# CORE-STABLE KIDNEY EXCHANGE VIA ALTRUISTIC DONORS

GERGELY CSÁJI, THÀNH NGUYEN,

Abstract. Kidney exchange creates gains by pooling patient–donor pairs across hospitals and countries, but cooperation may unravel when coalitions can profitably withdraw. We introduce the *supplemented core*, in which the platform uses voluntarily registered altruistic donors to restore stability. Because an altruistic donor adds a kidney without adding another patient, the donor provides an in-kind instrument for relaxing participation constraints when monetary transfers are unavailable. In worst-case compatibility graphs, the number of donors required can grow linearly with market size, even with arbitrary cycle length bounds. Under the standard random-graph model, however, a logarithmic number suffices for pairwise and bounded cyclic exchange, in expectation and with high probability. Calibrated simulations show that against coalitions of at most four organizations the weak-core is either already nonempty or can be restored with a single additional donor. The more consequential practical problem is algorithmic: standard operational rules may select an unstable exchange even when a stable exchange exists. A small number of altruistic donors also closes this gap. Altruistic donors can therefore serve as a stability reserve, sustaining cooperation rather than merely increasing transplants.

## 1. Introduction

Kidney exchange creates substantial gains by bringing together patients who have willing but incompatible living donors. These gains depend on market thickness. A larger common pool creates more opportunities for compatible cycles and chains, but it also requires hospitals to place their donor–patient pairs in the common market. In practice, kidney exchange remains fragmented. Hospitals may withhold pairs because participation involves administrative, testing, and coordination costs (Agarwal et al. 2019). Even if these costs are reimbursed, a hospital may prefer to conduct some exchanges internally rather than accept the allocation chosen by the centralized platform. The same issue arises in international kidney exchange, where countries decide whether to participate in a common program (Biró et al. 2019). The problem is therefore not only how to find more transplants, but also how to make participation attractive to the organizations that control the donor–patient pairs.

The core provides a natural benchmark for this participation problem. An exchange is in the weak core if no coalition of organizations can use the donor–patient pairs and altruistic donors under its control to make every member strictly better off. Core stability is stronger than individual rationality: it protects not only against the withdrawal of one organization, but also against coordinated withdrawal by several organizations. This is important when hospitals or countries can rematch their pairs outside the centralized program. The difficulty is that the weak core may

*Date*: August 21, 2026.
Nguyen: Daniels School of Business, Purdue University; nguye161@purdue.edu.
Csáji: HUN-REN KRTK KTI, and Eötvös Loránd University; csaji.gergely@krtk.hun-ren-hu.

be empty. Because kidney exchange is discrete, there may be no feasible exchange that satisfies all coalitional participation constraints.

This paper proposes a new way for the platform to restore stability: supplement selected exchanges with voluntarily registered altruistic donors. Altruistic donors are usually valued for initiating chains and increasing the number of transplants; we identify a distinct role. Because an altruistic donor contributes a kidney without adding another patient who must receive one, the donor gives the platform additional transplant capacity that can be used to satisfy an organization's participation constraint. Since monetary payments are prohibited and cannot substitute for a kidney, this capacity acts as an in-kind subsidy.[1]

The platform therefore faces a resource-allocation problem: it must deploy a scarce reserve of altruistic donors to satisfy the participation constraints that threaten cooperation, while using as few donors as possible. One additional transplant can have a disproportionate effect because the same organization may be pivotal in several potential blocking coalitions; improving that organization's allocation can eliminate several deviations at once. Our analysis characterizes the size of the donor reserve needed to guarantee a core-stable exchange and identifies when this requirement remains small relative to the size of the market.

We formalize this question through the *supplemented core*. We say that an exchange belongs to the $d$-supplemented core if the centralized platform can use at most $d$ additional altruistic donors and the resulting exchange is not blocked by any coalition of participating organizations. The usual core is the special case $d = 0$. The smallest value of $d$ measures how many additional donors are needed to make stable participation feasible.

We model kidney exchange as a directed graph. Each vertex represents a donor–patient pair or a non-directed altruistic donor, and a directed edge indicates that one donor can give a kidney to another vertex's patient. Each organization controls a distinct set of donor–patient pairs and may also control its own altruistic donors. Feasible exchanges consist of cycles and chains, and each organization values an exchange by the number of its patients who receive a transplant.

The additional donors used in the supplemented core are coordinated by the centralized platform, but they are not owned by it. They should be interpreted as non-directed donors who voluntarily register with the centralized program and agree to have their recipients selected through its matching process. This arrangement is already used in practice. For example, non-directed donors may enter the OPTN Kidney Paired Donation Pilot Program independently, while the national program performs the matching and participating transplant centers remain responsible for evaluation and transplantation.[2] Organizations may also have altruistic donors registered with them; these donors remain part of their endowments and are available when coalitions consider deviating. The additional donors are unavailable to a deviating coalition because they registered with the common program rather than with any member of that coalition. As in other static models of kidney exchange, the analysis is conditional on the set of donors who remain active at the time of the match. Every donor retains the right to withdraw.

[1]The analogy is to injecting outside resources to relax the no-deficit constraint underlying the Myerson–Satterthwaite impossibility (Myerson and Satterthwaite 1983).

[2]See the OPTN Kidney Paired Donation Pilot Program, https://www.hrsa.gov/optn/patients/organ-donation/living-donation/kidney-paired-donation-pilot-program-kpdpp.

We present three main theoretical results and use simulations to study their practical implications. The analysis separates two sources of instability. The first is intrinsic: a stable integral exchange may not exist. The second is algorithmic: a stable exchange exists, but the platform's allocation rule selects a different exchange that can be blocked. Additional altruistic donors address both problems. They restore existence when the core is empty and provide the flexibility needed to stabilize the outcome selected by an operational allocation rule.

Our first result shows that the problem can be severe. Without restrictions on the compatibility graph, the number of additional donors required for stability can grow linearly with the size of the market. This lower bound holds for pairwise exchange, for three-way exchange, and even for arbitrary cycle length bounds. Core nonexistence is therefore not merely a consequence of short cycle-length limits. In worst-case instances, the trivial policy of providing a separate donor for every unmatched patient is optimal up to a constant factor.

Our second result shows that in pairwise exchange, the number of additional donors required is at most the maximum number of *independent odd cycles* in the mutual compatibility graph. A collection of vertex-disjoint odd cycles is independent if no mutual compatibility edge connects different cycles. These cycles capture the obstruction to converting a fractional stable allocation into an integral exchange.

This characterization yields strong guarantees for economically relevant compatibility structures. Consider a finite-group random compatibility model in which donor–patient pairs are grouped by observable characteristics such as blood type, antibodies and sensitization. A directed compatibility edge from a donor in group $a$ to a patient in group $b$ is generated independently with probability $p_{ab}$. The probabilities may differ across groups, some may be zero because of medical incompatibility, and group sizes may be unequal. This is a standard and analytically tractable framework for representing the principal sources of heterogeneity in kidney-exchange pools. If the number of groups is fixed and all positive compatibility probabilities are bounded away from zero, the maximum number of independent odd cycles is $\mathcal{O}(\log |V|)$ in expectation and with probability approaching one. Hence, only $\mathcal{O}(\log |V|)$ additional donors are required. If the mutual compatibility graph is bipartite, there are no odd-cycle obstructions, and the weak core is nonempty without additional donors. We show that the bipartite result applies to a broad class of one-sided exchange economies with *binary assignment valuations*, extending Echenique et al. (2024).

Our third result extends the logarithmic guarantee beyond pairwise exchange. For any fixed cycle-length bound $\Delta \geq 3$, under the same heterogeneous random compatibility model, the expected number of additional donors is still $\mathcal{O}(\log |V|)$, and the same bound holds with probability approaching one. The implicit constant depends on $\Delta$, the number of groups, and the smallest positive compatibility probability, and may be larger than in the pairwise case. The important point is that the donor requirement remains logarithmic rather than linear in market size. Thus, the linear worst-case lower bound is not representative of this broad class of random compatibility environments, even when exchanges involve longer cycles.

The theoretical bounds count the suitably compatible donors used in the final exchange; they do not assume that the platform can choose donor characteristics. In practice, the required compatibilities must be found among donors who voluntarily register, so the number screened may

exceed the number ultimately used. If each newly registered donor has an independent probability bounded away from zero of satisfying a required compatibility, then the expected number screened is at most a constant multiple of the number used. The logarithmic bounds therefore remain logarithmic when expressed in terms of the size of the donor pool. Our simulations make no favorable selection assumption: altruistic donors are generated from empirically calibrated distributions and introduced in random order.

We complement the theoretical analysis with simulations based on a state-of-the-art kidney-exchange instance generator (Delorme et al. 2022, Pettersson and Trimble 2021). We first ask whether there exists a maximum-cardinality exchange that is stable against deviations by coalitions of at most four organizations. We impose this cap for both institutional and computational reasons. Coordinating a joint withdrawal becomes increasingly difficult as more hospitals or countries must agree on a deviation, while exhaustively testing all coalitions becomes computationally prohibitive as the number of organizations grows. Under this benchmark, we find a maximum-cardinality exchange that is weak-core stable against these coalitions without additional donors in 299 of the 300 simulated markets. In the remaining market, one randomly generated donor suffices. The required supplementation remains small under the stronger notions of the strong core and the transferable-utility core, evaluated under the same coalition-size cap. Thus, although core nonexistence can be severe in worst-case graphs, instability against practically relevant small coalitions is rare in the realistic instances we study.

Existence, however, does not imply stable implementation. Kidney-exchange programs typically select among feasible exchanges using a hierarchy of objectives, such as maximizing the number of transplants, favoring shorter cycles, preserving same-blood-type transplants, and matching hard-to-match patients (Biró et al. 2021). When we apply a representative lexicographic hierarchy without imposing core constraints, the selected exchange is blocked by a coalition of at most four organizations in as many as 20% of the simulated markets in some settings. Because a donor-free weak-core exchange exists in almost all of these markets, these failures arise primarily from algorithmic selection rather than from core nonexistence. A market may possess a stable and efficient exchange and nevertheless be operated unstably.

We next retain the same lexicographic hierarchy, test its output for blocking coalitions, and introduce additional altruistic donors only when the selected exchange is unstable. After each donor is added, we rerun the same allocation rule. The average number of additional donors required is at most 0.6 in every parameter setting, and at most three donors suffice in all but one simulated market; the exceptional market requires six. Supplementation therefore closes the gap between stable existence and stable implementation without requiring the platform to replace or reorder its established medical and operational objectives. It acts as a stability layer around the existing allocation rule.

We implement these procedures through an integer-programming framework. Our theoretical guarantees rely on Scarf's lemma and a rounding argument and therefore do not directly provide a practical algorithm. The computational framework uses a cycle formulation and searches for blocking coalitions. For the core-aware procedure, it iteratively adds the associated stability constraints and introduces altruistic donors when the constrained problem becomes infeasible. For

the practice-based procedure, it leaves the lexicographic rule unchanged, verifies its output, and adds donors only when a blocking coalition is found. For computational tractability, we restrict the stability tests to coalitions up to a specified size. The framework solves large instances efficiently on standard hardware and can certify stability against every coalition containing at most four organizations.

Taken together, the results show that stability is a property of both the market and its allocation rule. The compatibility graph determines whether a stable exchange exists; the algorithm determines whether one is selected. Supplementation repairs both failures: it restores existence when the core is empty and stabilizes the outcome when the operational rule selects outside the core.

This gives altruistic donors a role beyond creating transplants and initiating chains: they provide the flexibility needed to sustain participation by hospitals and countries. Approximately 300–400 altruistic donors participate annually in the United States, compared with about 6,000 donor–patient pairs in the UNOS database; the NKR data studied by Agarwal et al. (2019) contain 164 altruistic donors and 1,265 pairs. Our proposal requires neither centralized ownership nor compulsory transfer of these donors. It requires only that some non-directed donors voluntarily participate through the common platform. A small subset can then serve as a stability reserve when the selected exchange would otherwise be blocked. The organizational value of such a donor can exceed the single transplant directly enabled: by preventing withdrawal, the donor preserves the larger gains from pooling.

The paper is organized as follows. After discussing related work, we introduce the model, the supplemented core, and the methodological preliminaries used to establish our results. Section 3 provides worst-case lower bounds on the number of altruistic donors required. Section 4.1 studies pairwise exchange, while Section 4.2 analyzes cyclic exchange. Section 5 presents the computational analysis. Section 6 presents the application to one-sided exchange economies. Section 7 concludes. Proofs omitted from the main text are collected in the Appendix.

**Related Literature.** The success of kidney exchange has inspired a rich literature on market design, beginning with Roth et al. (2004). This literature studies different exchange structures, including pairwise and cyclic exchanges, and different strategic actors, ranging from individual donor–patient pairs to hospitals and countries. Incentives at the individual level are relatively well understood (Roth et al. 2004, Sönmez and Ünver 2013). Participation becomes more difficult when organizations control several pairs and can conduct exchanges outside the centralized program (Agarwal et al. 2019, Ashlagi and Roth 2012, 2014).

The main difference between our paper and the literature on kidney exchange among organizations is our positive approach to core stability. Existing work either emphasizes the possible nonexistence of the core or imposes weaker participation requirements. For example, Ashlagi and Roth (2014) focus on individual rationality, while Ashlagi et al. (2015) study incentive problems and multiplicative efficiency guarantees without imposing coalitional stability. Our model builds most directly on Csáji et al. (2024) and Kern et al. (2019), which establish negative results for the core. To the best of our knowledge, no previous paper quantifies how much additional donor supply is sufficient to restore the core in kidney exchange.

Our approach also relates to credit systems for international kidney exchange (Kern et al. 2019, Klimentova et al. 2021). These systems assign transplant targets to participating countries and carry deviations from the targets forward as credits. Credits address fairness across matching rounds. We instead study stability within a matching round: conditional on the currently available pairs and donors, no coalition should gain by leaving the program. The two approaches are complementary. Intertemporal credits compensate countries in future rounds, while altruistic donors provide additional flexibility in the current round.

There is also a large literature on dynamic kidney exchange, including Akbarpour et al. (2020), Anderson et al. (2017), Dickerson and Sandholm (2014). This literature mainly studies trade-offs among efficiency, waiting time, and market thickness. Our model is static because we focus on whether the allocation in a given matching round satisfies coalitional participation constraints.

From a computational perspective, the literature has developed several methods for finding good kidney-exchange allocations under operational constraints (Ashlagi et al. 2024, Druzsin et al. 2024, Mincu et al. 2021). These methods are mainly evaluated through simulations and generally do not provide theoretical guarantees for core stability. We obtain our theoretical results using the Relax-and-Round approach of Nguyen et al. (2021), Nguyen and Vohra (2018). Scarf's lemma provides a fractional stable benchmark, which we round to an integral exchange while bounding the additional donor supply required. Unlike work in which agents have small endowments (Jantschgi et al. 2025), hospitals in our model may control many donor–patient pairs. The rounding argument must therefore use the structure of the compatibility graph and the restrictions on feasible exchanges. We complement these existence results with an integer-programming framework for computing and verifying stable exchanges in realistic instances.

## 2. Model and Preliminaries

In this Section we describe our mathematical model for Kidney Exchange, motivated by the works of Kern et al. (2019) and Csáji et al. (2024). Unlike Roth et al. (2004), which models each agent as a donor-patient pair, we consider agents as organizations, such as hospitals in the U.S. exchange system or organizations in the international exchange system. We consider an exchange involving a set of $n$ organizations, denoted by $N = \{1, \dots, n\}$. Each organization $i \in N$ is endowed with a set $V^i$ consisting of donor-recipient pairs, unpaired patients, and altruistic donors. The sets $V^i$ are disjoint across organizations. Within each $V^i$, let $U^i \subset V^i$ denote the set of non-altruistic donor-patient pairs and unpaired patients. The remaining elements, $V^i \setminus U^i$, are altruistic donors willing to donate without requiring a kidney in return.

