified under uniform PRA for 4 hospitals each with lists of size 25 or larger.

---

[21] This increase in the number of excess pairs to allocate in the lottery from $q_2 - q_1$ to $q_2 - q_1 + k$ appears to have been over-looked in the analysis in Ashlagi and Roth (2014).

[22] A more formal analysis would be required to fully quantify the trade-offs between withholding and truthful reporting in Bonus, including the case of very large markets with hospitals with fixed list sizes.

[23] We use R (http://www.r-project.org/) for the code implementation. Maximum matchings are obtained through integer programming (IP) with Gurobi (http://www.gurobi.com/) and a free academic license. We follow a cycle formulation for the IP that performs weighted matching (each matched vertex has weight 1 by default). Random matchings are obtained by perturbing slightly (± 0.1) the weights of every cycle. Matchings that promote specific pair-types (crucial for both xCM and Bonus) are obtained by increasing the respective weights. An almost regular matching is achieved by increasing slightly (e.g. $+0.2$) the weights of OD-UD 2-cycles. The source code is available for download from http://www.eecs.harvard.edu/econcs/code/rgke.zip and the code base is also available through Github at https://github.com/ptoulis/kidney-exchange. Detailed instructions on how to reproduce all results of this section can be found in the package.

[24] For brevity, we will sometimes use the symbols O, R, S, and U to denote OD, R, S, and UD pairs respectively, where O refers to OD and not the O blood-type. In this way, a cycle will be denoted by the types of the pairs in the cycle; e.g., OUU denotes a 3-cycle among one OD pair and two UD pairs. Furthermore, for 3-cycles, we will use standard regular expression notation. For instance, OO? denotes any 3-cycle that contains two OD pairs, and O[RS] denotes any 3-cycle that contains one OD pair and any combination of two pairs from types R or S i.e., ORR, ORS or OSS.

Table 2
The amount of violation of R-, OD/UD- and S-perfect properties for graphs of various sizes, normalized by the size of a maximum matching on $\widetilde{G}_n$. 2-cycles only, and for uniform (U) and non-uniform PRA (nonU) models. Standard errors in parentheses. For example, for $n = 30$ the value 1.11 under OD/UD-perfect indicates that x OD pairs remained unmatched, where x represents 1.11% of the size of a maximum matching on $\widetilde{G}_{30}$. The OD/UD-perfect assumption requires that all OD pairs are matched.

PRA	Size (n)	R-perfect	OD/UD-perfect	S-perfect
U	30	4.89 (1.08)	1.11 (0.36)	7.30 (1.43)
	50	2.17 (0.51)	0.79 (0.20)	3.43 (0.67)
	70	1.05 (0.30)	0.46 (0.14)	1.92 (0.42)
	90	0.37 (0.16)	0.22 (0.09)	1.75 (0.32)
	130	0.00 (0.00)	0.13 (0.06)	1.22 (0.22)
nonU	30	12.25 (2.47)	10.39 (1.17)	26.20 (2.07)
	50	6.65 (0.73)	7.22 (0.88)	25.62 (1.92)
	70	5.03 (0.65)	4.30 (0.45)	19.36 (1.16)
	90	3.76 (0.53)	3.34 (0.37)	16.77 (1.03)
	130	2.19 (0.30)	1.77 (0.24)	12.63 (0.68)

Table 3
The amount of violation of 3R-, 3OD/UD- and 3S-perfect properties for graphs of various sizes, allowing for 3-cycles, and for both uniform (U) and non-uniform PRA (nonU) models. The AB–O statistic is the fraction of AB–O pairs that fail to match in 3-cycles with two UD pairs. Violations are normalized by the size of a maximum matching on $\widetilde{G}_n$ allowing for 3-cycles. Standard errors in parentheses.

PRA	Size (n)	3R-perfect	3OD/UD-perfect	AB–O	3S-perfect
U	30	7.00 (0.86)	2.06 (0.35)	0.00 (0.00)	8.54 (0.77)
	50	4.93 (0.76)	1.02 (0.17)	0.00 (0.00)	4.20 (0.38)
	70	3.30 (0.60)	0.55 (0.11)	0.00 (0.00)	2.69 (0.30)
	90	2.06 (0.40)	0.45 (0.09)	0.00 (0.00)	2.05 (0.18)
nonU	30	17.68 (1.96)	9.41 (1.28)	0.51 (0.18)	32.26 (2.54)
	50	6.99 (0.91)	4.76 (0.47)	0.24 (0.11)	18.53 (1.03)
	70	5.16 (0.62)	3.71 (0.34)	0.11 (0.05)	13.72 (0.80)
	90	3.63 (0.45)	2.87 (0.24)	0.07 (0.04)	12.20 (0.58)

Larger list sizes are required if tissue-type compatibility is non-uniform; e.g., individual hospitals would need to have as many as 130 pairs for the OD/UD-perfect property to be satisfied with 1% violation, and the S-perfect property requires combined lists of significantly more than 130 pairs.

The results allowing for 3-cycles are shown in Table 3. The AB–O statistic is the fraction of AB–O pairs that fail to match in a maximum matching restricted to 3-cycles involving two UD pairs. As in Table 2, all results are normalized by the cardinality of a maximum matching, this time allowing for 3-cycles. For uniform PRA, the 3OD/UD-perfect property is satisfied with 1% violation for graphs of size 50 or larger, and at 2% violation for graphs of size 30 or larger. Considering that Theorem 6 requires the 3R-perfect and 3S-perfect properties to hold for combined lists of all but one hospital, the required PM properties are satisfied at 2% violation with 4 hospitals, each with lists of size 30 or larger. The AB–O pairs can be easily matched, even for small graphs. With non-uniform PRA, the situation is similar to with 2-cycles, and larger lists are required; more than 90 pairs for an individual hospital to achieve 3OD/UD-perfect at 2% violation, and combined lists of significantly more than 90 pairs to satisfy the 3S-perfect property.