Compatibility graphs. Compatibility is modeled via a directed graph, denoted by $\mathcal{G} = (V, E)$, where $V = V^1 \cup .. \cup V^n$, and the set of edges, $E$, captures the compatibility of exchange. Specifically,

- For two non-altruistic donors—patient-donor pairs $u, v$, a directed edge $(u, v) \in E$ indicates that the donor in pair $u$ is compatible with the patient in pair $v$, meaning a transplant from $u$ to $v$ is feasible.
- To simplify notation and the description of feasible exchanges, we can model altruistic donors as special donor–patient pairs by introducing a dummy patient who is compatible with any donor. This allows us to define directed edges following the same rules as before.

- Similarly, we can model unpaired patients by adding a dummy donor who is only compatible with the dummy patient associated with the altruistic donors.

We examine the following model of *random compatibility graphs* (Delorme et al. 2022). Donor-recipient pairs are partitioned into a finite set of groups $\Phi$, based on characteristics such as blood type and the sensitization levels of both the donor and the patient.

For any two distinct donor-recipient pairs $u$ and $v$ belonging to groups $a$ and $b$, respectively, the directed compatibility edge $(u, v)$ is added independently with probability $p_{ab}$. Self-arcs $(u, u)$ are excluded, but we impose no restriction on the diagonal probabilities $p_{aa}$, which describe compatibility between distinct pairs belonging to the same group. We write

$$p := \min\{p_{ab} : p_{ab} > 0\}.$$

The number of groups $g := |\Phi|$ and the minimum positive probability $p$ are treated as constants.

Exchanges. We consider two models of kidney exchange: **pairwise exchange** and **cyclic exchange**. In pairwise exchange, only mutually compatible donor-patient pairs can trade kidneys. Cyclic exchange allows longer cycles, where each donor gives to the next patient and receives a kidney from the previous donor. We assume a maximum cycle length of $\Delta$, with $\Delta = 2$ corresponding to pairwise exchange. In practice, $\Delta = 3$ is typical, while values above 5 are rare due to the logistical and medical challenges of coordinating multiple simultaneous surgeries.[3]

Let $\mathcal{C}_\Delta$ denote the cycles of length at most $\Delta$. Given a subset of vertices $W \subseteq V$, denote the induced compatibility graph on $W$ as $\mathcal{G}[W]$. We call $\mathcal{E}$ an exchange among $W$ if $\mathcal{E}$ is a union of disjoint cycles of length at most $\Delta$ in the induced compatibility graph $\mathcal{G}[W]$. We say that $y^* \in [0,1]^{\mathcal{C}_\Delta}$ is a *fractional exchange* if it holds that for any vertex $v \in V$, $\sum_{c \in \mathcal{C}_\Delta | v \in c} y^*_c \leq 1$.

Given an exchange $\mathcal{E} = (V_\mathcal{E}, E_\mathcal{E})$, the utility of an organization $i$ is given by the number of their patients receiving a kidney, that is

$$u_i(\mathcal{E}) := |V_\mathcal{E} \cap U^i|.$$

Overall, we refer to the economy as a *partition exchange economy*, represented by the tuple $(\mathcal{G}, \Delta, \{V^i, U^i\}_{i=1}^n)$.

Definition 2.1. *Given a partition exchange economy, an exchange $\mathcal{E}$ is* individually rational, *if $u_i(\mathcal{E}) \geq u_i(\mathcal{E}_i)$ for any $i \in N$ and exchange $\mathcal{E}_i \subseteq \mathcal{G}[V^i]$. That is, no player can improve its utility with an exchange of its own.*

Definition 2.2. *Given a partition exchange economy, an exchange $\mathcal{E}$ among $V$, and a coalition $P \subseteq N = \{1, .., n\}$, we say that $P$ is a* (strong) blocking coalition *to $\mathcal{E}$, if there exists an exchange $\mathcal{E}'$ in the graph $\mathcal{G}[\cup_{i \in P} V_i]$ such that $u_i(\mathcal{E}') > u_i(\mathcal{E})$ for all $i \in P$. We say that $\mathcal{E}$ is in the (weak) core, if there is no blocking coalition to $\mathcal{E}$.*

[3]Even in the case of altruistic donors, cycles become chains, and operational difficulties with chains are less severe. Longer chains are also rare in practice (Agarwal et al. (2019)). Moreover, chains of size 10–20 typically result from the accumulation of exchanges over an extended period. For example, the National Kidney Registry reported a 25-transplant chain that was built up over several months. Our model aims to capture the clearing mechanism over a relatively shorter period, during which longer chains can be broken into smaller segments. The donor at the end of a chain in one period can then be viewed as an altruistic donor initiating a new chain in the next period.

Note that in the definition above, some papers in the literature refer to such a block as a strong block and the corresponding concept as the weak core. For convenience, in this paper we often refer to the weak core as simply the core.

As we show in this paper, the core can be empty in both pairwise and cyclic exchanges. To obtain a positive result, we either impose a restriction on the compatibility graph—discussed in Section 6—or relax the notion of the core by introducing what we call the supplemented core. To this end, we introduce a special agent, indexed by 0, representing the centralized market designer, with $V^0$ denoting the designer's endowment of *altruistic* donors. Let $\mathcal{G}^{+V^0}$ be the extended compatibility graph that incorporates the additional donors in $V^0$. The designer *is not a player of the game and is not allowed to participate in any coalition.* Their sole role is to stabilize the exchange by providing additional resources when necessary.

Next we introduce our main concept.

Definition 2.3. *Given additional altruistic donors $V^0$, an exchange $\mathcal{E}^*$ in the extended compatibility graph $\mathcal{G}^{+V^0}$ is a $V^0$-supplemented core if it is not blocked by any coalition $P \subseteq N = \{1, .., n\}$. If $|V^0| \leq d$, we also refer to $\mathcal{E}^*$ as a $d$-supplemented core exchange.*

Definition 2.4. *Given a set of additional altruistic donor $V^0$, an exchange $\mathcal{E}^*$ in the extended compatibility graph $\mathcal{G}^{+V^0}$ is* Pareto optimal *if there does not exist another exchange $\mathcal{E}'$ in $\mathcal{G}^{+V^0}$ such that all organizations are weakly better off and at least one organization is strictly better off.*

We show in a simple example how an additional altruistic donor can guarantee the existence of a core exchange.

Example 2.1. *Consider the partition economy illustrated in Figure 1, featuring three organizations represented by red triangle, green square, and blue circle, respectively. This instance, from Csáji et al. (2024), demonstrates an empty core. However, adding a single altruistic donor (the black vertex) makes the bold matching $\mathcal{E}$ a core outcome. To see this, observe that any organization alone can cover all but 3 of its vertices, any two organizations can cover all but 2 of their vertices and the three organizations together can cover all but 5 of their vertices. Since in the bold $V^0$-supplemented exchange $\mathcal{E}$, one organization has 2 vertices unmatched, and the others only one, no single organization, nor the 3 organizations together block $\mathcal{E}$. Furthermore, no 2 organizations block either, as then they could leave at most $(2-1)+(1-1)=1$ of their vertices unmatched, which is impossible.*

Remark 2.1. *As this example illustrates, and as our main results formalize, stability can be restored by adding altruistic donors who are compatible with patients left unmatched by the rounding procedure. The deterministic bounds count the compatible donors ultimately used and therefore abstract from the process of finding them. In practice, the platform cannot choose donors' compatibility profiles. Suppose, however, that each newly registered donor, independently of previous arrivals, has a probability bounded away from zero of satisfying any outstanding compatibility requirement. In this case, the expected number of donors that must be screened for each successful match is then bounded by a constant. Consequently, the expected number of donors that must be screened is at most a constant multiple of the number ultimately used, so the logarithmic bounds established below*

*remain logarithmic. Our simulations impose this more realistic restriction: altruistic donors are generated randomly and added in random order rather than selected to have particular compatibility profiles.*

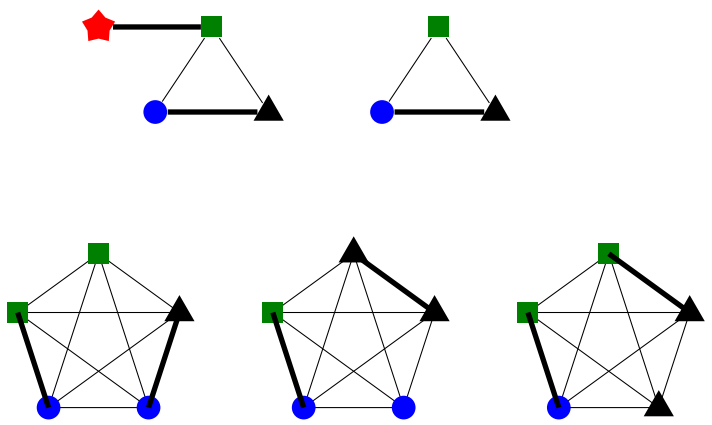


FIGURE 1. A partition exchange economy with pairwise exchanges that has no core but admits a 1-supplemented core. The three organizations are represented by vertices with distinct shape (circle, triangle, square). The star vertex in the upper-left corner represents $V^0$. The bold matching is a supplemented core.

It is well known that in exchange economies with satiated utilities, the strong core is an unreasonably demanding solution concept.

REMARK 2.2. *Consider the following example. Organization $A$ has $k$ altruistic donors and no patients. There are $k+1$ other organizations, $B_1, \ldots, B_{k+1}$, each owning $k$ donor–patient pairs. Every altruistic donor of $A$ is compatible with every patient of every $B_i$, and the donor in each pair owned by a $B_i$ is compatible with the dummy patient associated with every altruistic donor. No exchange is possible among the organizations $B_1, \ldots, B_{k+1}$ without the altruistic donors. In any exchange using only the $k$ donors initially owned by $A$, at least one organization $B_i$ has fewer than $k$ patients matched. The coalition $\{A, B_i\}$ then weakly blocks: $A$ remains indifferent because it has no patients, while $B_i$ can use all $k$ donors of $A$ to match all of its patients. Hence, in any strong-core allocation, every organization $B_i$ must have all $k$ of its patients matched. This requires $k(k+1)$ altruistic donors in total. Since $A$ initially owns $k$ such donors, at least $k^2$ additional altruistic donors are necessary. Conversely, $k^2$ additional universally compatible altruistic donors are sufficient. Thus, the number of additional donors required for the strong core is $\Theta(k^2)$, which is of the same order as the total number of vertices in the original instance, namely $k^2 + 2k$. By contrast, no additional donor is needed for the weak core. Since organization $A$ has no patients, it can never be strictly better off, and no coalition consisting only of the organizations $B_1, \ldots, B_{k+1}$ can perform any exchange. Hence any maximum-cardinality allocation of the original $k$ altruistic donors is Pareto optimal and belongs to the weak core.*

The requirement in Remark 2.2 is clearly unrealistic. Therefore, we focus on the weak core and Pareto optimality, rather than the strong core. Remark 2.2 also shows that to have a strong core

solution, we may need to add $|V^i|$ many additional donors for some of the organizations $i$ – note that adding $|V^i|$ many additional donors for $i$ ensures that $i$ cannot strictly improve anymore.

2.1. **Supplementary Lemmas.** Our results rely on Scarf's lemma together with a rounding argument Nguyen and Vohra (2018). We prove below three technical lemmas that underpin the analysis. The first is Scarf's lemma, the second shows how to round a fractional solution to an integral one, and the third is a basic linear-algebraic fact from linear programming that we use to analyze the quality of the rounding procedure.

For a matrix $Q$, let $Q_{ij}$ denote the $j$-th element of the $i$-th row of $Q$, $Q_i$ denote the $i$-th row of $Q$ and $Q_{\cdot j}$ denote the $j$-th column of $Q$.

LEMMA 2.1 (Scarf (1967)). *Let $Q$ be an $n \times m$ nonnegative matrix, such that every column of $Q$ has a nonzero element and let $q \in \mathbb{R}^n_+$. Suppose that every row $i$ has a strict ordering $\succ_i$ on those columns $j$ for which $Q_{ij} > 0$. Then there is an extreme point of $\{x \in \mathbb{R}^m \mid Qx \leq q,\ x \geq 0\}$, that dominates every column in some row, where we say that $x \geq 0$ dominates column $j$ in row $i$, if $Q_{ij} > 0$, $(Q_i)^T x = q_i$ and $k \succ_i j$ for all $j \neq k \in \{1, \ldots, m\}$, such that $Q_{ik}x_k > 0$. Also, this extreme point can be found algorithmically.*

Let $\gamma^i_c = |U^i \cap c|$ for a cycle $c$ and organization $i$. Recall that $\mathcal{C}_\Delta$ denotes the set of cycles of length at most $\Delta$. The following key Lemma – building on Scarf's Lemma – provides an existential guarantee for a special fractional solution: one that we can round off to find a $V^0$-supplemented core.

LEMMA 2.2 (Rounding Lemma). *Given a partition exchange economy with cycles of size at most $\Delta$, there exists a fractional exchange $y^*$ in $\mathcal{G}$ such that: for any set $V^0$ of altruistic donors and any (deterministic) exchange $\mathcal{E}^*$ in $\mathcal{G}^{+V^0}$, the exchange $\mathcal{E}^*$ belongs to the $V^0$-supplemented core if*

$$u_i(\mathcal{E}^*) \geq \Big\lfloor \sum_{c \in \mathcal{C}_\Delta} \gamma^i_c y^*_c \Big\rfloor \text{ for all } i \in N.$$

*Proof.* Proof. Let us create an instance of Scarf's Lemma as follows. For the matrix $Q$, we create a row for each vertex $v \in V$. Furthermore, for each coalition $\varnothing \neq P \subseteq N$ and each nonzero utility vector $(u_1, \ldots, u_{|P|})$ which is attainable by a feasible exchange $\varnothing \neq \mathcal{E} \subseteq E[P]$ (that is $u_i(\mathcal{E}) = u_i$ for $i \in P$), we create a column $(P, \mathcal{E})$ by choosing an arbitrary such exchange $\mathcal{E}$. The entry in the intersection of a row $v$ and a column $(P, \mathcal{E})$ is 1, if $v$ is covered by $\mathcal{E}$ (also implying that $v \in V^i$ for some $i \in P$ by $\mathcal{E} \subseteq E[P]$) and 0 otherwise. The bounding vector $q$ is chosen to be 1 for all rows. Both are nonnegative, hence they satisfy the conditions of Scarf's Lemma.

Then, we create the strict ranking required for the rows over their nonzero elements. For a vertex $v \in V^i$, this is created in a way such that $(P, \mathcal{E}) \succ_i (P', \mathcal{E}')$, whenever it holds that $u_i(\mathcal{E}) > u_i(\mathcal{E}')$. If $u_i(\mathcal{E}) = u_i(\mathcal{E}')$, then we break the ties in a way such that $(P, \mathcal{E}) \succ_i (P', \mathcal{E}')$, whenever $|P| < |P'|$.

By Scarf's Lemma, there exists an extreme point of this system that dominates every column in some row. Choose this solution $x^*$ in a way such that $|x^*|_1 = \sum_{P,\mathcal{E}} x^*_{P,\mathcal{E}}$ is maximal.

Define $y^*_c = \sum_{P,\mathcal{E} | c \in \mathcal{E}} x^*_{P,\mathcal{E}}$ for all cycles $c \in \mathcal{C}_\Delta$. Then, $\sum_{c | v \in c} y^*_c = \sum_{c | v \in c} \sum_{P,\mathcal{E} | c \in \mathcal{E}} x^*_{P,\mathcal{E}} = \sum_{P,\mathcal{E} | \mathcal{E} \text{ covers } v} x^*_{P,\mathcal{E}} = Q_v x^* \leq 1$, so $y^*$ is indeed a fractional exchange.

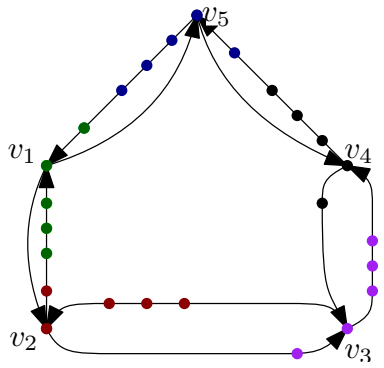



FIGURE 2. An instance, where the core is empty for any $\Delta \geq 6$, including $\Delta = \infty$. Colors denote the organizations.