We now validate Theorem 2, which provides an analytical expression for $\mu(n)$, the expected cardinality of a maximum matching in $\widetilde{G}_n$, and leads to the square-root law. The analytical expression for the number of matches in a maximum matching of $\widetilde{G}_n$ is $\mu(n) \approx 0.556n - 0.338\sqrt{n} - 2$, for a graph of size n (see Section 3). This is for 2-cycles only and with uniform PRA. Table 4 compares the average number of matches in simulation against this expression. The results are averaged over 10 000 trials. To test the fit, we run a linear regression of the number matches, and obtain $\mu(n) = 0.567n - 0.572\sqrt{n} - 0.9$.[25] The prediction is generally very close, but usually over-estimating slightly due to violations in the PM assumptions.

7.2. Multi-hospital kidney exchange with uniform PRA

In this section, we compare the incentive and welfare properties of rCM, xCM and Bonus, considering only uniform PRA.

7.2.1. Strategic properties
The individual strategies we consider are truthful behavior (t) and a canonical deviation (c), which is the strategy when a hospital first finds an almost-regular (maximum) matching and then reports only unmatched pairs to the mechanism.

---

[25] The coefficients on n and $\sqrt{n}$ are statistically significant with p-values less than 10^{-5}, with standard errors 0.007 and 0.137 for the coefficients of n and $\sqrt{n}$ respectively.

Table 4

Validation of the square-root law: Average size of maximum matching on $\widetilde{G_n}$ for various sizes n compared to the theoretical expression developed in Section 3.

Size (n)	Observation average #matches	Prediction $\mu(n)$
25	10.153	10.210
50	22.746	23.410
75	36.784	36.773
100	50.351	50.220
125	62.724	63.721
150	76.691	77.260
175	90.018	90.829
200	104.380	104.420

Table 5

Incentive analysis under the uniform PRA model, allowing for 2-cycles only, with $m = 6$ hospitals each having $n = 12$ patient–donor pairs. The entry in the strategy column reflects the strategy type for which average utility (std error), and cumulative number of OD, R, S and UD pairs matched is reported, considering all 1000 trials.

Mechanism	Profile	Strategy	Avg. utility	#OD	#R	#S	#UD
rCM	tttttt	t	5.90 (0.03)	1629	2088	1778	1584
	tttttc	c	6.68 (0.06)	1604	2267	1802	2347
	tccccc	t	4.54 (0.05)	1492	1656	1476	371
	cccccc	c	5.57 (0.03)	1462	1878	1580	1201
xCM	tttttt	t	5.85 (0.03)	1618	2125	1811	1458
	tttttc	c	5.81 (0.03)	1600	2137	1765	1468
	tccccc	t	5.59 (0.07)	1485	1840	1582	1246
	cccccc	c	5.57 (0.03)	1457	1879	1583	1201
Bonus	tttttt	t	5.67 (0.03)	1537	2081	1778	1410
	tttttc	c	6.19 (0.06)	1571	1994	1784	2075
	tccccc	t	4.75 (0.06)	1405	1889	1511	424
	cccccc	c	5.50 (0.03)	1428	1870	1574	1179

A strategy profile for 6 hospitals is denoted by profile tttttt, tttttc, and so forth. We adopt the canonical deviation because this is the deviation for which a vulnerability of Bonus is identified (see Section 6).

In Table 5, we report the average utility (= number of matches, including those matched in recourse) for the hospital with the indicated strategy, along with the standard errors. We also report the total number of OD, R, S and UD pairs matched across all trials by a hospital with the indicated strategy.

We can make the following observations:

- rCM is not *ex post* incentive compatible. In particular, the canonical deviation provides an average utility of 6.68 vs 5.90 when the other hospitals are truthful. This deviation remains beneficial compared to truthful reporting when others also deviate. A deviating hospital matches more UD and R pairs by free-riding on the reports of others.
- The canonical deviation is not useful in xCM at least on average (compare 5.81 with 5.85), which is consistent with the theoretical analysis of the mechanism.
- The Bonus mechanism is not *ex post* incentive compatible for this number of hospitals. In particular, the canonical deviation provides an average utility of 6.19 vs 5.67 when the other hospitals are truthful. This deviation remains beneficial compared to truthful reporting when others also deviate. A deviating hospital matches more UD pairs by free-riding, which is consistent with the theoretical analysis of the strategic vulnerability of the mechanism.

Table 6 provides an analogous incentive analysis of the mechanisms, now allowing for 3-cycles. The qualitative results are largely unchanged. The main difference is that we now see a small benefit in xCM for deviating to the canonical strategy. This can be explained by noting that the PM properties required for Theorem 6 are not well supported until the combined lists of all but one hospital have more than 90 pairs, and the lists of single hospitals have more than 30 pairs.

Table 7 summarizes the incentive properties of the various mechanisms, varying the number of hospitals and size of hospital lists and considering both 2-cycle and 3-cycle settings. We fix the other hospitals to report complete lists, and record the average ratio (and standard error) of the number of matches achieved by a hospital from adopting the canonical strategy compared to the number of matches when reporting its complete list. We make the following observations:

- Neither rCM nor Bonus are *ex post* incentive compatible for any of these (m, n) combinations, irrespective of 2- or 3-cycles. The rCM and Bonus have qualitatively similar incentive properties in these environments.

Table 6
Incentive analysis under the uniform PRA model allowing for 3-cycles, with $m = 6$ hospitals each having $n = 12$ donor–patient pairs. The entry in the strategy column reflects the strategy type for which average utility (std error), and cumulative number of OD, R, S and UD pairs matched is reported, considering all 1000 trials.

Mechanism	Profile	Strategy	Avg. utility	# OD	#R	#S	#UD
rCM	tttttt	t	6.61 (0.04)	795	1262	1073	838
	tttttc	c	7.37 (0.09)	814	1301	1093	1211
	tccccc	t	4.93 (0.08)	830	970	980	177
	cccccc	c	5.90 (0.04)	807	1069	981	684
xCM	tttttt	t	6.29 (0.04)	785	1238	1015	736
	tttttc	c	6.41 (0.09)	810	1244	1066	724
	tccccc	t	5.80 (0.10)	825	1092	868	693
	cccccc	c	5.81 (0.04)	806	1068	932	678
Bonus	tttttt	t	5.85 (0.04)	762	1059	984	700
	tttttc	c	6.61 (0.09)	810	1096	1064	996
	tccccc	t	4.75 (0.08)	795	1054	804	194
	cccccc	c	5.77 (0.04)	801	1058	925	675

Table 7
Incentive summary (uniform PRA): Average ratio (std error) of the total number of pairs matched by a hospital that uses the canonical deviation compared to the total number matched when truthful, and when other hospitals are truthful. Ratios larger than one indicate that the canonical deviation is useful.

Mechanism	$(m, n) =$ (# hospitals, # pairs per hospital.)		
	(4, 18)	(6, 12)	(12, 6)
rCM	1.148 (0.007)	1.133 (0.008)	1.141 (0.014)
3-rCM	1.124 (0.008)	1.113 (0.011)	1.090 (0.013)
xCM	0.995 (0.008)	0.994 (0.009)	1.021 (0.015)
3-xCM	1.022 (0.010)	1.018 (0.013)	0.997 (0.015)
Bonus	1.116 (0.008)	1.091 (0.009)	1.091 (0.015)
3-Bonus	1.158 (0.011)	1.131 (0.014)	1.058 (0.016)

Table 8
Comparison of the number of OU and RR 2-cycles in xCM and Bonus when all hospitals are truthful and $m = 6$, $n = 12$. Bonus performs fewer OU cycles because of its lottery procedure.

Mechanism	PRA	#OU	#RR	#total pairs
xCM-all-t	uniform	8753	6218	21 043
	non-uniform	7226	5260	16 081
Bonus-all-t	uniform	8461	6218	20 420
	non-uniform	5486	5260	13 925

- Even though these environments do not meet the PM properties required for the theoretical incentive properties of xCM and 3-xCM (the hospital lists are too small, and the combined lists of all but one hospital are too small), the incentive properties are qualitatively better than the other mechanisms. The canonical deviation does not appear to be useful in xCM with 2-cycles, but may be marginally useful in 3-xCM.

7.2.2. Welfare properties

Table 8 presents an overall comparison of the number of OU 2-cycles and RR 2-cycles in xCM and Bonus given truthful reports. This is for $m = 6$, $n = 12$, and considering both uniform and non-uniform PRA. The xCM mechanism matches more OU cycles than Bonus even though the OD and UD pairs are not pooled except in Step 4 of xCM. The mechanisms match the same number of R pairs.

Table 9 illustrates the efficiency of each mechanism with 2-cycles, and providing hospitals with the canonical strategy when this is beneficial. For xCM we present the results for truthful reports (all-t). For Bonus and rCM we present the results for canonical reports (all-c). In addition to welfare (= total number of matches, averaged over all trials), we report the total number of different types of cycles matched across all trials, as well as the fraction of cycles that are matched within the mechanism rather than in recourse. For a baseline, we provide the statistics for maximum matchings obtained with and without pooling on truthful graphs. The analysis reveals an overall welfare ranking of

xCM-all-t > rCM-all-c > Bonus-all-c.

Table 9

Average welfare (std err) for each mechanism as well as total counts of different types of 2-cycles matched across all trials. Uniform PRA model, with $m = 6$ hospitals each having $n = 12$ patient–donor pairs. Below each cycle count is the fraction of cycles that are matched by the mechanism rather than in recourse.

Mechanism	Welfare	OO	OR	OS	OU	RR	SS
no pooling	24.06 (0.15)	97	885	922	8662	3579	2695
max matching	35.40 (0.18)	5	168	95	9504	6180	5287
rCM-all-c	33.40 (0.18)	79	673	730	7211	5300	4376
		2.5	9.7	5.5	5.6	46.8	49.8
xCM-all-t	35.07 (0.17)	185	315	274	8753	6218	5298
		100	100	100	100	100	100
Bonus-all-c	33.01 (0.17)	83	627	706	7074	5297	4369
		2.4	0	0	3.8	47.2	50

Table 10

Average welfare (std error) for each mechanism as well as total counts of different types of 3-cycles that match across all trials. Uniform PRA model, with $m = 6$ hospitals each having $n = 12$ patient–donor pairs. Below each cycle count is the fraction of cycles that are matched by the mechanism rather than in recourse.

Mechanism	Welfare	OO?	O[RS]	ORU	OSU	OUU	RRS	SSS
no pooling	29.73 (0.24)	153	572	1205	1479	81	573	369
max matching	39.51 (0.23)	1	50	1694	93	428	1686	1073
3-rCM-all-c	35.40 (0.20)	164	480	1203	1623	83	1055	505
		0	3.54	2.74	0.92	0	46.16	27.52
3-xCM-all-t	37.56 (0.22)	299	20	2408	200	170	0	1263
		100	100	100	100	100	–	100
3-Bonus-all-c	34.62 (0.20)	175	478	1167	1586	82	568	515
		1.14	0	0	0	0	0	28.93

Table 11

Total count of matched pair-types and matched 2-cycle types for each mechanism across all trials. Uniform PRA model, with $m = 6$ hospitals each having $n = 12$ donor–patient pairs.