Let $\mathcal{E}^*$ be an exchange using some $d \geq 0$ altruistic donors, such that $u_i(\mathcal{E}^*) \geq \lfloor \sum_{c \in \mathcal{C}_\Delta} \gamma_c^i y_c^* \rfloor$ for all $i \in N$.

Suppose there is a blocking coalition $P$ with an exchange $\mathcal{E}_P$, where all $i \in P$ improve. By construction, there is an exchange $\mathcal{E}'_P$, such that $P, \mathcal{E}'_P$ has a column in $Q$ and $u_i(\mathcal{E}_P) = u_i(\mathcal{E}'_P)$ for all $i \in P$. Hence, as $x^*$ was a dominated solution, it dominated $P, \mathcal{E}'_P$ in some row $v$. Let $j$ be the organization for which $v \in V^j$. For this $j \in N$, we have that $\sum_{P',\mathcal{E}'} Q_v x^*_{P',\mathcal{E}'} = 1$ and that for any $P', \mathcal{E}'$ with $j \in P'$ and $x^*_{P',\mathcal{E}'} > 0$ it holds that $u_j(\mathcal{E}') \geq u_j(\mathcal{E}'_P)$.

Hence, by the construction of $y^*$, we must have $\sum_{c \in \mathcal{C}_\Delta} \gamma_c^j y_c^* = \sum_{P',\mathcal{E}'} \sum_{c \in \mathcal{E}'} \gamma_c^j x^*_{P',\mathcal{E}'} = \sum_{P',\mathcal{E}'} u_j(\mathcal{E}') x^*_{P',\mathcal{E}'} \geq \sum_{P',\mathcal{E}' | \mathcal{E}' \text{ covers } v} u_j(\mathcal{E}') x^*_{P',\mathcal{E}'} \geq Q_v x^* \cdot \min_{P',\mathcal{E}'} \{u_j(\mathcal{E}') \mid Q_{v,(P',\mathcal{E}')} x^*_{P',\mathcal{E}'} > 0\} \geq u_j(\mathcal{E}'_P)$.

Therefore, we get that $u_j(\mathcal{E}^*) \geq \lfloor \sum_{c \in \mathcal{C}_\Delta} \gamma_c^j y_c^* \rfloor \geq u_j(\mathcal{E}'_P) = u_j(\mathcal{E}_P)$, contradicting that $P$ is a strong blocking coalition. □

We will also use the following well-known lemma from operations research (Schrijver (1998)).

LEMMA 2.3. *Let $z^*$ be an everywhere strictly positive extreme point of $\{Qz \leq q, z \geq 0\}$. Then the number of variables equals the maximum number of linearly independent tight rows of $Q$.*

## 3. Lower Bounds

In this section, we establish lower bounds on the number of additional donors required for pairwise exchange, three-way exchange, unbounded exchange, and $\Delta$-exchange for any fixed $\Delta$.

In the case of pairwise exchange, we can discard all compatibility arcs $(u, v)$ such that $(v, u)$ is not a compatibility arc, and only consider the undirected graph $\mathcal{G}$ of the mutual compatibilities. Hence, in this case, we replace the set $E$ of the edges of $\mathcal{G}$ by the possible pairwise exchanges $\mathcal{C}_2$. Therefore, the feasible exchanges are now assumed to be the matchings of the graph $\mathcal{G}$.

EXAMPLE 3.1. *We give a simple example, where the weak core is empty for any bound $\Delta \geq 6$, in particular even for unbounded exchanges, illustrated in Figure 2. (For pairwise exchanges, a no-instance was given by Csáji et al. (2024)). We have five organizations, $1, 2, 3, 4, 5$, each with five vertices in $V^i$. The compatibility graph is a five-cycle of mutual compatibilities with subdivided arcs,*

*where each organization controls exactly one of the vertices of the five-cycle. The one controlled by $i$, we call $v_i$. Then, $v_2v_1$ is subdivided with four vertices, three belonging to organization 1, and one to organization 2. Arc $v_2v_3$ is subdivided with one vertex, belonging to organization 3, and arc $v_3v_2$ is subdivided with three vertices, each belonging to organization 2. Arc $v_3v_4$ is subdivided with three vertices belonging to organization 3, and arc $v_4v_3$ with one vertex belonging to organization 4. Arc $v_4v_5$ is subdivided with four vertices, three belonging to organization 4, and one to organization 5. Arc $v_5v_1$ is subdivided with four vertices, three belonging to organization 5, and one to organization 1.*

*If one of the large cycles containing all of $\{v_1, v_2, v_3, v_4, v_5\}$ is chosen, then it is the only exchange cycle and either the coalition $\{1, 2\}$ or $\{4, 5\}$ blocks the exchange. Otherwise, if only cycles between consecutive $v_i$-s are chosen, then there exists some $j$ with $v_j$ uncovered. Then, the coalition $\{j-1, j\}$ blocks the exchange.*

More generally, we show that for any bound $\Delta$, $\Omega(|V|)$ donors may be required for the core, even with three organizations.

THEOREM 1. *Fix any exchange bound $\Delta$. Then, there exists a constant $C_\Delta$ and arbitrary large partition exchange economies such that the $C_\Delta V$-supplemented core is empty in them, even if the number of players is $n = 3$.*

*Proof.* **Proof for Pairwise Exchanges.** We create an instance of a partition exchange economy with pairwise exchanges as follows. Let $K_l$ be the clique on $l$ vertices - i.e. the graph where any two vertices are connected. The graph $\mathcal{G}$ of the partition exchange economy consists of $3a$ vertex-disjoint $K_3$ graphs (i.e. triangles) $K_3^1, K_3^2, \ldots, K_3^{3a}$ and $9a$ vertex-disjoint $K_5$ graphs $K_5^1, K_5^2, \ldots, K_5^{9a}$, where $a \in \mathbb{N}$ is a parameter. We have that $|V| = 9a + 45a = 54a$, and the size of each organization is $18a$.

There are three organizations $1, 2, 3$. Each triangle $K_3^j$ contains one vertex from each organization. Furthermore, $K_5^{(j-1)\cdot 3a+1}, \ldots, K_5^{(j-1)\cdot 3a+3a}$ contains 2 vertices from organizations $j$ and $j+1$ and 1 vertex from $j+2$ for $j \in [3]$, where addition is taken modulo 3 (i.e. $2+1=3$, but $2+2=1$).

It is easy to see that for any pairwise exchange (matching), at least $12a$ vertices must remain unmatched, at least one for each disjoint clique in the graph.

Furthermore, if two organizations join, then we claim that together they can create a matching where at most $6a$ of their vertices remain unmatched. Indeed, they can match all their vertices in the triangles as there is always an edge between them. Among $K_5^1, \ldots, K_5^{9a}$, there are $6a$ cliques, where one of them has an odd number of vertices. In such cliques, they must leave one of their vertices unmatched, but this can be an arbitrary one of their vertices, as any two vertices within a $K_5^j$ are connected. In the cliques $K_5^j$, where both of them have an even number of vertices, all can be matched.

Finally, a single organization can also create a matching, leaving only $6a$ of its vertices unmatched. This holds because among the $9a$ cliques $K_5^1, \ldots, K_5^{9a}$, it has an even number of vertices in $6a$ of them, so all of them can be matched there. Hence, only $3a$ vertices in the $K_5^j$s and $3a$ in the triangles will be unmatched.

Let $d := a$. Suppose for the contrary that there exists a $3(d-1)$-supplemented core solution $\mathcal{E}$. By our above observations, the two worst-off organizations together have at least $\frac{2}{3} \cdot (12a - (3d-3)) =$

$8a - 2d + 2 = 6a + 2$ unmatched vertices. Say one of them has $c \geq 0$ and the other at least $8a-2d+2-c$. As each organization can create a matching with at most $6a$ unmatched vertices, we must have that $c \leq 6a$ and $8a-2d+2-c \leq 6a$, so $2a-2d+2 \leq c$. Hence, $2 \leq c \leq 6a = 8a-2d+2-2$ and so both organizations must have at least 2 vertices unmatched. Hence, both of these two organizations can improve their situation in a matching among themselves, where only $6a < 6a+2$ vertices are unmatched, contradicting the fact that $M$ is in the $3(d-1)$-supplemented core.

As we have that $d = a$ and hence $d = \frac{|V|}{54}$, the statement follows.

For $\Delta \geq 3$, the proof is deferred to the Appendix. □ □

## 4. Upper Bounds

4.1. **Pairwise Exchange.** We first consider pairwise exchange. Our main result on the existence of the supplemented core relies on the following concept of independent odd cycles.

Definition 4.1. *Given an undirected graph $\mathcal{G} = (V, E)$, a set of odd-length cycles $\mathcal{C} = \{c_1, \ldots, c_l\}$ is called* independent*, if $c_i \cap c_j = \emptyset$ for $i \neq j$ and there exists no $u \in c_i, v \in c_j$, $j \neq i$ such that $(u, v) \in E$. Let $\nu(\mathcal{G})$ denote the maximum number of independent odd cycles of $\mathcal{G}$.*

Theorem 2. *In a partition exchange economy with pairwise exchanges, the $\nu(\mathcal{G})$-supplemented core is always nonempty and contains a Pareto-optimal solution. As a special case, if $\mathcal{G}$ is bipartite, then the core is nonempty and contains a Pareto-optimal solution.*

*Proof.* Proof. By Lemma 2.2, there exists a fractional exchange $y^*$, such that if $u_i(\mathcal{E}^*) \geq \lfloor \sum_{c \in \mathcal{C}_2} \gamma_c^i y_c^* \rfloor = \lfloor \sum_{e \in E} \gamma_e^i y_e^* \rfloor$ for all $i \in N$ for some exchange $\mathcal{E}^*$ using $d$ additional donors, then $\mathcal{E}^*$ is in the $d$-supplemented core. Let $y^*(E(U^i)) := \sum_{e \in E} \gamma_e^i y_e^*$.

We create a polyhedron $\mathcal{P} = \{Az \leq a, Bz \geq b, z \geq 0\}$ as follows. The variables $z$ correspond to the edges of $\mathcal{G}$. Furthermore, the matrix $A$ is just the incidence matrix of $\mathcal{G}$ with $a \equiv 1$. In the matrix $B$, we have a row for each organization, such that $B_{i,e} = \gamma_e^i = |e \cap U^i|$. It is easy to see that the row $B_i$ of an organization $i$ is the sum of the rows $A_v$ for $v \in U^i$. Every entry of $B$ is from $\{0, 1, 2\}$. The vector $b$ is set such that $b_i = \lfloor y^*(E(U^i)) \rfloor$.

The polyhedron $\mathcal{P}$ is clearly nonempty, as $y^* \in \mathcal{P}$.

Consider the following procedure. Let $z^*$ be an extreme point of $\mathcal{P}$ that maximizes $\sum_{e \in E} (\sum_{i \in N} \gamma_e^i) z_e^*$. If there is an integer coordinate $z_e^*$ in $z^*$, then we delete that column from $A$ and $B$ and set $a := a - z_e^* \cdot A_{\cdot,e}$ and $b := b - z_e^* \cdot B_{\cdot,e}$, where $Q_{\cdot,j}$ denotes columns $j$ for a matrix $Q$. Also, if a row becomes all zero, then we delete it. Furthermore, if $a_v$ becomes 0, then we can set all variables with a nonzero entry in the row $A_v$ to 0 and delete them. Hence, we can assume that $a$ is always 1 for every coordinate. If some row or column got deleted, then we update $z^*$ to an extreme point of the new polyhedron $\mathcal{P}$ that maximizes $\sum_{e \in E} (\sum_{i \in N} \gamma_e^i) z_e^*$.

If at some point, all variables become integral, then it gives an exchange $\mathcal{E}$ in the core by Lemma 2.2.

Otherwise, suppose that all variables of the extreme point $z^*$ are fractional. Then, we can use Lemma 2.3 for the (updated) polyhedron $\mathcal{P}$ to show some properties that are satisfied by $z^*$.

Credit each variable $z_e^*$ of $z^*$ with 1 token. Then, we redistribute the tokens as follows. If $u \in V^i \setminus U^i$ for some $i \in N$, or $A_u z^* = 1$, then we give $\frac{1}{2}$ token to $A_u$. Else, $u \in U^i$, for some $i$,

$A_u z^* < 1$ and we give $\frac{1}{2}$ token to $B_i$. Similarly, if $v \in V^j \setminus U^j$ for some $j \in N$, or $A_v z^* = 1$, then we give $\frac{1}{2}$ token to $A_v$. Otherwise, $v \in U^j$ for some $j$, $A_v z^* < 1$ and we give $\frac{1}{2}$ token to $B_j$. It is easy to see that each $z^*_e$ component distributes at most 1 token.

Furthermore, each tight row $A_v$ gets at least one token, since each coordinate is $\{0, 1\}$ in $A_v$ and all coordinates are fractional, so there are at least two edges $e_1, e_2$ that give $\frac{1}{2}$ token to $A_v$.

Take a tight row $B_i$, such that not all $A_v$ rows with $v \in U^i$ are tight. Otherwise, the row $B_i$ would be generated by the tight rows $A_v$ for $v \in U^i$.

Then, take a vertex $v \in U^i$ with $A_v z^* < 1$. By the tightness of $B_i$, we have that there are at least two $v \in U^i$ with $A_v z^* < 1$, since $b_i = B_i z^* = \sum_{v \in U^i} A_v z^*$ and $b_i$ is integer. Therefore, $B_i$ gets at least one token too.

By Lemma 2.3 we get that this is only possible, if each tight row that got at least one token got exactly one token, and if a row got less than one token, then it got 0.

From this, we conclude that there are at most two fractional edges of $z^*$ incident to any vertex $v$ with $A_v z^* = 1$. We claim that this is also true for a vertex $v$ with $A_v z^* < 1$. Indeed, if $v \in U^i$ for an organization $i$ such that $B_i$ is tight, then all fractional edges incident to $v$ give a $\frac{1}{2}$ token to $B_i$'s row, so there are at most two. Finally, if $v \notin U^i$ for any $i \in N$, then $\frac{1}{2}$ of $z^*_e$'s token would have been given to $A_v$ for any fractional edge $e$ with $v \in e$, hence as only tight rows received tokens, there can be no fractional edges incident to $v$. Hence, the fractional edges of $z^*$ correspond to a disjoint set of cycles and paths in the graph $\mathcal{G}[\cup_{i \in N} U^i]$. Also, no paths can be odd-length, as otherwise it contradicts the fact that $z^*$ maximizes $\sum_{e \in E} (\sum_{i \in N} \gamma^i_e) z^*_e$. Furthermore, there can be no even length cycle $C$ either, since otherwise $z^*$ is not extreme: $z^* = \frac{1}{2}(z_1 + z_2)$, where $z_1 = z^* - \varepsilon\chi^{C_1} + \varepsilon\chi^{C_2}$ and $z_2 = z^* + \varepsilon\chi^{C_1} - \varepsilon\chi^{C_2}$ for some sufficiently small enough $\varepsilon > 0$ and $C = C_1 \cup C_2$ is chosen such that the edges of $C$ alternate between $C_1$ and $C_2$. For even length paths, the endpoints $v$ necessarily have non-tight $A_v$ rows, and thus by the tight token distribution argument the endpoints belong to organizations with tight $B_i$ rows who control exactly two endpoints of (by the above, necessarily even) paths. Hence, we can take the union of all such paths and create a set of even cycles $C$ (via contracting the two endpoints controlled by each organization), and obtain that $z^*$ is not extreme: again, $z^* = \frac{1}{2}(z_1 + z_2)$, where $z_1 = z^* - \varepsilon\chi^{C_1} + \varepsilon\chi^{C_2}$ and $z_2 = z^* + \varepsilon\chi^{C_1} - \varepsilon\chi^{C_2}$ for some sufficiently small $\varepsilon > 0$ ( $C = C_1 \cup C_2$ is chosen such that the edges of $C$ alternate between $C_1$ and $C_2$).

Hence, we conclude that the fractional edges of $z^*$ gives a vertex disjoint set of odd cycles $\mathcal{C}$. If $\mathcal{G}$ is bipartite, then this is a contradiction, which shows that the core is nonempty in this case, and $z^*$ corresponds to a Pareto-optimal exchange in the core.