Mechanism	#O	#R	#S	#U	#OU	#RR	#SS
no pooling	6216	6075	6288	5205	2252	1512	1176
max matching	6337	10046	8538	6684	4041	2444	1752
3-rCM-all-c	6458	8547	7836	5478	2368	23145	1454
	1.8	30.2	19.5	1.7	1.9	3.3	20.5
3-xCM-all-t	6258	9847	8073	5874	2646	3701	1999
	100	100	100	100	100	100	100
3-Bonus-all-c	6409	8464	7398	5422	2368	2798	1468
	0.9	29.6	14.5	1.0	2.1	44.8	21.4

Relative to the welfare obtained without pooling, the most important gains in xCM come from the matching of additional RR and SS cycles.[26]

Tables 10 and 11 illustrate the welfare also allowing 3-cycles, again adopting all-t for 3-xCM and all-c for 3-rCM and 3-Bonus. Table 10 provides a total count of the different types of 3-cycles in the matchings. Table 11 provides a summary count of the total number of OD, R, S, and UD pair-types matched, and the main types of 2-cycles matched. The results are qualitatively unchanged from the setting with 2-cycles, with the same welfare ranking between mechanisms. There is also a significant benefit from pooling.[27]

Table 12 summarizes the welfare in each mechanism across the three different environments, and compares with the welfare without pooling. The welfare is computed as the average fraction of pairs matched relative to those in a maximum matching when all lists are pooled. The efficiency of xCM-all-t is between 98% and 99% for 2-cycles and 93% and 95% for 3-cycles. The efficiency of Bonus under canonical deviations is generally worse than rCM under canonical deviations.

7.3. Multi-hospital kidney exchange with non-uniform PRA

Table 13 summarizes the incentives to deviate to the canonical strategy in each mechanism, presenting this as a fraction of the utility from truthful reports, and considering non-uniform PRA. For xCM with 2-cycles there is no incentive to deviate.

---

[26] The main inefficiency in xCM arises because fewer OU pairs and thus fewer UD pairs are matched. Bonus also matches RR and SS pairs, but because hospitals use canonical deviations the other types of matches occur in recourse.

[27] The main inefficiency in 3-xCM-all-t relative to the maximum matching occurs because fewer UD and S pairs match in xCM, due to fewer OD-UD cycles and RRS cycles, respectively. It seems likely that the efficiency of 3-xCM could be further improved when 3S-perfect does not hold by facilitating RRS cycles.

Table 12
Welfare summary (uniform PRA): Average ratio (std error) of the total number of pairs matched relative to the size of a maximum matching.

Mechanism	$(m, n) =$ (# hospitals, # pairs per hospital.)		
	(4, 18)	(6, 12)	(12, 6)
no pooling	0.771 (0.006)	0.682 (0.005)	0.505 (0.005)
3-no pooling	0.845 (0.008)	0.759 (0.007)	0.535 (0.005)
rCM-all-c	0.957 (0.007)	0.943 (0.007)	0.943 (0.007)
3-rCM-all-c	0.908 (0.008)	0.893 (0.008)	0.875 (0.006)
xCM-all-t	0.994 (0.001)	0.991 (0.001)	0.982 (0.001)
3-xCM-all-t	0.959 (0.001)	0.951 (0.002)	0.935 (0.002)
Bonus-all-c	0.952 (0.007)	0.933 (0.007)	0.912 (0.007)
3-Bonus-all-c	0.899 (0.008)	0.872 (0.008)	0.823 (0.006)

Table 13
Incentive summary (non-uniform PRA): Average ratio (std error) of the total number of pairs matched by a hospital that uses the canonical deviation compared to the total number matched when truthful, and when other hospitals are truthful.

Mechanism	$(m, n) =$ (# hospitals, # pairs per hospital.)		
	(4, 18)	(6, 12)	(12, 6)
rCM	1.121 (0.008)	1.099 (0.011)	1.059 (0.016)
3-rCM	1.064 (0.008)	1.079 (0.011)	1.074 (0.012)
xCM	1.008 (0.009)	0.992 (0.011)	0.990 (0.016)
3-xCM	1.009 (0.011)	1.016 (0.013)	1.030 (0.015)
Bonus	1.123 (0.010)	1.089 (0.013)	1.061 (0.020)
3-Bonus	1.167 (0.014)	1.152 (0.017)	1.099 (0.020)

Table 14
Welfare summary (non-uniform PRA): Average ratio (std error) of the total number of pairs matched relative to the size of a maximum matching.

Mechanism	$(m, n) =$ (# hospitals, # pairs per hospital.)		
	(4, 18)	(6, 12)	(12, 6)
rCM-all-c	0.947 (0.005)	0.946 (0.005)	0.948 (0.005)
3-rCM-all-c	0.890 (0.005)	0.887 (0.005)	0.890 (0.005)
xCM-all-t	0.977 (0.001)	0.976 (0.001)	0.976 (0.001)
3-xCM-all-t	0.891 (0.001)	0.887 (0.001)	0.892 (0.001)
Bonus-all-c	0.864 (0.005)	0.831 (0.005)	0.785 (0.005)
3-Bonus-all-c	0.791 (0.004)	0.741 (0.004)	0.666 (0.004)

For xCM with 3-cycles there may be a very small benefit, but the results are often within the standard error. On the other hand, rCM and Bonus remain vulnerable to canonical deviations. Table 14 presents a summary of the welfare properties, relative to the maximum matching on the pooled lists. We adopt all-t for xCM and all-c for rCM and Bonus. With 2-cycles, the welfare ranking remains xCM-all-t > rCM-all-c > Bonus-all-c. For 3-cycles, the welfare of 3-rCM-all-c and 3-xCM-all-t are very similar, confirming that larger hospitals and more hospitals are important for 3-xCM in non-uniform PRA environments.

7.4. Additional experiments with many, small-sized hospitals

In this set of experiments we compare xCM and Bonus in environments for which the theory suggests that Bonus should be more suited. For this purpose, we increase the number of hospitals and reduce their individual sizes, considering 2-cycles and the uniform PRA model. All standard errors were computed on 1000 bootstrap samples.

We start with $m = 24$ hospitals, with $n = 10$ pairs each, and set hospitals H_2 through H_{24} to truthful reporting. Then, under Bonus, hospital H_1 matches +0.2 (se = 0.03) more pairs/experiment in canonical deviation than in truthful reporting. In contrast, under the xCM mechanism, H_1 achieves −0.06 (se = 0.02) fewer matches if it canonically deviates. There is also a statistically significant difference in the equilibrium efficiency of the two mechanisms: Bonus achieves 121.5 (se = 0.28) total matches on average under canonical deviation by all hospitals, whereas xCM achieves 130.6 (se = 0.28) total matches on average under truthful reporting by all hospitals.

Similarly, for $m = 30$ and $n = 8$ pairs each, a deviating hospital achieves +0.1 (se = 0.03) more matched pairs when deviating under Bonus but incurs a loss of −0.05 pairs (se = 0.02) under xCM. Furthermore, Bonus achieves 119.8 (se = 0.3) total matches on average under the canonical deviation by all hospitals, whereas xCM achieves 130.03 (se = 0.3)

total matches on average under truthful reporting by all hospitals. While the incentive issues in `Bonus` are mitigated as the number of hospital increases, `Bonus` still remains vulnerable to deviations for up to 30 hospitals.

Appendix A. Additional proofs

Theorem 4. *The* xCM *mechanism is EPIC and 2-way efficient for properties (i) S-perfect and R-perfect on compatibility graphs the size of every marginal economy and larger, and (ii)* OD/UD-*perfect on every hospital's compatibility graph.*

Lemma 6. *Fix any hospital h. Assuming the S-perfect and R-perfect property for compatibility graphs the size of the marginal economy without h and larger, and the* OD/UD-*perfect property for the reports of hospitals other than h, then for any strategy s'_h of h, the* xCM *mechanism satisfies the following properties:*

1. *In Step 1 with regard to S pairs, and for any type* T ∈ {O, A, B, AB}*:*
   (i) *the matching μ_{xCM} is maximum on the subgraph of combined graph $G'_\oplus$ restricted to T–T pairs,*
   (ii) *hospital h matches all except for at most one reported T–T pair, and if the parity of $\tau'_\oplus$ is even then it matches all reported T–T pairs, and*
   (iii) *hospital h matches all reported T–T pairs with non-zero probability when at least one hospital $j \neq h$ reports an odd number of T–T pairs.*
2. *In Step 2 with regard to R pairs, and assuming $\alpha'_\oplus \geq \beta'_\oplus$, then,*
   (i) *all reported B–A pairs are matched, and the matching μ_{xCM} is maximum on the subgraph of combined graph $G'_\oplus$ restricted to R pairs,*
   (ii) *if $(\alpha'_h \leq \beta'_h)$ or $(\alpha'_\oplus = \beta'_\oplus)$ then all reported A–B pairs of hospital h are matched, else the expected number of reported A–B pairs matched is,*

$$\beta'_h + \min\big(\alpha'_h - \beta'_h, g_{\mathrm{uniform}}(z_{-h,\mathrm{AB}}, x_{\mathrm{BA}})\big),$$

   where $z_{-h,\mathrm{AB}}$ and x_{BA} are as defined in the xCM *mechanism.*
2′. *Symmetrically for the case that $\beta'_\oplus > \alpha'_\oplus$.*
3. *In Steps 3 and 4, any additional pairs that are matched involve an* OD *pair of hospital h.*

Proof. Consider hospital h. Let G'_{-h} denote the combined graph from hospitals other than h.

Consider any T ∈ {O, A, B, AB}.

1 (ii): If $\tau'_\oplus$ is even, then a maximum matching on the subgraph is perfect by the S-perfect property, and satisfies constraints (7). In particular, hospital h matches all its reported T–T pairs. If $\tau'_\oplus$ is odd, consider two cases. If τ'_h is odd, then the following almost-perfect matching exists by the S-perfect property and satisfies all constraints (7): match every T–T pair of every hospital except h and all but one T–T pair of hospital h. If τ'_h is even, then some other hospital $j \neq h$ must have an odd parity T–T graph and the following almost-perfect matching exists by the S-perfect property and satisfies all constraints (7): match every T–T pair of all hospitals other than h and j, every pair but one of j, and all pairs of h. Because there exists an almost-perfect matching that satisfies the constraints, then the matching μ_T computed in Step 1 must leave h with at most one unmatched T–T pair. This also establishes 1 (i) since in both cases there is a maximum matching that satisfies constraints (7).

1 (iii): Let $j \neq h$ denote a hospital with an odd parity T–T graph, and consider the interesting case, where the parity $\tau'_\oplus$ is odd (otherwise, there is a perfect matching). Whatever the parity of τ'_h, the almost-perfect matching that matches every T–T pair of every hospital except j exists (by the S-perfect property), and satisfies all constraints. Because of this, and since the maximum matching must leave one T–T unmatched and Step 1 selects matching μ_T at random, the probability that all T–T pairs of h are matched is non-zero.

2: For the R-subgraph, and the case $\alpha'_\oplus \geq \beta'_\oplus$. A maximum matching includes all B–A pairs by the R-perfect property.

2 (i): We construct a maximum matching that satisfies constraints (8) with $q = 0$, and is essentially equivalent to the matching computed by xCM (under the R-perfect property). For this, order the hospitals such $(\alpha'_j \leq \beta'_j)$ for hospitals $j \in \{1, \ldots, r\}$, and not for $j \in \{r+1, \ldots, m\}$. Consider the graph $G^\dagger$ induced by including the following number of A–B and B–A pairs for each hospital (and corresponding edges),

$$(\alpha'_1, \beta'_1), \ldots, (\alpha'_r, \beta'_r), (\beta'_{r+1} + y_{r+1,\mathrm{AB}}, \beta'_{r+1}), \ldots, (\beta'_m + y_{m,\mathrm{AB}}, \beta'_m),$$

where quotas $y_{j,\mathrm{AB}}$ are as defined by xCM. In particular, $\sum_{j>r} y_{j,\mathrm{AB}} = \sum_{j \leq r}(\beta'_j - \alpha'_j)$, so that $\sum_{j \leq r} \alpha'_j + \sum_{j>r}(\beta'_j + y_{j,\mathrm{AB}}) = \sum_j \beta'_j$, and $G^\dagger$ is a balanced bipartite graph. By the R-perfect property, there exists a perfect matching on $G^\dagger$. Moreover, this matching is maximum on the R-subgraph because it includes all B–A pairs. Because all B–A pairs are matched, the matching satisfies the B–A constraints (8) with $q = 0$. The A–B constraints are satisfied for $j \leq r$, because $m_j + \max(0, y_{j,\mathrm{AB}} + \delta_j) = m_j + \delta_j + y_{j,\mathrm{AB}} = \alpha'_j + 0$. The A–B constraints are satisfied for $j > r$ because $m_j + \max(0, y_{j,\mathrm{AB}} + \delta_j) = m_j + \delta_j + y_{j,\mathrm{AB}} = \beta'_j + y_{j,\mathrm{AB}}$. Because the A–B constraints hold with equality, the maximum matching determined by xCM must match the same number of A–B and B–A pairs for each hospital.