Suppose that $\mathcal{G}$ is not bipartite.

We can suppose that $\mathcal{C}$ is an independent set of odd cycles. Otherwise, if there exists an edge $(u, v)$ with $u \in c_{j_1}, v\rho_{j_2}$, $j_1 \neq j_2$, then all vertices in $c_{j_1} \cup c_{j_2}$ can be covered with a matching. Hence, we may choose such a matching, and iterate this procedure until no such case exists.

From this, it follows that we can round $z^*$ to a matching $\mathcal{E}$ that satisfies that at most $\nu(\mathcal{G})$ vertices are unmatched from these cycles. Hence, with $\nu(\mathcal{G})$ additional donors, we can guarantee that the final supplemented exchange $\mathcal{E}^*$ satisfies $u_i(\mathcal{E}^*) \geq \lfloor y^*(E(U^i)) \rfloor$, so by Lemma 2.2 it gives a solution in the $\nu(\mathcal{G})$-supplemented core. If $\mathcal{E}^*$ is not Pareto-optimal, then we can take a Pareto-improving exchange that maximizes total utility among such improvements. It still satisfies the

required utility lower bounds, so is in the $\nu(\mathcal{G})$-supplemented core, and it additionally satisfies Pareto-optimality too. □ □

The proof of the following remark can be found in the Appendix.

REMARK 4.1. *The $\nu(\mathcal{G})$-supplemented core exchange in Theorem 2 can be chosen in such a way that each organization $i$ requires at most $\lceil \frac{|U^i|}{3} \rceil$ additional donors.*

Next, we show that in a random compatibility graph, the expected maximum number of independent odd cycles grows logarithmically with the size of the graph. This shows that in a typical random compatibility graph, compared with the market size, only a small number of altruistic donors are needed to stabilize the exchange.

LEMMA 4.1. *Consider a compatibility graph that is generated according to the random graph model. Assuming a constant number $g$ of groups, for any distribution of the vertices to the groups, in a random kidney exchange graph $\mathcal{G}$, the expected value satisfies $\mathbb{E}[\nu(\mathcal{G})] = \mathcal{O}(\log |V|)$ and as the size of graph increases, the probability that $\nu(\mathcal{G}) = \mathcal{O}(\log |V|)$ approaches 1.*

From Lemma 4.1 and Theorem 2 we immediately get the following logarithmic bound on the altruistic donors needed by the system.

THEOREM 3. *Consider a partition exchange economy with pairwise exchanges, where the compatibility graph on the original vertices $V$, including the existing altruistic donors in $V \setminus U$, is generated according to the random graph model with a constant number of groups. Then, there is a constant $C = C(g, p)$ such that, if $d(\mathcal{G})$ denotes the minimum number of* additional *altruistic donors required for a supplemented-core exchange, then*

$$\mathbb{E}[d(\mathcal{G})] \leq C \log |V| \quad \textit{and} \quad \Pr\left(d(\mathcal{G}) \leq C \log |V|\right) \xrightarrow{|V| \to \infty} 1.$$

*The expectation and probability are taken over the realization of the random compatibility arcs. The supplemented-core exchange can be chosen Pareto-optimal.*

4.2. **Cyclic Exchange.** Next, we consider cyclic exchanges. We first study the practically most relevant case in which cycles have length at most three. In Appendix C.2, we show that the same logarithmic bound extends to every fixed upper bound on cycle length.

The pairwise argument from Section 4.1 does not extend directly. With pairwise exchanges, the support of a fractional extreme point can be reduced to independent odd cycles. With three-way exchanges, fractional extreme points may have substantially more complicated supports. Our argument therefore rounds the aggregate numbers of cycles of different patterns rather than the support of the fractional exchange itself.

THEOREM 4. *Consider a partition exchange economy with $\Delta = 3$, where the compatibility graph on the original vertices $V$, including the existing altruistic donors in $V \setminus U$, is generated according to the random compatibility model. Let $m := |V|$, and let $d(\mathcal{G})$ denote the minimum number of additional altruistic donors required for a supplemented-core exchange. There is a constant $C = C(g, p)$ such that*

$$\mathbb{E}[d(\mathcal{G})] \leq C \log(m + 2) \quad \textit{and}$$

$$\Pr\left(d(\mathcal{G}) \le C\log(m+2)\right) \ge 1-(m+2)^{-4}.$$

*The expectation and probability are taken over the realization of the random compatibility arcs. Moreover, the supplemented-core exchange can be chosen Pareto-optimal.*

The proof rests on the following rounding result. Its uniformity is important because the fractional exchange obtained from Lemma 2.2 depends on the realized compatibility graph.

LEMMA 4.2 (Logarithmic aggregate-pattern rounding). *Let $m := |V|$. There is a constant $C_0 = C_0(g,p)$ such that, with probability at least $1-(m+2)^{-4}$, the following statement holds simultaneously for every fractional exchange $y$ using cycles of size at most three: there is an integral exchange $\mathcal{E}$, using only the original vertices in $V$, such that*

$$\sum_{i\in N}\left(\left\lfloor\sum_{c\in\mathcal{C}_3}\gamma_c^i y_c\right\rfloor - u_i(\mathcal{E})\right)^+ \le C_0\log(m+2).$$

The formal proof is given in Appendix C.1. We describe its structure here. A cycle pattern records, for each position in a two- or three-cycle, the random-graph group of the corresponding vertex and whether that vertex represents a utility-bearing patient or an existing altruistic donor. Because the number of groups and the cycle-length bound are fixed, there are only constantly many patterns.

For each pattern, we retain only the total fractional weight assigned to cycles of that pattern and round this total down to an integer. Flooring the pattern weights removes only a constant number of utility-bearing cycle positions. A transportation argument then assigns the remaining integer numbers of positions to organizations. The integrality of the transportation problem ensures that this assignment creates only a constant aggregate loss relative to the organizations' fractional utilities. Existing altruistic donors enter the relevant capacity constraints, but an uncovered altruistic slot is not charged as a utility loss.

It remains to realize the assigned positions as vertex-disjoint cycles in the compatibility graph. For

$$h = K\log(m+2),$$

the random graph satisfies, with high probability, a uniform connectivity property: every sufficiently large pair of admissible vertex sets contains the arcs needed for a two-cycle, and every sufficiently large family of first-stage paths can be completed using a sufficiently large set of third vertices. Hall's theorem then realizes all but $O(h)$ copies of each pattern. Since there are only constantly many patterns, the total unrealized utility is $O(\log m)$.

*Proof.* Proof of Theorem 4. Let $m := |V|$. By Lemma 2.2, there is a fractional exchange $y^*$ such that it is sufficient to construct an augmented integral exchange giving every organization $i$ at least

$$b_i := \left\lfloor\sum_{c\in\mathcal{C}_3}\gamma_c^i y_c^*\right\rfloor$$

transplants.

On the high-probability event in Lemma 4.2, the lemma may be applied to the graph-dependent exchange $y^*$ because its conclusion holds simultaneously for every fractional exchange. It gives an

integral exchange $\mathcal{E}$, using only the original vertices, for which

$$\sum_{i\in N}\big(b_i - u_i(\mathcal{E})\big)^+ \le C_0 \log(m+2).$$

For each organization $i$, define its remaining deficit by

$$e_i := \big(b_i - u_i(\mathcal{E})\big)^+.$$

Because $b_i \le |U^i|$, organization $i$ has at least $e_i$ uncovered utility-bearing patients. Add $e_i$ altruistic donors and assign them to $e_i$ distinct uncovered patients of organization $i$. Under the dummy-patient representation, each new altruistic donor and the vertex associated with the uncovered patient form a directed two-cycle. The resulting augmented exchange $\mathcal{E}^+$ therefore satisfies

$$u_i(\mathcal{E}^+) \ge u_i(\mathcal{E}) + e_i \ge b_i \qquad (i \in N).$$

Lemma 2.2 then implies that $\mathcal{E}^+$ belongs to the supplemented core. Consequently, on the high-probability event,

$$d(\mathcal{G}) \le \sum_{i\in N} e_i \le C_0 \log(m+2).$$

After increasing the constant if necessary, this proves the stated high-probability bound.

For every realization of the compatibility graph, the trivial bound

$$d(\mathcal{G}) \le |U| \le m$$

holds, since one additional altruistic donor can be assigned to every utility-bearing patient. Hence,

$$\begin{aligned}\mathbb{E}[d(\mathcal{G})] &\le C_0 \log(m+2) + m \Pr\left(d(\mathcal{G}) > C_0 \log(m+2)\right)\\ &\le C_0 \log(m+2) + m(m+2)^{-4}\\ &\le C \log(m+2)\end{aligned}$$

for a constant $C = C(g,p)$.

Finally, if the resulting supplemented-core exchange is not Pareto-optimal, replace it by a Pareto-dominating exchange that maximizes total utility in the same augmented compatibility graph. Any coalition blocking the new exchange would also block the original one. The replacement therefore remains in the supplemented core and is Pareto-optimal. □ □

Extension to longer cycles. The same construction extends to every fixed cycle-length bound $\Delta$. The number of augmented patterns remains constant, with the constant now depending on $\Delta$, and a pattern can be realized by extending vertex-disjoint partial paths one slot at a time. Each matching stage loses at most $O(\log m)$ paths, and there are only constantly many stages and patterns. Thus, for every fixed $\Delta$, the required supplementation remains $\mathcal{O}_{g,p,\Delta}(\log(m+2))$ in expectation and with probability at least $1-(m+2)^{-4}$. The complete argument is given in Appendix C.2.

## 5. Numerical Results

In this section, we use simulations to study the existence of the core and the stability of allocation rules used in practice. For computational feasibility, we test deviations by coalitions of at most four organizations. We therefore use "stable" in this numerical section to mean stable against coalitions

of size at most four. When there are at most five organizations, this check is sufficient for the full core because we restrict attention to maximum-cardinality exchanges, so the grand coalition cannot block. For larger numbers of organizations, the results should be interpreted as stability against small coalitions. This is also an economically relevant benchmark because coordinating a deviation becomes more difficult as the number of participating organizations increases. We first ask whether a maximum-cardinality exchange in the weak core exists. We then examine whether a representative lexicographic rule selects such an exchange. A stable exchange may exist even when the rule selects an unstable one.

The weak core is nonempty in almost every simulated market. In the only case requiring supplementation, one altruistic donor suffices. The lexicographic rule, however, selects an exchange blocked by a coalition of at most four organizations in as many as 20% of instances in some settings. Adding a small number of voluntarily registered altruistic donors and reapplying the same rule almost always restores stability.

This finding suggests a simple practical use of supplementation. The platform applies its existing allocation rule, checks the selected exchange for blocking coalitions, and introduces additional donors only if the exchange is unstable. Supplementation can therefore improve stability without changing the platform's medical, ethical, and logistical priorities.

In addition to the weak core, we examine two more demanding stability notions. We first define the TU core, the strongest of the three notions considered in our numerical analysis.

Definition 5.1. *Given an exchange $\mathcal{E}$ and a coalition $P \subseteq N = \{1, \dots, n\}$, we say that $P$ is a* TU blocking coalition to $\mathcal{E}$ *if there exists an exchange $\mathcal{E}'$ in the graph $\mathcal{G}[\cup_{i \in P} V_i]$ such that*

$$\sum_{i \in P} u_i(\mathcal{E}') > \sum_{i \in P} u_i(\mathcal{E}).$$

*We say that $\mathcal{E}$ is in the* TU core *if it has no TU blocking coalition.*

Thus, an exchange is in the TU core if every coalition receives at least as many transplants in the platform's exchange, in total, as it could obtain by withdrawing and exchanging on its own. The TU core implies the strong core, but the converse need not hold. A strong blocking coalition must make every member weakly better off and at least one member strictly better off. By contrast, a TU blocking coalition needs only to increase the coalition's total number of transplants and may make some of its members worse off.

Definition 5.2. *Given additional altruistic donors $V^0$, an exchange $\mathcal{E}^*$ in the extended graph $\mathcal{G}^{+V^0}$ is in the* $V^0$-supplemented TU core *if it is not TU blocked by any coalition $P \subseteq N$. If $|V^0| \leq d$, we also refer to $\mathcal{E}^*$ as a* $d$-supplemented TU-core exchange.

5.1. **Algorithms Under Study.** We generate instances of partition exchange economies using the state-of-the-art generator of Delorme et al. (2022) and Pettersson and Trimble (2021), which refines the widely used Saidman generator of Saidman et al. (2006). This approach is consistent with previous computational studies of kidney exchange (Benedek et al. 2023, 2024, Druzsin et al. 2024). We generate 10 base instances, each containing 1,000 patient–donor pairs and 50 available altruistic donors.

We study two classes of allocation procedures. The first class consists of *core-aware procedures*, which search among maximum-cardinality exchanges for an outcome satisfying the weak-, strong-, or TU-core constraints. These procedures address whether a stable efficient exchange exists. The second is a *practice-based lexicographic procedure*, which first optimizes a representative hierarchy of objectives used by kidney exchange programs and only afterward tests the selected exchange for weak-core stability. This procedure addresses whether operational objectives select a stable outcome and, when they do not, how many additional altruistic donors are needed to stabilize the outcome without changing the priority order.

All methods operate on a precomputed cycle database containing every feasible 2- and 3-cycle in the directed compatibility graph.

Cycle enumeration. Given a set of patient–donor vertices $V$ and adjacency lists $A^+$, we enumerate all simple directed cycles of length $\Delta \in \{2, 3\}$. Each cycle $c$ stores (i) its ordered vertex tuple, (ii) its length $|c|$, (iii) the per-organization incidence counts $\gamma_c^i$, (iv) the number $a_c$ of altruistic donors on the cycle, if any, (v) the number of same-blood-type edges, and (vi) a hardness score equal to $\max_{v \in c} \rho(v)^{-1}$, where $\rho(v)$ is the indegree of vertex $v$. These statistics provide the objective weights for the integer programs described below.

TU-core optimization with pre-added altruists. For computational speed, the TU-core procedure makes a set of altruists available at the outset. To approximate observed availability, we add altruists equal to 5% of the number of patients, sampled uniformly at random from the generated donor pool. Let $\mathcal{C}_\Delta$ be the enumerated cycles, including those involving altruists, and let $x_c \in \{0, 1\}$ indicate whether cycle $c$ is selected. We first compute the baseline coverage $K$ by solving the maximum-packing problem on the graph without the additional altruists:

$$\max \sum_{c \in \mathcal{C}_\Delta} |c| x_c \qquad \text{s.t.} \qquad \sum_{c \ni v} x_c \leq 1 \quad \forall v \in V, \qquad x_c \in \{0, 1\}.$$

Holding the number of real-patient transplants at least at $K$, we then minimize altruist use:

$$\min \sum_{c \in \mathcal{C}_\Delta} a_c x_c,$$

subject to the vertex-disjointness constraints, the coverage constraint

$$\sum_{c \in \mathcal{C}_\Delta} (|c| - a_c) x_c \geq K,$$

and, for every tested coalition $S$, the TU-core constraint

$$\sum_{c \in \mathcal{C}_\Delta} \left( \sum_{i \in S} \gamma_c^i \right) x_c \geq \mathrm{opt}_S,$$

where $\mathrm{opt}_S$ is the maximum number of transplants that coalition $S$ can obtain in its induced subgraph. The resulting objective is lexicographic: maximize real-patient transplants first and minimize the number of altruists used second. We also record the number of distinct altruists in the selected exchange.

Strong-core heuristic. The strong-core procedure begins without additional altruists and solves the maximum-packing problem. It then restricts attention to exchanges of the same maximum

cardinality. After each solve, we test strong-core stability. For each coalition $S$, the deviation problem requires every $i \in S$ to receive at least its status-quo utility and requires the coalition as a whole to gain at least one transplant. If a blocking coalition is found, we add the core cut

$$\sum_c \left( \sum_{i \in S} \gamma_c^i \right) x_c \geq U_S^{\text{baseline}} + 1$$

and re-solve. We add the next randomly ordered altruist and reconstruct the cycle database only when the cut-augmented problem becomes infeasible. The procedure terminates when the selected exchange passes the strong-core test or the available donor pool is exhausted.