2 (ii): If $\alpha'_\oplus = \beta'_\oplus$ then $r = m$ and $\alpha'_j = \beta'_j$ for all hospitals j. If $\alpha'_h \leq \beta'_h$ then the matching includes all reported A–B pairs of hospital h by construction. Otherwise, if $\alpha'_\oplus > \beta'_\oplus$ and $\alpha'_h > \beta'_h$ then the expected number of reported A–B pairs matched is $\beta'_h = m_h + \delta_h$ plus the expected size of quota $y_{h,AB}$, which is $\min(\alpha'_h - \beta'_h, g_{uniform}(z_{-h,AB}, x_{BA}))$.

3: In regard to Step 3: by (1) and (2) there are no R–R or S–S edges in the remaining compatibility graph, and thus the only possible matches involve OD pairs. In regard to Step 4: there are no OD pairs remaining from other hospitals because of the OD/UD-assumption, and no remaining (cross-hospital) S–S matchings or R–R matchings because of Step 1 and Step 2 of xCM and the R-perfect and S-perfect assumptions. □

The expectation in the statement of the following lemma is taken with respect to randomization internal to xCM.

Lemma 7 (*Best response in* xCM). *Fixing the reports of other hospitals, if a strategy s_h of hospital h with graph G_h,*

(i) *matches every OD pair in G_h with a UD pair in G_h, and maximizes across all strategies the expected number of UD pairs in G_h matched with OD pairs of other hospitals,*

(ii) *maximizes across all strategies the expected number of S pairs in G_h matched, not counting S pairs that match with OD pairs of G_h, and*

(iii) *maximizes across all strategies the expected number of R pairs in G_h matched, not counting R pairs that match with OD pairs of G_h, then*

strategy s_h is a best-response.

Proof. Let N_R and N_S denote the number of R and S pairs matched by s_h. The only possible improvement is some strategy s'_h that changes the matching so that $k > 0$ of its OD pairs match with some of the R or S pairs in G_h. This must decrease the number of its UD pairs that match by at least k (by (i)). Suppose it increased the number of its R or S pairs that match by more than k. Then more than $N_R + N_S$ pairs match without using an OD pair of h, a contradiction with (ii) and (iii). This completes the proof. □

Proof of Theorem 4. Consider hospital h, and fix the (truthful) reports G_{-h} of the other hospitals. To establish the EPIC property, we show that a truthful strategy for h meets the requirements of Lemma 7.

(i) Immediate, by OD/UD-perfect and Step 3, and property (3) of Lemma 6, which precludes any OD pairs of other hospitals matching with a UD pair of h.

(ii) Immediate, because there are no matches that do not involve an OD pair of h in Steps 3 or 4 of xCM by property (3) of Lemma 6.

Left to argue is that truthful reporting maximizes the number of S pairs that match in Step 1. Consider some type T–T for some $T \in \{O, A, B, AB\}$. First, if $\tau_\oplus$ is even then with the truthful strategy, h matches all T–T pairs by Lemma 6 property (1) (ii), and truthful reporting is optimal. Consider $\tau_\oplus$ is odd, and two cases. (Case 1) τ_h is even. In this case, there is at least one $j \neq h$ with an odd parity number of T–T pairs and h will match all T–T pairs with non-zero probability when reporting truthfully by Lemma 6 property (1) (iii). It is worse to send an odd number and hold back an odd number, because all those sent by h would match by Lemma 6 property (1) (ii) and at least one would be unmatched in the recourse action. Holding back an even number of T–T is worse because this reduces the number of maximum matchings computed in Step 1 of xCM that include all T–T pairs of h, reducing the probability that μ_T includes all reported T–T pairs (since a maximum matching is selected at random). (Case 2) τ_h is odd. If at least one $j \neq h$ has an odd parity number of T–T pairs then the analysis proceeds as in case 1, but considering sending an even number and holding back an odd number, and then holding back an even number. Otherwise, if all the other hospitals have an even number of T–T pairs then it is possible that h has one reported pair unmatched by xCM with probability one. But now if h sends an even number it matches all it sends by Lemma 6 property (1) (ii) but necessarily fails to match at least one of the odd number it holds back. Sending an odd number and holding back an even number leaves unchanged that one reported pair is unmatched by xCM with probability one.

(R pairs) In regard to R pairs, there are no matches that do not involve an OD pair in h in Steps 3 or 4 of xCM by property (3) of Lemma 6. Left to argue is that truthful reporting maximizes the number of R pairs matched in Step 2. Assume without loss of generality that $\alpha_\oplus \geq \beta_\oplus$, and further assume that $\alpha_h > \beta_h$ and $\alpha_\oplus > \beta_\oplus$, because hospital h will otherwise match all pairs by Lemma 6 (2 (ii)) when reporting truthfully. Given this, then by Lemma 6, (2) (ii), when reporting truthfully the hospital matches in expectation $2\beta_h + \min(\alpha_h - \beta_h, g_{-h}(x_{BA}))$ R pairs, where shorthand $g_{-h}(x_{BA}) = g_{uniform}(z_{-h,AB}, x_{BA})$, and x_{BA} is the total number of excess supply of B–A pairs reported by other hospitals. The interesting case is when $\alpha_h - \beta_h > g_{-h}(x_{BA})$ (otherwise it again matches all its R pairs), so that truthful reporting matches an expected number of $2\beta_h + g_{-h}(x_{BA}) = \#truth$ pairs.

Consider some non-truthful report of R pairs. Let m'_h denote the number of R pairs of each type matched before additional assignment from the uniform rule; i.e., the quantity m^*_h in xCM. Let $q'_{h,AB} = m'_h + c_{AB}$ and $q'_{h,BA} = m'_h + c_{BA}$ denote the expected, total number of reported A–B and B–A pairs of h matched in Step 2, where $c_{AB}, c_{BA} \geq 0$ denote additional matches through the use of the uniform rule. We proceed by case analysis:

(Case 1) $\alpha'_\oplus \geq \beta'_\oplus$, so that A–B pairs remain on the long-side. In this case, the best possible outcome for h is that it matches an expected number, $2(\beta_h - q'_{h,BA}) + q'_{h,AB} + q'_{h,BA} = 2\beta_h + q'_{h,AB} - q'_{h,BA} = 2\beta_h + c_{AB} - c_{BA}$ R pairs in total, where

this represents matching every unmatched B–A pair with an A–B pair in Step 3 and recourse. The adjusted supply of excess B–A pairs in the uniform rule is $x_{BA} + c_{BA}$, i.e., the same number from other hospitals and any additional pairs of h that are in excess supply and matched. Based on this, we have $c_{AB} = g_{-h}(x_{BA} + c_{BA}) \leq g_{-h}(x_{BA}) + c_{BA}$, where the inequality is by Lemma 3, property (1). Therefore the best possible outcome is $\beta_h + c_{AB} - c_{BA} \leq 2\beta_h + g_{-h}(x_{BA}) = \#truth$.

(Case 2) $\alpha'_\oplus < \beta'_\oplus$, so that B–A pairs are now on the long-side, and all reported A–B pairs match, by the R-perfect property. Because the number of A–B pairs and B–A pairs matched must be equal, we have:

$$\alpha'_h + \sum_{j \neq h} \alpha_j \leq q'_{h,BA} + \sum_{j \neq h} \beta_j \tag{A.1}$$

where the right-hand side is the number of B–A pairs of h matched and the total number of B–A pairs in the supply from others. This simplifies to,

$$\alpha'_h - q'_{h,BA} \leq \sum_{j \neq h} \beta_j - \sum_{j \neq h} \alpha_j \leq \sum_{j \neq h} \beta_j - \sum_{j \in L} q_{j,AB} - \sum_{j \notin \{L \cup h\}} \alpha_j, \tag{A.2}$$

where $q_{j,AB} \leq \alpha_j$ is the expected number of A–B pairs matched to j at truth, and $L = \{j \neq h, \alpha_j > \beta_j\}$ are the hospitals long A–B. But now we can also write the balance of A–B pairs and B–A pairs matched at truth,

$$\beta_h + g_{-h}(x_{BA}) + \sum_{j \in L} q_{j,AB} + \sum_{j \notin \{L \cup h\}} \alpha_j = \sum_j \beta_j, \tag{A.3}$$

where all B–A pairs are matched and the left-hand side is the expected number of A–B pairs matched. This simplifies to, $g_{-h}(x_{BA}) = \sum_{j \neq h} \beta_j - \sum_{j \in L} q_{j,AB} - \sum_{j \notin \{L \cup h\}} \alpha_j$, and so from (A.2) we have $\alpha'_h - q'_{h,BA} \leq g_{-h}(x_{BA})$. Now, the best possible outcome for h in this case is that it matches, in expectation, $2(\beta_h - q'_{h,BA}) + \alpha'_h + q'_{h,BA} = 2\beta_h + \alpha'_h - q'_{h,BA} \leq 2\beta_h + g_{-h}(x_{BA}) = \#truth$. The first term assumes that all B–A pairs that are unmatched can be matched with A–B pairs in Step 3 or a recourse action, and adds the number of A–B and B–A pairs matched.

(OD-UD pairs) Finally, note that the truthful strategy matches all OD pairs with UD pairs in Step 3 because of the OD/UD-perfect property of each individual hospital's compatibility graph.

For 2-way efficiency, by Lemma 6 the matching is maximum on the subgraphs of the combined graph restricted to T–T pairs (for all blood-types T) and restricted to R pairs. This follows from the S-perfect and R-perfect properties. In addition, every OD pair is matched to a UD pair. The 2-way efficiency property follows by Theorem 1, since the matching is regular on the combined graph. □

Appendix B. Detail for tests of PM properties

In matchings with only 2-cycles, we compute the violations of the PM properties in simulation as follows:

- For the R-perfect property, we obtain G_R the R-subgraph of $\widetilde{G}_n$ and then remove a subset of the long-side at random such that the resulting R-subgraph is balanced. We perform a random matching on the balanced subgraph and then record the number of unmatched R pairs as a fraction of the number of pairs in a maximum matching in $\widetilde{G}_n$.
- For the OD/UD-perfect property, we compute a maximum matching on the subgraph of OD and UD pairs and record the number of unmatched OD pairs as a fraction of the number of pairs in a maximum matching in $\widetilde{G}_n$.
- For the S-perfect property, we obtain G_S, the S-subgraph of $\widetilde{G}_n$, and for every component T–T, we remove one pair at random if there is an odd number of pairs and perform a maximum matching. We record the total number of unmatched S pairs (determined in this way) as a fraction of the number of pairs in a maximum matching in $\widetilde{G}_n$.