Weak-core heuristic. The weak-core procedure is identical to the strong-core procedure except that a blocking coalition must make every member strictly better off. It uses the same iterative cut generation and adds a randomly ordered altruist only when the cut-augmented problem becomes infeasible. A strong-core solution obtained without additional altruists also certifies weak-core stability, which we exploit in the computational pipeline.

Lexicographic weak-core optimization. This procedure is our operational benchmark. It separates the objectives used to *select* an exchange from the criterion used to *verify* stability. We first solve a four-stage lexicographic integer program subject only to feasibility and vertex-disjointness:

(1) maximize the number of real-patient transplants;
(2) subject to (1), maximize the number of cycles, thereby favoring shorter cycles;
(3) subject to (1)–(2), maximize the number of edges on which donor and recipient blood types match; and
(4) subject to (1)–(3), maximize the number of hard-to-match patients receiving a transplant, using the hardness score defined above.

This hierarchy reflects commonly used objectives in kidney exchange programs (Biró et al. 2021), although the precise objectives and their ordering vary across programs. Each stage fixes the objective values attained at the preceding stages.

The lexicographic program does not impose core constraints. After solving it, we test whether its selected exchange is weak-core stable. If the exchange is blocked, we add the next altruistic donor from the randomly ordered generated pool, rebuild the cycle database, and solve the same four-stage lexicographic program again. We repeat this process until the output is stable or no donor remains. Thus, the intervention changes neither the objectives nor their priority order. It augments the feasible exchange environment with voluntarily registered donors until the established allocation rule selects a stable outcome. Because donors are generated from realistic distributions and introduced in random order, the procedure does not assume that the platform can choose their compatibility characteristics.

5.2. **Simulation Framework.** We vary the following parameters:

- **Base instances:** 10 XML instances generated under the default configuration of Delorme et al. (2022) and Pettersson and Trimble (2021).
- **Number of organizations:** $n \in \{5, 10, 15, 20, 30\}$. Patient–donor vertices are assigned to organizations using a uniform assignment with Dirichlet noise to allow heterogeneous organization sizes.

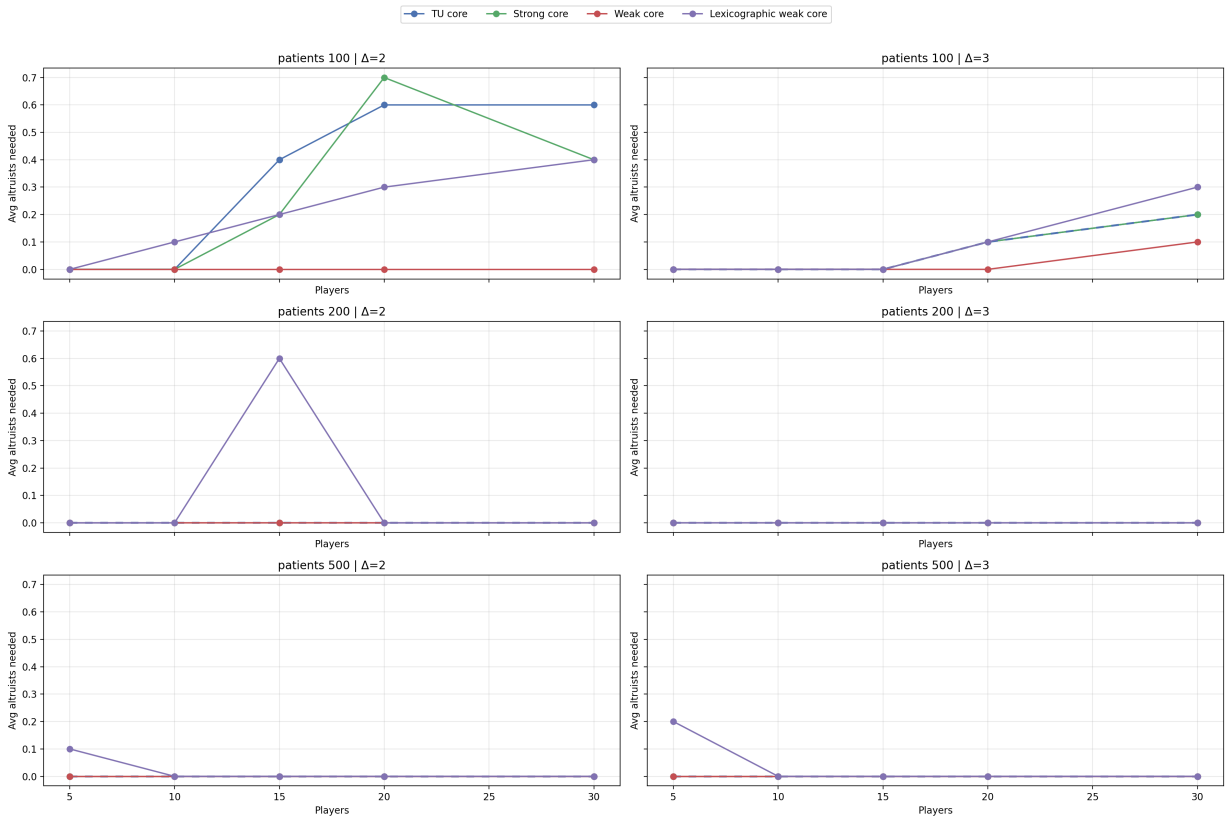



FIGURE 3. Average number of additional altruistic donors required under each allocation procedure and simulation setting.

- **Pool size:** 100, 200, or 500 patient–donor vertices, sampled without replacement from each base instance.
- **Cycle-length cap:** $\Delta \in \{2, 3\}$.
- **Coalition-size cap:** blocking coalitions contain at most four organizations.

This gives $5 \times 2 \times 3 \times 10 = 300$ simulated markets for each procedure. All simulations were run on an AMD Ryzen 7735HS processor with 16GB RAM.

5.3. **Results.** Figures 3–5 report, respectively, the average number of altruists required, the maximum number required, and the fraction of simulated markets requiring supplementation. The results reveal a sharp distinction between the existence of stable exchanges and the stability of exchanges selected by the practice-based lexicographic rule.

Stable solutions usually exist. Our first benchmark asks whether stability is feasible when the platform may search among maximum-cardinality exchanges for a core outcome. The answer is strongly positive. Across the 300 simulated markets, the core-aware procedure finds a weak-core exchange without additional altruists in 299 cases. In the remaining case, one donor suffices. Importantly, this donor is drawn from the randomly ordered pool of generated altruists; it is not chosen to possess a particular compatibility. Intrinsic nonexistence of the weak core therefore appears to be rare in the compatibility environments represented by our simulations.

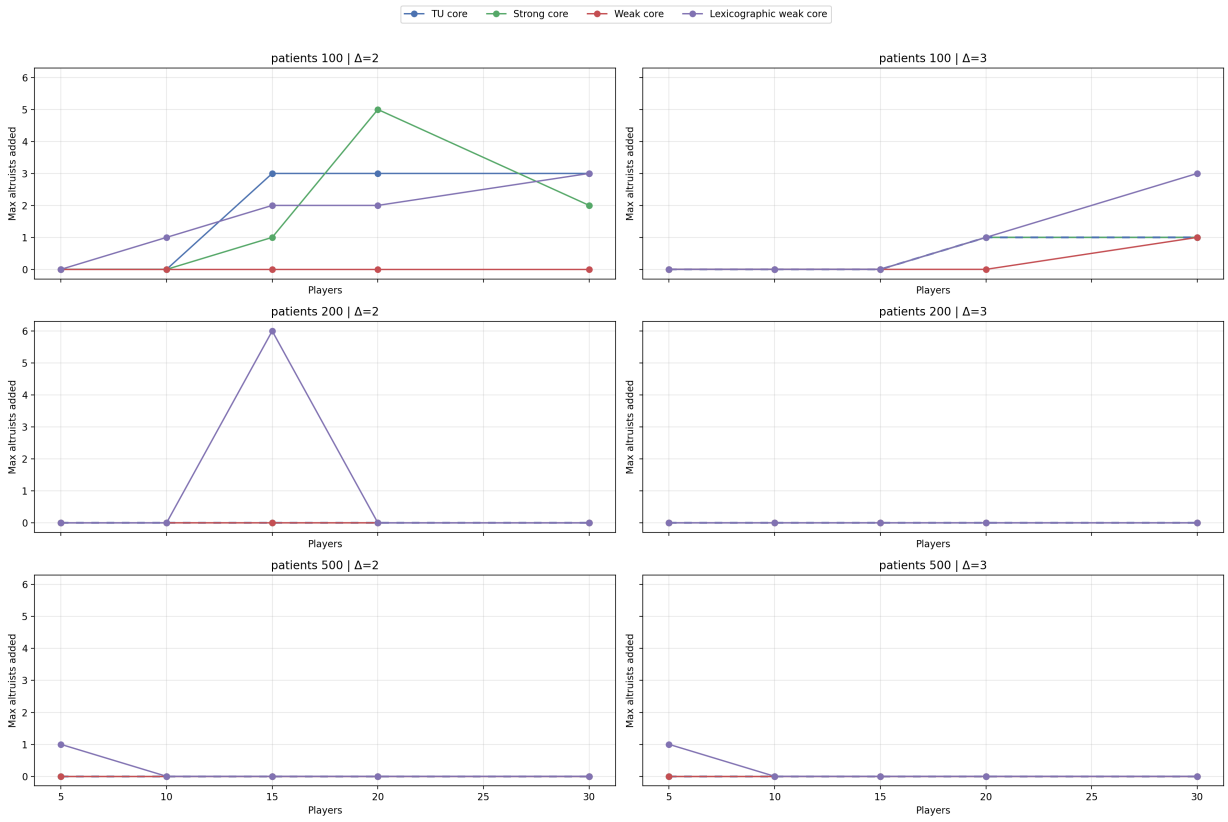



FIGURE 4. Maximum number of additional altruistic donors required under each allocation procedure and simulation setting.

This finding is reassuring, but it does not make the supplemented-core concept irrelevant. Nonemptiness means that at least one stable exchange exists; it does not imply that an allocation rule pursuing other objectives will select that exchange.

Operational objectives need not select a stable exchange. The practice-based lexicographic results make this distinction visible. For pools of 100 patient–donor pairs, the lexicographic rule selects an exchange that requires supplementation for weak-core stability in approximately 10% of the simulations. With pairwise exchange and more than 15 organizations, this fraction reaches 20%. We also observe unstable lexicographic outcomes in a small number of simulations with 200 and 500 patient–donor pairs.

Because a donor-free weak-core exchange exists in 299 of the 300 simulated markets, almost all of these failures are failures of selection rather than failures of existence. In other words, an exchange that is stable against the tested coalitions is available, but the operational hierarchy selects a different maximum-cardinality exchange that those organizations can block. This is an algorithmic stability gap: solving the medical and operational optimization problem does not by itself solve the participation problem.

A small amount of supplementation closes the algorithmic stability gap. We next retain the same four lexicographic priorities and introduce randomly ordered altruistic donors one at a time. The required intervention is small. The average number of additional donors is at most 0.6 in every

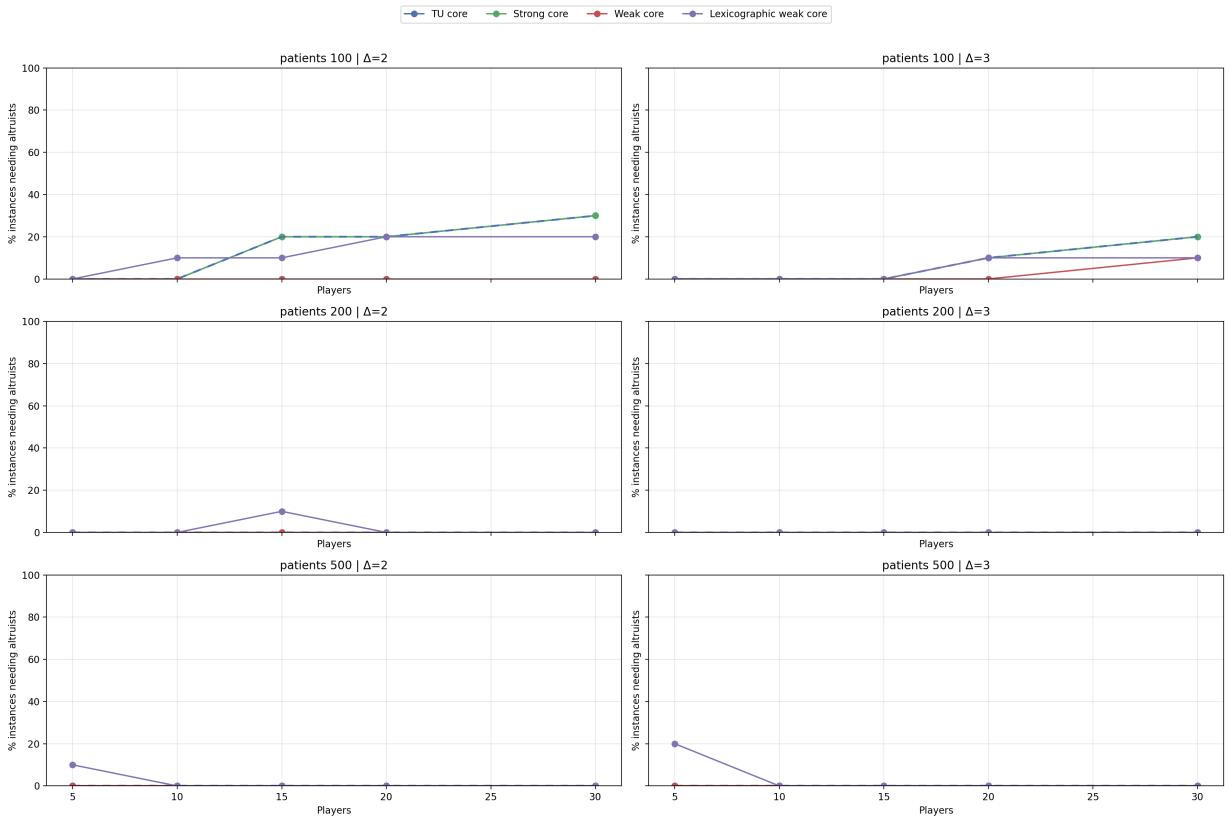



FIGURE 5. Percentage of simulated markets in which additional altruistic donors are required under each allocation procedure. The TU-core data points are not visible where they coincide with the strong-core data points.

parameter setting. In all but one simulated market, at most three donors are sufficient for the lexicographic rule to select a weak-core-stable exchange. The sole exception has 200 patient–donor pairs, 15 organizations, and $\Delta = 2$, and requires six donors.

These results are important operationally. The platform need not discard or reorder objectives that may reflect medical priorities, logistical concerns, or existing governance arrangements. It can continue to use the same allocation rule and coordinate a small number of voluntarily registered altruistic donors when the rule's initial output fails the stability test. Supplementation therefore functions as a modular stability intervention, rather than as a wholesale redesign of the allocation algorithm.

Stronger stability notions. The strong- and TU-core results reinforce the same conclusion. For pools of 200 and 500 patient–donor pairs, the core-aware procedures find strong- and TU-core exchanges without additional altruists in every parameter setting. With 100 pairs and $\Delta = 2$, the average number of donors required remains below 0.7; the maximum is three for the TU core and five for the strong-core heuristic. The larger maximum for the strong core in one instance may be an artifact of the heuristic, whose accumulated cuts can exclude feasible strong-core solutions. With $\Delta = 3$, the average is below 0.2 and the maximum is one donor in every setting.

Across the stability notions, larger pools and the availability of 3-cycles generally reduce the need for supplementation. Holding pool size fixed, the frequency with which supplementation is required tends to increase with the number of organizations, reflecting the larger collection of participation constraints that the selected exchange must satisfy.

Overall, the simulations deliver two complementary messages. Stable maximum-cardinality exchanges are ordinarily available, and the rare failure of weak-core existence is repaired by a single randomly generated donor. At the same time, stability is not automatic: a practice-based lexicographic rule can select an exchange that is blocked even when a donor-free core exchange exists. In those cases, a small number of additional altruistic donors allows the platform to preserve its established objective hierarchy while making the implemented exchange stable. The practical value of the supplemented core thus lies not only in restoring existence, but also in bridging the gap between stable allocations and the algorithms used to select them.

## 6. Applications to One-Sided Exchange Economies

We next apply the bipartite case of Theorem 2 to an indivisible-goods exchange economy. The result extends the core-existence theorem of Echenique et al. (2024). The proof uses an auxiliary partition exchange economy in which goods are represented by altruistic-donor vertices and the positions in an assignment valuation are represented by patient–donor-pair vertices. We make the endowment and blocking correspondence explicit below.

There is a finite set of goods

$$M = M^1 \dot\cup \cdots \dot\cup M^n,$$

where $M^i$ is the endowment of agent $i \in N$. Agent $i$ has a valuation $v_i : 2^M \to \mathbb{R}$. An allocation is a partition $M = M_1^* \dot\cup \cdots \dot\cup M_n^*$.

Definition 6.1. *An allocation $M^* = (M_1^*, \ldots, M_n^*)$ is in the weak core if there do not exist a coalition $S \subseteq N$ and a partition*

$$\bigcup_{i\in S} M^i = \dot\bigcup_{i\in S} M_i'$$

*such that $v_i(M_i') > v_i(M_i^*)$ for every $i \in S$.*

Echenique et al. (2024) consider categorical economies. In such an economy, the goods are partitioned into categories, $M = O^1 \dot\cup \cdots \dot\cup O^K$, and each agent $i$ has a set $G_i \subseteq M$ of acceptable goods. Its valuation is

$$v_i(X) = \sum_{k=1}^{K} \min\bigl\{|X \cap O^k \cap G_i|, 1\bigr\}.$$

Thus, an agent obtains one unit of utility from an acceptable good in a category and obtains no additional utility from receiving a second acceptable good in the same category. These preferences are additively separable and dichotomous.

Theorem 5 (Echenique et al. (2024)). *A categorical economy with additively separable and dichotomous preferences has a nonempty weak core.*

We establish core existence for the larger class of binary assignment valuations.

DEFINITION 6.2 (Binary assignment valuation). *A valuation $v : 2^M \to \mathbb{R}$ is a* binary assignment valuation *if there are a finite set of positions $J$ and a matrix $\alpha \in \{0,1\}^{M\times J}$ such that, for every $X \subseteq M$,*

$$v(X) = \max_{z} \left\{ \sum_{g\in X}\sum_{j\in J} \alpha_{gj} z_{gj} : \begin{array}{ll} \sum_{j\in J} z_{gj} \le 1 & \text{for every } g \in X, \\ \sum_{g\in X} z_{gj} \le 1 & \text{for every } j \in J, \\ z_{gj} \in \{0,1\} & \text{for every } g \in X,\ j \in J \end{array} \right\}.$$

*Equivalently, $v(X)$ is the size of a maximum matching between the goods in $X$ and the positions in $J$, where $g$ can be assigned to $j$ if and only if $\alpha_{gj} = 1$.*

The restriction $\alpha_{gj} \in \{0,1\}$ distinguishes this class from the more general assignment valuations studied by Milgrom (2009). Every categorical valuation above is a binary assignment valuation. For agent $i$, take one position $j_k$ for each category $O^k$ and set

$$\alpha^i_{gj_k} = 1 \quad \Longleftrightarrow \quad g \in O^k \cap G_i.$$

Matching goods to these positions gives exactly $\sum_k \min\{|X \cap O^k \cap G_i|, 1\}$.

The inclusion is strict. Consider three goods $a, b, c$ and the valuation $v(X) = \min\{|X|, 2\}$. This is a binary assignment valuation: take two positions and let every good be acceptable at both positions. It cannot be represented by categorical preferences. Because each singleton has value one, all three goods would have to be acceptable. Because every pair has value two, each pair would then have to lie in different categories. The three goods would therefore lie in three different categories, implying $v(\{a, b, c\}) = 3$, a contradiction.

THEOREM 6. *If every agent has a binary assignment valuation, then the weak core is nonempty.*

*Proof.* Proof. For each agent $i$, fix a representation of $v_i$ by a set of positions $J^i$ and a binary matrix $\alpha^i \in \{0,1\}^{M\times J^i}$. We may take the sets $J^i$ to be pairwise disjoint. We construct an auxiliary partition exchange economy with pairwise exchange.

For each good $g \in M^i$, create an altruistic-donor vertex, also denoted by $g$, and assign it to organization $i$. For each position $j \in J^i$, create a patient–donor-pair vertex, also denoted by $j$, and assign it to the same organization. Thus, organization $i$ owns

$$V^i = M^i \dot\cup J^i.$$

This is the endowment correspondence: a coalition $S$ controls precisely the goods in $\bigcup_{i\in S} M^i$ and the positions in $\bigcup_{i\in S} J^i$. The altruistic-donor vertices here represent the agents' endowed goods; they are not additional donors supplied by the platform.

Let $J = \bigcup_{i\in N} J^i$. The graph is bipartite, with vertex classes $M$ and $J$. For $g \in M$ and $j \in J^i$, include the edge $\{g, j\}$ if and only if $\alpha^i_{gj} = 1$; include no other edges. In the directed kidney-exchange representation, the donor at $g$ is compatible with the patient at $j$ exactly when $\alpha^i_{gj} = 1$, while the donor paired with $j$ is compatible with the dummy patient attached to $g$. Hence, each edge $\{g, j\}$ is a feasible pairwise exchange. Matching $g$ to $j \in J^i$ has the economic interpretation of assigning good $g$ to agent $i$ through position $j$.

The graph is bipartite. Theorem 2 therefore gives a matching $\mathcal{E}^*$ in the weak core of the auxiliary partition exchange economy. We now convert this matching into an allocation of the goods. For each agent $i$, let

$$\widehat{M}_i = \big\{g \in M : \{g, j\} \in \mathcal{E}^* \text{ for some } j \in J^i\big\}.$$

Because $\mathcal{E}^*$ is a matching, the sets $\widehat{M}_i$ are pairwise disjoint. Let $R = M \setminus \bigcup_{i \in N} \widehat{M}_i$ be the set of unmatched goods and return each such good to its original owner. Define

$$M_i^* = \widehat{M}_i \cup (R \cap M^i), \qquad i \in N.$$

The sets $M_i^*$ form a partition of $M$. Moreover, the edges of $\mathcal{E}^*$ incident to $J^i$ give a feasible assignment of the goods in $\widehat{M}_i$ to distinct positions in $J^i$. Consequently,

$$(*) \qquad v_i(M_i^*) \;\geq\; |\widehat{M}_i| \;=\; u_i(\mathcal{E}^*).$$

Equality is not needed: returning an unmatched good may raise an agent's valuation, which only makes that agent harder to improve upon.

It remains to show that the construction preserves coalitional deviations. Suppose, toward a contradiction, that a coalition $S \subseteq N$ strongly blocks $M^*$. Then $S$ can repartition its original endowment as

$$\bigcup_{i \in S} M^i = \dot{\bigcup_{i \in S}} M_i'$$

so that $v_i(M_i') > v_i(M_i^*)$ for every $i \in S$. For each $i \in S$, choose an assignment attaining $v_i(M_i')$. Retain the pairs $(g, j)$ with $\alpha_{gj}^i = 1$. These pairs form a matching between $M_i'$ and $J^i$ of size $v_i(M_i')$. Because the bundles $M_i'$ are disjoint and the position sets $J^i$ are disjoint, the union of these matchings over $i \in S$ is a feasible matching $\mathcal{E}'$ in the subgraph induced by

$$\bigcup_{i \in S} V^i = \left(\bigcup_{i \in S} M^i\right) \dot{\cup} \left(\bigcup_{i \in S} J^i\right).$$

For every $i \in S$, this matching gives $u_i(\mathcal{E}') = v_i(M_i') > v_i(M_i^*) \geq u_i(\mathcal{E}^*)$, where the last inequality follows from $(*)$. Thus $S$ blocks $\mathcal{E}^*$ in the auxiliary partition exchange economy, contradicting the choice of $\mathcal{E}^*$. Hence $M^*$ is in the weak core. □

## 7. Conclusions

Core stability has generally been viewed as a demanding and often unattainable benchmark for kidney exchange among organizations. The core may be empty, and even when it is nonempty, finding and implementing a stable exchange can be difficult. Our results offer a more optimistic perspective. Exact core stability can often be achieved through a modest intervention: supplementing the exchange with altruistic donors. Rather than weakening the stability requirement or relying on monetary transfers, the platform expands the feasible set just enough to satisfy organizations' participation constraints. The supplemented core therefore turns stability from a passive property of the market into an objective that the platform can actively pursue.

The theoretical and computational results provide complementary support for this approach. The worst-case bounds show that stability can be costly in arbitrary compatibility graphs. In random and type-based environments, however, the required supplementation remains small as

the market grows. The simulations go further: core-stable outcomes are usually available, and outcomes selected by practical lexicographic rules can be stabilized with only a few realistically generated donors. Thus, the supplemented core is not merely an existence result. It provides a way to reconcile stability with the clinical and operational priorities already used by kidney exchange programs.

A practical next step is to translate the supplemented core into an operating protocol. A platform could maintain a reserve of voluntarily committed altruistic donors, test proposed exchanges for blocking coalitions, and deploy donors only when needed to restore stability. In a dynamic setting with repeated clearing, the reserve may be partly self-replenishing: an altruist-initiated chain can leave a bridge donor to initiate a later chain, so using an altruist today need not reduce future stabilizing capacity one-for-one. Further computational work should incorporate stability into existing allocation software and extend the guarantees to larger coalitions and richer operational constraints.

## References


Agarwal, N., I. Ashlagi, E. Azevedo, C. R. Featherstone, and Ö. Karaduman (2019). Market failure in kidney exchange. *American Economic Review 109*(11), 4026–4070.

Akbarpour, M., S. Li, and S. O. Gharan (2020). Thickness and information in dynamic matching markets. *Journal of Political Economy 128*(3), 783–815.

Anderson, R., I. Ashlagi, D. Gamarnik, and Y. Kanoria (2017). Efficient dynamic barter exchange. *Operations Research 65*(6), 1446–1459.

Ashlagi, I., Á. Cseh, D. Manlove, A. Ockenfels, and W. Pettersson (2024). Designing a kidney exchange program in germany: simulations and recommendations. *Central European Journal of Operations Research*, 1–27.

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

Ashlagi, I. and A. E. Roth (2012). New challenges in multihospital kidney exchange. *American Economic Review 102*(3), 354–359.

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

Benedek, M., P. Biró, G. Csáji, M. Johnson, D. Paulusma, and X. Ye (2023). Computing balanced solutions for large international kidney exchange schemes when cycle length is unbounded. *arXiv preprint arXiv:2312.16653*.

Benedek, M., P. Biró, D. Paulusma, and X. Ye (2024). Computing balanced solutions for large international kidney exchange schemes. *Autonomous Agents and Multi-Agent Systems 38*(1), 18.

Biró, P., B. Haase-Kromwijk, T. Andersson, E. I. Ásgeirsson, T. Baltesová, I. Boletis, C. Bolotinha, G. Bond, G. Böhmig, L. Burnapp, et al. (2019). Building kidney exchange programmes in europe—an overview of exchange practice and activities. *Transplantation 103*(7), 1514–1522.

Biró, P., J. Van de Klundert, D. Manlove, W. Pettersson, T. Andersson, L. Burnapp, P. Chromy, P. Delgado, P. Dworczak, B. Haase, et al. (2021). Modelling and optimisation in european kidney exchange programmes. *European Journal of Operational Research 291*(2), 447–456.

Csáji, G., T. Király, and Z. Mészáros-Karkus (2024). The ntu partitioned matching game for international kidney exchange programs. *arXiv preprint arXiv:2409.01452*.

Delorme, M., S. Garcia, J. Gondzio, J. Kalcsics, D. Manlove, W. Pettersson, and J. Trimble (2022). Improved instance generation for kidney exchange programmes. *Computers & Operations Research 141*, 105707.

Dickerson, J., A. Kazachkov, A. Procaccia, and T. Sandholm (2017). Small representations of big kidney exchange graphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 31.

Dickerson, J. P. and T. Sandholm (2014). Balancing efficiency and fairness in dynamic kidney exchange. In *AAAI Workshop: Modern Artificial Intelligence for Health Analytics*, Volume 14, pp. 08.

Druzsin, K., P. Biró, X. Klimentova, and R. Fleiner (2024). Performance evaluation of national and international kidney exchange programmes with the enckep simulator. *Central European Journal of Operations Research 32*(4), 923–943.

Echenique, F., S. Goel, and S. Lee (2024). Stable allocations in discrete exchange economies. *Journal of Economic Theory 222*, 105921.

Jantschgi, S., T. Nguyen, and A. Teytelboym (2025). Competitive combinatorial exchange. In *Proceedings of the 26th ACM Conference on Economics and Computation (EC '25)*, Stanford, CA, USA. ACM.

Kern, W., P. Biró, D. Pálvölgyi, and D. Paulusma (2019). Generalized matching games for international kidney exchange. In *18th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2019*, pp. 413–421. Association for Computing Machinery.

Klimentova, X., A. Viana, J. P. Pedroso, and N. Santos (2021). Fairness models for multi-agent kidney exchange programmes. *Omega 102*, 102333.

Milgrom, P. (2009). Assignment messages and exchanges. *American Economic Journal: Microeconomics 1*(2), 95–113.

Mincu, R.-S., P. Biró, M. Gyetvai, A. Popa, and U. Verma (2021). Ip solutions for international kidney exchange programmes. *Central European Journal of Operations Research 29*, 403–423.

Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*(2), 265–281.

Nguyen, H., T. Nguyen, and A. Teytelboym (2021). Stability in matching markets with complex constraints. *Management Science 67*(12), 7438–7454.

Nguyen, T. and R. Vohra (2018). Near-feasible stable matchings with couples. *American Economic Review 108*(11), 3154–3169.

Pettersson, W. and J. Trimble (2021). Kidney matching tools data set generator. https://wpettersson.github.io/kidney-webapp/.

Roth, A. E., T. Sönmez, and M. U. Ünver (2004). Kidney exchange. *The Quarterly journal of economics 119*(2), 457–488.

Saidman, S. L., A. E. Roth, T. Sönmez, M. U. Ünver, and F. L. Delmonico (2006). Increasing the opportunity of live kidney donation by matching for two-and three-way exchanges. *Transplantation 81*(5), 773–782.

Scarf, H. E. (1967). The core of an n person game. *Econometrica: Journal of the Econometric Society*, 50–69.

Schrijver, A. (1998). *Theory of linear and integer programming*. John Wiley & Sons.

Sönmez, T. and M. U. Ünver (2013). Market design for kidney exchange. *The Handbook of Market Design 93137*.

## Appendix A. Omitted Proofs for Lower Bounds

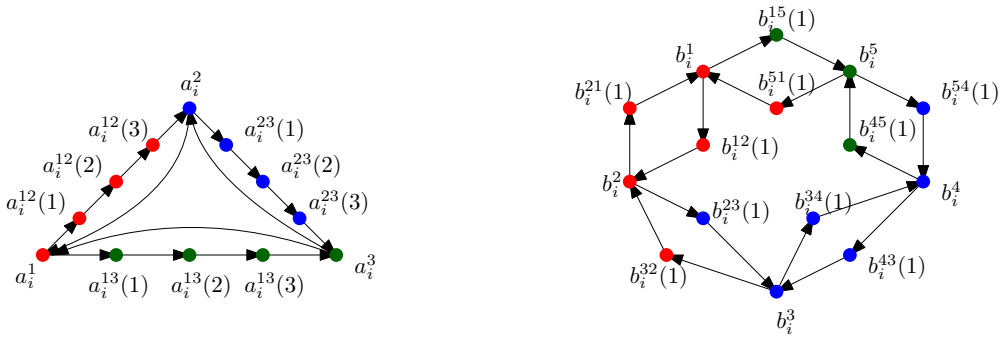



FIGURE 6. An illustration for Theorem 1 with $\Delta = 5$. A type-A graph is on the left and a type-B graph is on the right. The index $i$ is such that $i \equiv 0 \bmod 2$ and $i \equiv 1 \bmod 3$. Player 1 owns the red, player 2 the blue and player 3 the green vertices.