In the presence of 3-cycles, we compute the violations of the assumed PM properties in our simulations as follows:

- For the R-perfect property, we obtain the extended R-subgraph of $\widetilde{G}_n^3$ and remove R pairs (if needed) in order to make a balanced graph of R and virtual R pairs. For example if $\#A–B > \#B–A + \#virtual\ B–A$, we will remove $(\#A–B - \#B–A - \#\ virtual\ B–A)$ pairs from the A–B pairs, uniformly at random. Similar to 2-cycles, we compute a maximum matching allowing for 3-cycles that maximize the number of matched R pairs, and record the number of unmatched R pairs as a fraction of the number of pairs in a maximum matching on $\widetilde{G}_n^3$.
- For the OD/UD-perfect property, we first obtain all UD and AB–O pairs that can be maximally matched in OUU matches. We remove these pairs from $\widetilde{G}_n^3$, and then keep only the remaining OD/UD pairs and perform a *2-cycle* maximum matching. We record the number unmatched OD pairs in the final step and number of unmatched AB–O pairs in the first step, as a fraction of the number of pairs in a maximum matching on $\widetilde{G}_n^3$.
- For the S-perfect property, we obtain G_S, the S-subgraph of $\widetilde{G}_n^3$, and for every component T–T, we perform a maximum matching with 3-cycles. We record the total number of unmatched S pairs (computed in this way) as a fraction of the number of pairs in a maximum matching on $\widetilde{G}_n^3$.

Appendix C. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.geb.2015.01.001.

References

Abraham, D.J., Blum, A., Sandholm, T., 2007. Clearing algorithms for barter exchange markets: enabling nationwide kidney exchanges. In: Proceedings of the 8th ACM Conference on Electronic Commerce. ACM, p. 304.

Ashlagi, I., Roth, A.E., 2014. Free riding and participation in large scale, multi-hospital kidney exchange. Theoretical Econ. 9, 817–863. http://dx.doi.org/10.3982/TE1357, http://econtheory.org/ojs/index.php/te/article/viewFile/20140817/11620/358.

Ashlagi, I., Fischer, F., Kash, I.A., Procaccia, A.D., 2013. Mix and match: a strategyproof mechanism for multi-hospital kidney exchange. Games Econ. Behav. http://dx.doi.org/10.1016/j.geb.2013.05.008, http://www.sciencedirect.com/science/article/pii/S089982561300081X.

Bertsimas, D., Farias, V.F., Trichakis, N., 2013. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. Operations Res. 61 (1), 73–87.

Delmonico, F.L., Arnold, R., Scheper-Hughes, N., Siminoff, L.A., Kahn, J., Youngner, S.J., 2002. Ethical incentives for organ donation. N. Engl. J. Med. 346 (25), 2002–2005.

Dickerson, John P., Procaccia, Ariel D., Sandholm, Tuomas, 2012a. Dynamic matching via weighted myopia with application to kidney exchange. In: AAAI.

Dickerson, John P., Procaccia, Ariel D., Sandholm, Tuomas, 2012b. Optimizing kidney exchange with transplant chains: theory and reality. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol. 2. International Foundation for Autonomous Agents and Multiagent Systems, pp. 711–718.

Ehlers, L., Klaus, B., 2003. Probabilistic assignments of identical indivisible objects and uniform probabilistic rules. Rev. Econ. Design 8 (3), 249–268.

Erdos, P., Rényi, A., 1964. On random matrices. Magyar Tud. Akad. Mat. Kutató Int. Közl 8, 455–461.

Feynman, Richard Phillips, Leighton, Robert B., Sands, Matthew Linzee, 1964. The Feynman Lectures on Physics, vol. 1. Addison–Wesley Publishing Company.

Montgomery, R.A., Zachary, A.A., Ratner, L.E., Segev, D.L., Hiller, J.M., Houp, J., Cooper, M., Kavoussi, L., Jarrett, T., Burdick, J., et al., 2005. Clinical results from transplanting incompatible live kidney donor/recipient pairs using kidney paired donation. JAMA 294 (13), 1655.

Moulin, H., 2002. The proportional random allocation of indivisible units. Soc. Choice Welfare 19, 381–413.

Rees, M., Pelletier, R., Mulgaonkar, S., Laskow, D., Nibhanupudy, B., Kopke, J., Roth, A., Ünver, U., Sandholm, T., Rogers, J., 2008. Report from a 60 transplant center multiregional kidney paired donation program: 2. Transplantation 86 (2S), 1.

Rees, M.A., Kopke, J.E., Pelletier, R.P., Segev, D.L., Rutter, M.E., Fabrega, A.J., Rogers, J., Pankewycz, O.G., Hiller, J., Roth, A.E., 2009. A nonsimultaneous, extended, altruistic-donor chain. N. Engl. J. Med. 360 (11), 1096.

Ross, L.F., Woodle, E.S., 2000. Ethical issues in increasing living kidney donations by expanding kidney paired exchange programs. Transplantation 69 (8), 1539.

Roth, A.E., Sönmez, T., Ünver, M.U., 2005. Notes on forming large markets from small ones: participation incentives in multi-center kidney exchange. Unpublished.

Roth, A.E., Sönmez, T., Ünver, M.U., 2007. Efficient kidney exchange: coincidence of wants in markets with compatibility-based preferences. Amer. Econ. Rev. 97 (3), 828–851.

Saidman, S.L., Roth, A.E., Sönmez, T., Ünver, M.U., Delmonico, F.L., 2006. Increasing the opportunity of live kidney donation by matching for two- and three-way exchanges. Transplantation 81 (5), 773.

Sasaki, H., 1997. Randomized uniform allocation mechanism and single-peaked preferences of indivisible good. Technical report. Waseda University, Japan.

Sprumont, Y., 1991. The division problem with single-peaked preferences: a characterization of the uniform allocation rule. Econometrica 59 (2), 509–519.

Toulis, Panos, Parkes, David C., 2011. A random graph model of kidney exchanges: efficiency, individual-rationality and incentives. In: Proc. ACM Conference on Electronic Commerce. EC'11.

Ünver, M.U., 2010. Dynamic kidney exchange. Rev. Econ. Stud. 77 (1), 372–414.

Wikipedia, 2010. Blood type [online; accessed 22-December-2010].

Zenios, S.A., Woodle, E.S., Friedman Ross, L., 2001. Primum non nocere: avoiding harm to vulnerable wait list candidates in an indirect kidney exchange. Transplantation 72 (4), 648.