### A.1. **Proof of Theorem 1.**

*Proof.* **Proof of Nonexistence for General $\Delta$.** We create an instance of a partition exchange economy with exchange bound $\Delta$ as follows.

The graph $\mathcal{G}$ of the partition exchange economy consists of $60x$ vertex-disjoint type-A graphs and $180x$ vertex-disjoint type-B graphs, where $x \in \mathbb{N}$ is a parameter.

Type-A graphs have $3(\Delta-1)$ vertices and are essentially achieved from a triangle of two-way compatibility arcs, by subdividing some arcs. The $i$-th type-A graph $\mathcal{G}_i^A$ has vertices $a_i^1, a_i^2, a_i^3$, $a_i^{12}(1), \ldots, a_i^{12}(\Delta-2)$, $a_i^{13}(1), \ldots, a_i^{13}(\Delta-2)$, $a_i^{23}(1), \ldots, a_i^{23}(\Delta-2)$. The set of compatibility arcs are $\{(a_i^1, a_i^{12}(1)), (a_i^{12}(1), a_i^{12}(2)), \ldots, (a_i^{12}(\Delta-3), a_i^{12}(\Delta-2)), (a_i^{12}(\Delta-2), a_i^2), (a_i^2, a_i^1)\}$, $\{(a_i^2, a_i^{23}(1)), (a_i^{23}(1), a_i^{23}(2)), \ldots, (a_i^{23}(\Delta-3), a_i^{23}(\Delta-2)), (a_i^{23}(\Delta-2), a_i^3), (a_i^3, a_i^2)\}$ and $\{(a_i^1, a_i^{13}(1)), (a_i^{13}(1), a_i^{13}(2)), \ldots, (a_i^{13}(\Delta-3), a_i^{13}(\Delta-2)), (a_i^{13}(\Delta-2), a_i^3), (a_i^3, a_i^1)\}$.

If $i$ is even, then player 1 owns the vertices $a_i^1, a_i^{12}(1), \ldots, a_i^{12}(\Delta-2)$, player 2 owns $a_i^2, a_i^{23}(1), \ldots, a_i^{23}(\Delta-2)$ and player 3 owns $a_i^3, a_i^{13}(1), \ldots, a_i^{13}(\Delta-2)$. If $i$ is odd, then player 1 owns the vertices $a_i^1, a_i^{13}(1), \ldots, a_i^{13}(\Delta-2)$, player 2 owns $a_i^2, a_i^{12}(1), \ldots, a_i^{12}(\Delta-2)$ and player 3 owns $a_i^3, a_i^{23}(1), \ldots, a_i^{23}(\Delta-2)$.

Type-B graphs $\mathcal{G}_B^i$ have $10\lfloor\frac{\Delta-2}{2}\rfloor+5$ vertices and are essentially achieved from a five long cycle of two-way compatibility arcs, by subdividing the arcs. Let us use the notation $\delta = \lfloor\frac{\Delta-2}{2}\rfloor$. The $i$-th type-B graph $\mathcal{G}_i^B$ has vertices $\{b_i^l \mid l \in [5]\}$ and $\{b_i^{l(l+1)}(1), \ldots, b_i^{l(l+1)}(\delta), b_i^{(l+1)l}(1), \ldots, b_i^{(l+1)l}(\delta) \mid l \in [5]\}$, where we let $5+1 := 1$ in the upper indices (if $\delta = 0$, then no such vertices are added). The set of compatibility arcs are $\{(b_i^l, b_i^{l(l+1)}(1)), (b_i^{l(l+1)}(1), b_i^{l(l+1)}(2)), \ldots, (b_i^{l(l+1)}(\delta), b_i^{l+1}), (b_i^{l+1}, b_i^{(l+1)l}(1)), (b_i^{(l+1)l}(1), b_i^{(l+1)l}(2)), \ldots, (b_i^{(l+1)l}(\delta), b_i^l) \mid l \in [5]\}$, where again we let $5+1 := 1$ in the upper indices.

If $i \equiv 1 \bmod 3$, then player 1 owns $b_i^1, b_i^{51}(1), \ldots, b_i^{51}(\delta), b_i^{21}(1) \ldots, b_i^{21}(\delta), b_i^2, b_i^{12}(1), \ldots, b_i^{12}(\delta)$, $b_i^{32}(1), \ldots, b_i^{32}(\delta)$, player 2 owns $b_i^3, b_i^{23}(1), \ldots, b_i^{23}(\delta), b_i^{43}(1), \ldots, b_i^{43}(\delta), b_i^4, b_i^{34}(1), \ldots, b_i^{34}(\delta)$, $b_i^{54}(1)$, $b_i^{54}(\delta)$, and player 3 owns $b_i^5, b_i^{45}(1), \ldots, b_i^{45}(\delta), b_i^{15}(1), \ldots, b_i^{15}(\delta)$.

If $i \equiv 2 \bmod 3$, then player 2 owns $b_i^1, b_i^{51}(1), \ldots, b_i^{51}(\delta), b_i^{21}(1) \ldots, b_i^{21}(\delta), b_i^2, b_i^{12}(1), \ldots, b_i^{12}(\delta)$, $b_i^{32}(1), \ldots, b_i^{32}(\delta)$, player 3 owns $b_i^3, b_i^{23}(1), \ldots, b_i^{23}(\delta), b_i^{43}(1), \ldots, b_i^{43}(\delta), b_i^4, b_i^{34}(1), \ldots, b_i^{34}(\delta)$, $b_i^{54}(1)$, $b_i^{54}(\delta)$, and player 1 owns $b_i^5, b_i^{45}(1), \ldots, b_i^{45}(\delta), b_i^{15}(1), \ldots, b_i^{15}(\delta)$.

If $i \equiv 0 \bmod 3$, then player 3 owns $b_i^1, b_i^{51}(1), \ldots, b_i^{51}(\delta), b_i^{21}(1) \ldots, b_i^{21}(\delta), b_i^2, b_i^{12}(1), \ldots, b_i^{12}(\delta)$, $b_i^{32}(1), \ldots, b_i^{32}(\delta)$, player 1 owns $b_i^3, b_i^{23}(1), \ldots, b_i^{23}(\delta), b_i^{43}(1), \ldots, b_i^{43}(\delta), b_i^4, b_i^{34}(1), \ldots, b_i^{34}(\delta)$, $b_i^{54}(1)$, $b_i^{54}(\delta)$, and player 2 owns $b_i^5, b_i^{45}(1), \ldots, b_i^{45}(\delta), b_i^{15}(1), \ldots, b_i^{15}(\delta)$.

We have that $|V| = 60x \cdot 3(\Delta - 1) + 180x \cdot (10\delta + 5) = O(x)$. For an illustration of type-A and type B graphs, see Figure 6.

It is easy to see that the feasible exchange cycles for a type-A graph $\mathcal{G}_i^A$ are the $\Delta$-long cycles containing either $\{a_i^1, a_i^{12}(1), \ldots, a_i^{12}(\Delta-2), a_i^2\}$, or $\{a_i^2, a_i^{23}(1), \ldots, a_i^{23}(\Delta-2), a_i^3\}$ or $\{a_i^1, a_i^{13}(1), \ldots, a_i^{13}(\Delta-2), a_i^3\}$. Furthermore, exactly one can be chosen in any exchange, as any two of them have a vertex in common.

The feasible exchange cycles for a type-B graph $\mathcal{G}_i^B$ are the cycles containing $\{(b_i^l, b_i^{l(l+1)}(1), \ldots, b_i^{l(l+1)}(\delta), b_i^{l+1}, b_i^{(l+1)l}(1), \ldots, b_i^{(l+1)l}(\delta)\}$ for some $l \in [5]$. Here, at most two, non-adjacent exchange cycles can be chosen simultaneously.

It follows that in total, the 3 players together can cover only $60x\Delta + 180x(4\delta + 4)$ vertices, which is either $840x(\delta + 1)$ or $840x(\delta + 1) + 60x$, depending on whether $\Delta$ is even or odd respectively.

A single player can achieve an exchange only in those type-B graphs, where he owns $4\delta + 2$ vertices. In each of them, it can cover $2\delta + 2$ of his vertices, totaling $120x \cdot 2(\delta + 1) = 240x(\delta + 1)$.

If two players join, then we claim that together they can form an exchange in the type-A graphs, where $60x\Delta$ of their vertices are covered in each. In odd indexed cases, $\Delta - 1$ are covered from one player and 1 of the other, and in even indexed cases, the situation is reversed. Hence, both players have $30x\Delta$ vertices covered among type-A graphs. This arrangement is uniquely best for both of them among type-A graphs. This is either $60x(\delta + 1)$ or $60x(\delta + 1) + 30x$, depending on whether $\Delta$ is even, or odd, respectively.

For any two players, among the type-B graphs, there are $120x$ of them, where one of them only controls $2\delta + 1$ vertices, while the other controls $4\delta + 2$. In such cases, they can cover $2\delta + 2$ vertices, in two ways. Either $\delta + 1$ from each, or $2\delta + 2$ from one, and 0 from the other. In the $60x$ type-B graphs, where both have $4\delta + 2$ vertices, they can cover $2\delta + 2$ from both of them, which is uniquely optimal. Hence, in total they can cover $120x(2\delta + 2) + 60x(4\delta + 4) = 480x(\delta + 1)$. Observe, that among the type-B graphs, they can distribute the cycles in a way such that both of them receives a multiple of $(\delta + 1)$ transplants and both of them receives at least $180x(\delta + 1)$.

Suppose the $3x$-supplemented core is nonempty, where we have seen that $3x = O(|V|)$. Since any directed path in the compatibility graph has length less than $10\delta + 10 = 10(\delta + 1)$, it follows that even with these altruistic donors, the total number of transplants is at most $840x(\delta+1)+30x(\delta+1)$ or $840x(\delta + 1) + 30x(\delta + 1) + 60x$, based on the parity of $\Delta$. Hence, the two players who are worst off receive at most $580x(\delta + 1)$ or $580x(\delta + 1) + 40x$ transplants respectively.

Since neither of them blocks by themselves, each has at least $240x(\delta+1)$. If they join, then they can each get $60x(\delta + 1)$ or $60x(\delta + 1) + 30x$ among the type-A graphs, depending on whether $\Delta$ is even or odd, respectively. Furthermore among the type-B graphs, as we have observed, they can divide $480x(\delta + 1)$ in a way such that both of them receive a multiple of $(\delta + 1)$ transplants and both of them receives at least $180x(\delta + 1)$ (and so at most $300x(\delta + 1)$.

We get that no matter the parity of $\Delta$, the two players can have, in total, at least $20x(\delta+1)$ more transplants together, than what they received originally. As both of them had at least $240x(\delta + 1)$

transplants originally, the other had at most $340x(\delta+1)$ (or $340x(\delta+1)+40x$). Since $\delta+1\geq 1$, in both cases, we get that they need at most $290x(\delta+1)$ from the type-B graphs to match their original utility.

Therefore, using again that in any arrangement where both of them receives at most $300x(\delta+1)$ is possible among the type-B graphs, they form a blocking coalition, as they can divide the extra $20x(\delta+1)$ donors in a way that both of them improve by at least $10x(\delta+1)$. This is a contradiction, showing that the $3x = O(|V|)$-supplemented core is empty. □ □

## Appendix B. Omitted Proofs for Pairwise Exchange

### B.1. **Proof of Remark 4.1.**

*Proof.* Proof. In the proof of Theorem 2, we chose an arbitrary matching that leaves only one vertex from each odd cycle $c\in\mathcal{C}$.

We will show that if we choose this matching properly, then it satisfies the claimed additional constraint.

Create a graph $\mathcal{H}$ as follows. The set of vertices of $\mathcal{H}$ is $N$ and there is an arc $ij$, if there exists an odd cycle $c\in\mathcal{C}$, where $i$ has a covered vertex and $j$ has an uncovered vertex. Let $A$ be the set of those organizations such that they have strictly less than $\lceil|U^i\cap\mathcal{C}|/3\rceil$ vertices uncovered and $B$ be the set of those organizations such that they have strictly more than $\lceil|U^i\cap\mathcal{C}|/3\rceil$ vertices uncovered. If $B$ is empty, we are done. If $B$ is nonempty, then by the maximality of the chosen matching, $A$ is also nonempty.

We claim that there is a path from $A$ to $B$ in $\mathcal{H}$. Suppose otherwise. Then, the set of vertices $\hat{B}$ not reachable from $A$ all satisfy that they have strictly more than $\sum_{i\in\hat{B}}|U^i\cap\mathcal{C}|/3$ uncovered vertices. Furthermore, as there is no incoming edge to $\hat{B}$ in $\mathcal{H}$, it follows that any odd cycle that contains an uncovered vertex of a organization from $\hat{B}$ satisfies that all other vertices also belongs to someone from $\hat{B}$. As every odd cycle has at least 3 vertices and each odd cycle satisfies $c\subset\mathcal{G}[\cup_{i\in N}U^i]$ as we observed in the proof of Theorem 2, we get that if we sum the number of uncovered vertices of each organization in $\hat{B}$, then it is at most one-third of the total number of vertices of these organizations in $\mathcal{C}$, contradicting the above observation about $\hat{B}$.

Hence, in this case we can find an $A-B$ path in $\mathcal{H}$ along which we can create a new matching, where someone from $A$ has one less vertex covered and someone from $B$ has one more vertex covered. Iterating this, we arrive at a point, where $B$ will be empty.

Hence, for each organization, $\frac{|U^i|}{3}$ additional donors suffice. □ □

### B.2. **Proof of Lemma 4.1.**

*Proof.* Proof Let $n:=|V|$ and condition on an arbitrary assignment of the vertices to the $g$ groups in $\Phi$. Recall that

$$p:=\min\{p_{ab}:p_{ab}>0\}>0.$$

No assumption is imposed on the diagonal probabilities $p_{aa}$.

For each vertex $v$, define its augmented class as

$$\tau(v)=(a,r)\in\Phi\times\{\mathrm{u},\mathrm{a}\},$$

where $a$ is the group of $v$, and $r = \mathrm{u}$ if $v$ is utility-bearing and $r = \mathrm{a}$ if $v$ represents an existing altruistic donor. Thus, there are at most $2g$ augmented classes.

For two augmented classes $\theta$ and $\theta'$, let $q_{\theta\theta'}$ denote the probability that two distinct vertices in these classes are mutually compatible. Each of the two required directed arcs is either deterministic, because its target is a dummy patient, or appears independently with probability at least $p$. Consequently,

$$q_{\theta\theta'} > 0 \quad \Longrightarrow \quad q_{\theta\theta'} \geq p^2.$$

In particular, for two utility-bearing classes corresponding to groups $a$ and $b$,

$$q_{\theta\theta'} = p_{ab} p_{ba},$$

including $q_{\theta\theta} = p_{aa}^2$ when $a = b$. Moreover, mutual-edge indicators associated with distinct unordered pairs of vertices are independent.

Let $\nu'(\mathcal{G})$ denote the maximum number of independent edges in the mutual-compatibility graph; equivalently, $\nu'(\mathcal{G})$ is the maximum size of an induced matching. Let

$$X(\mathcal{G}) := \nu'(\mathcal{G}).$$

Selecting one edge from each cycle in an independent collection of odd cycles gives an induced matching. Therefore,

$$\nu(\mathcal{G}) \leq X(\mathcal{G}).$$

It is thus sufficient to show that $\mathbb{E}[X] = \mathcal{O}(\log n)$ and that $X = \mathcal{O}(\log n)$ with probability approaching one.

Let

$$h := \binom{2g+1}{2},$$

which bounds the number of unordered pairs of augmented classes. Suppose that $\mathcal{G}$ contains an induced matching of size $k$. By the pigeonhole principle, this matching contains at least $\ell := \lceil \frac{k}{h} \rceil$ edges whose endpoints have the same unordered pair of augmented classes, say $\{\theta, \theta'\}$.

Consider a fixed collection of $\ell$ disjoint candidate edges of this class-pair type. If $q_{\theta\theta'} = 0$, the probability that these candidate edges form an induced matching is zero. Suppose therefore that $q_{\theta\theta'} > 0$.

First, let $\theta \neq \theta'$, and write the candidate edges as $\{x_i, y_i\}$, where $x_i$ belongs to $\theta$ and $y_i$ belongs to $\theta'$. For these edges to form an induced matching, every off-matching pair $\{x_i, y_j\}$ with $i \neq j$ must be a nonedge. There are $\ell(\ell-1)$ such pairs.

If $\theta = \theta'$, all $2\ell$ endpoints belong to the same augmented class. In this case, inducedness requires all nonmatching pairs among the endpoints to be absent. In particular, it requires at least $\ell(\ell-1)$ distinct potential edges to be absent.

Thus, in either case,

$$\Pr(\text{the fixed candidate edges form an induced matching}) \leq (1-p^2)^{\ell(\ell-1)} \leq \exp\{-p^2 \ell(\ell-1)\}.$$

There are at most $n^{2\ell}$ collections of $\ell$ candidate edges and at most $h$ possible unordered class-pair types. Hence,

$$\Pr(X \geq h\ell) \leq h n^{2\ell} \exp\{-p^2 \ell(\ell-1)\}.$$

Set

$$\ell_0 := \left\lceil \frac{8}{p^2} \log(n+2) \right\rceil \qquad \text{and} \qquad k_0 := h\ell_0.$$

For all sufficiently large $n$,

$$p^2 \ell_0(\ell_0 - 1) \geq 4\ell_0 \log(n+2).$$

It follows that

$$\Pr(X \geq k_0) \leq h \exp\{-2\ell_0 \log(n+2)\} \leq n^{-3}.$$

Therefore,

$$\Pr\big(\nu(\mathcal{G}) \leq k_0\big) \geq \Pr(X \leq k_0) \longrightarrow 1.$$

Since $X \leq n$, we also have

$$\begin{aligned} \mathbb{E}[\nu(\mathcal{G})] &\leq \mathbb{E}[X] \\ &\leq k_0 + n \Pr(X \geq k_0) \\ &\leq h \left\lceil \frac{8}{p^2} \log(n+2) \right\rceil + n^{-2} = \mathcal{O}_{g,p}(\log n). \end{aligned}$$

The bounds are uniform over the assignment of vertices to groups, so they continue to hold without conditioning on that assignment. □

## Appendix C. Logarithmic bound beyond pairwise exchange

C.1. **Three-Way Exchanges in the Random Compatibility Model.** This appendix proves Lemma 4.2. The proof first aggregates a fractional exchange by finitely many augmented cycle patterns, then rounds the aggregate pattern counts, and finally realizes the rounded counts in the random compatibility graph.

*Augmented classes and cycle patterns.* Let

$$V = \bigcup_{i \in N} V^i, \qquad U = \bigcup_{i \in N} U^i, \qquad A := V \setminus U,$$

where $A$ is the set of existing altruistic donors, represented by vertices with dummy patients. Let $m := |V|$.

Recall that the set of random-graph groups is denoted by $\Phi$, where $|\Phi| = g$. Every vertex belongs to one group. If the patient represented by $v$ is not a dummy patient, then, for vertices $u$ and $v$ in groups $a$ and $b$, respectively, the arc $(u, v)$ appears independently with probability $p_{ab}$. If $v \in A$, then its patient is a dummy patient and $(u, v)$ is present deterministically for every $u \neq v$.

We impose no restriction on the diagonal probabilities $p_{aa}$ and define $p := \min\{p_{ab} : p_{ab} > 0\}$. Both $g$ and $p > 0$ are constants. For a vertex $v$, let

$$\tau(v) = \begin{cases} \mathrm{u}, & v \in U, \\ \mathrm{a}, & v \in A. \end{cases}$$

An *augmented class* is a pair

$$\theta = (a, r) \in \Phi \times \{\mathrm{u}, \mathrm{a}\},$$

where $a$ is the random-graph group and $r$ records whether the vertex is utility-bearing or altruistic. Let $V_\theta$ denote the vertices of augmented class $\theta$.

For augmented classes $\theta = (a, r)$ and $\theta' = (b, r')$, we say that the transition $\theta \to \theta'$ is *admissible* if either $r' = \mathrm{a}$, in which case the target patient is dummy and the arc is deterministic, or $r' = \mathrm{u}$ and $p_{ab} > 0$.

An *augmented cycle pattern* of length $k \in \{2, 3\}$ is a cyclic equivalence class of ordered tuples

$$\pi = (\theta_1, \ldots, \theta_k)$$

such that every transition $\theta_j \to \theta_{j+1}$ is admissible, where indices are taken modulo $k$. Tuples that differ only by cyclic rotation represent the same pattern. For each pattern, fix one representative. Let $\mathcal{P}$ denote the set of augmented cycle patterns, and let $P := |\mathcal{P}|$. Since $g$ is constant,

$$P \leq (2g)^2 + (2g)^3 = \mathcal{O}(g^3).$$

For $\pi \in \mathcal{P}$, let $|\pi| \in \{2, 3\}$ denote its length. For an augmented class $\theta$, let $\mu_\theta^\pi$ denote the number of occurrences of $\theta$ in $\pi$. Define

$$w(\pi) := \sum_{\theta=(a,\mathrm{u})} \mu_\theta^\pi,$$

the number of utility-bearing slots of $\pi$. A pattern may contain the same augmented class more than once; this can occur either because same-group compatibility has positive probability or because a transition enters an altruistic vertex and is therefore deterministic.

*A uniform random-graph event.*

Definition C.1. *For an integer $h \geq 1$, let $\mathcal{R}_h$ be the event that the following properties hold.*

- *(R1) For every pair of augmented classes $\theta, \theta'$ for which the transition $\theta \to \theta'$ is admissible and random, and for all disjoint sets $X \subseteq V_\theta$, $Y \subseteq V_{\theta'}$ with $|X|, |Y| \geq h$, there is an arc from $X$ to $Y$.*
- *(R2) For every augmented length-two pattern represented by $(\theta_1, \theta_2)$ and all disjoint sets $X \subseteq V_{\theta_1}$ and $Y \subseteq V_{\theta_2}$ with $|X|, |Y| \geq h$, there are $x \in X$ and $y \in Y$ such that $x \to y \to x$ is a directed two-cycle.*
- *(R3) For every augmented length-three pattern represented by $(\theta_1, \theta_2, \theta_3)$, every set $M \subseteq V_{\theta_1} \times V_{\theta_2}$ of pairwise vertex-disjoint ordered pairs satisfying $x \to y$ for every $(x, y) \in M$, and every set $Z \subseteq V_{\theta_3}$ disjoint from the vertices occurring in $M$, if $|M|, |Z| \geq h$, then there are $(x, y) \in M$ and $z \in Z$ such that $y \to z \to x$.*

The disjointness conditions above ensure that the relevant directed-edge variables are distinct even when two or more augmented classes coincide.

Transitions whose target is altruistic are deterministic and therefore cannot cause any of these properties to fail. Every random transition in an admissible pattern has probability at least $p$.

Lemma C.1. *There is a constant $K = K(g, p)$ such that, for $h := \lceil K \log(m + 2) \rceil$, we have*

$$\Pr(\mathcal{R}_h) \geq 1 - (m + 2)^{-4}.$$

*Proof.* Proof. Write $L := \log(m+2)$ and choose $K \geq 18/p^2$. We establish each property for fixed augmented classes or a fixed pattern and then take a union bound over the constantly many choices.

First consider (R1). If the transition is deterministic, the assertion holds automatically. Otherwise, fix sets $X, Y$ of sizes $s, t$, respectively. The probability that there is no arc from $X$ to $Y$ is at most

$$(1-p)^{st} \leq \exp(-pst) \leq \exp(-p^2 st).$$

There are at most $m^{s+t}$ choices for $X$ and $Y$. Since $s, t \geq h$,

$$st \geq \frac{h}{2}(s+t).$$

Consequently,

$$\begin{aligned}(s+t)\log m - p^2 st &\leq (s+t)L - \frac{p^2 h}{2}(s+t) \\ &\leq -8(s+t)L.\end{aligned}$$

Thus, the probability that (R1) fails for fixed augmented classes is at most

$$\begin{aligned}\sum_{s,t\geq h} \exp\big((s+t)\log m - p^2 st\big) &\leq \sum_{s,t\geq h} (m+2)^{-8(s+t)} \\ &= \mathcal{O}\big((m+2)^{-16}\big).\end{aligned}$$

For (R2), fix sets $X, Y$ of sizes $s, t$. For any $x \in X$ and $y \in Y$, each required arc is either deterministic or appears independently with probability at least $p$. Hence, the probability that $x \to y \to x$ is a directed two-cycle is either one or at least $p^2$. The events corresponding to distinct pairs $(x, y)$ depend on disjoint directed-edge variables. Therefore,

$$\begin{aligned}&\Pr(\text{there is no directed two-cycle between } X \text{ and } Y) \\ &\leq (1-p^2)^{st} \leq \exp(-p^2 st).\end{aligned}$$

The same calculation as for (R1) gives a failure probability $\mathcal{O}((m+2)^{-16})$ for each fixed length-two pattern.

Finally, consider (R3). Fix a family $M$ of $s$ pairwise vertex-disjoint ordered pairs and a set $Z$ of size $t$. For a fixed $(x, y) \in M$ and $z \in Z$, each of the arcs $y \to z$ and $z \to x$ is either deterministic or occurs independently with probability at least $p$. The probability that both arcs exist is therefore either one or at least $p^2$. As $(x, y)$ ranges over $M$ and $z$ ranges over $Z$, all non-deterministic directed-edge variables used by these completion events are distinct. Hence,

$$\Pr(\text{no pair in } M \text{ can be completed using } Z) \leq \exp(-p^2 st).$$

There are at most $m^{2s}$ choices for $M$ and at most $m^t$ choices for $Z$. Since $s, t \geq h$,

$$st \geq \frac{h}{3}(2s+t).$$

Therefore,

$$\begin{aligned}(2s+t)\log m - p^2 st &\leq (2s+t)L - \frac{p^2 h}{3}(2s+t) \\ &\leq -5(2s+t)L.\end{aligned}$$

The failure probability for a fixed length-three pattern is at most

$$\sum_{s,t\geq h}(m+2)^{-5(2s+t)} = \mathcal{O}\big((m+2)^{-15}\big).$$

There are only $\mathcal{O}(g^3)$ augmented patterns. A union bound proves the result for all sufficiently large $m$; increasing $K$ if necessary handles the remaining values of $m$. □ □

*Realizing prescribed pattern counts.* We use the following consequence of Hall's theorem.

LEMMA C.2. *Let $H = (L, R; F)$ be a bipartite graph with $|L| = |R| = q$. Suppose that there are no sets $X \subseteq L$ and $Y \subseteq R$ such that $|X|, |Y| \geq h$ and there is no edge between $X$ and $Y$. Then $H$ has a matching of size at least $q - h$.*

*Proof.* Proof. Let a maximum matching have size $q - d$. By the deficiency form of Hall's theorem, there is a set $X \subseteq L$ such that

$$d = |X| - |N_H(X)|.$$

Set $Y := R \setminus N_H(X)$. There is no edge between $X$ and $Y$, and

$$d = |X| + |Y| - q.$$

If $|X| < h$, then $d \leq |X| < h$. Similarly, if $|Y| < h$, then $d \leq |Y| < h$. Otherwise, $X$ and $Y$ contradict the assumed absence of a large empty rectangle. Hence $d < h$, proving the claim. □ □

LEMMA C.3. *Assume that $\mathcal{R}_h$ holds. Let*

$$\pi = (\theta_1, \ldots, \theta_k) \in \mathcal{P}, \qquad k \in \{2, 3\},$$

*and suppose that, for every occurrence of a slot in $\pi$, we are given a set of exactly $q$ vertices of the corresponding augmented class. Assume that the sets assigned to distinct slots are pairwise disjoint, even when two slots have the same augmented class.*

*Then $\mathcal{G}$ contains vertex-disjoint directed cycles of pattern $\pi$ that leave at most $2h$ supplied vertices uncovered in each slot. Consequently, at most*

$$2h|\pi| \leq 6h$$

*supplied vertices are left uncovered in total.*

*Proof.* Proof. If $q < 2h$, discard all supplied vertices. The claimed bound follows immediately. Suppose therefore that $q \geq 2h$.

First let $k = 2$, and let the two supplied slot sets be $X_1$ and $X_2$. Construct a bipartite graph between $X_1$ and $X_2$, where $x_1 \in X_1$ and $x_2 \in X_2$ are adjacent if

$$x_1 \to x_2 \to x_1$$

is a directed two-cycle. By (R2), this bipartite graph has no empty rectangle with both sides of size at least $h$. Lemma C.2 gives a matching of size at least $q - h$. Thus, at most $h$ vertices remain uncovered in each slot.

Now let $k = 3$, and denote the supplied slot sets by $X_1, X_2, X_3$. Consider first the transition from the first slot to the second. If this transition is deterministic, choose an arbitrary perfect

matching from $X_1$ to $X_2$. Otherwise, by (R1) and Lemma C.2, the bipartite graph of arcs from $X_1$ to $X_2$ has a matching $M$ of size at least $q-h$.

Thus, in either case, we obtain a family $M$ of at least $q-h$ pairwise vertex-disjoint directed paths $x_1 \to x_2$. Choose an arbitrary set $X_3' \subseteq X_3$ with $|X_3'| = |M|$. Construct an auxiliary bipartite graph with left side $M$ and right side $X_3'$. A path $(x_1, x_2) \in M$ is adjacent to $x_3 \in X_3'$ if

$$x_2 \to x_3 \to x_1.$$

By (R3), this auxiliary graph has no empty rectangle with both sides of size at least $h$. Hence, Lemma C.2 gives a matching of size at least

$$|M| - h \geq q - 2h.$$

Every matched pair gives a directed cycle

$$x_1 \to x_2 \to x_3 \to x_1.$$

Therefore, at most $2h$ supplied vertices remain uncovered in each slot. □ □

*Aggregate-pattern rounding.*

LEMMA C.4. *Assume that $\mathcal{R}_h$ holds. Let $y$ be a fractional exchange using cycles of size at most three, and define*

$$b_i := \left\lfloor \sum_{c \in \mathcal{C}_3} \gamma_c^i y_c \right\rfloor \qquad (i \in N).$$

*There is an integral exchange $\mathcal{E}$, using cycles of size at most three and only vertices from the original pool $V$, such that*

$$\sum_{i \in N} \big(b_i - u_i(\mathcal{E})\big)^+ \leq 9Ph.$$

*Proof.* Proof. For each augmented pattern $\pi \in \mathcal{P}$, let

$$\lambda_\pi := \sum_{\substack{c \in \mathcal{C}_3: \\ c \text{ has augmented pattern } \pi}} y_c, \qquad q_\pi := \lfloor \lambda_\pi \rfloor.$$

For every augmented class $\theta$, define its fractional and rounded usages by

$$R_\theta := \sum_{\pi \in \mathcal{P}} \mu_\theta^\pi \lambda_\pi, \qquad S_\theta := \sum_{\pi \in \mathcal{P}} \mu_\theta^\pi q_\pi.$$

Since $q_\pi \leq \lambda_\pi$, we have $S_\theta \leq R_\theta$. Moreover, the vertex-capacity constraints of the fractional exchange imply

$$R_\theta \leq |V_\theta|.$$

Only utility-bearing slots matter for organizational utilities. Define the loss caused by flooring the augmented-pattern weights on utility-bearing slots as

$$\begin{aligned} \delta_U &:= \sum_{\theta=(a,\mathrm{u})} (R_\theta - S_\theta) \\ &= \sum_{\pi \in \mathcal{P}} w(\pi)(\lambda_\pi - q_\pi). \end{aligned}$$

Since $w(\pi) \leq |\pi| \leq 3$,

$$\delta_U < 3P. \tag{1}$$

Step 1: Assigning the rounded slots to organizations. For every organization $i$ and every utility-bearing augmented class $\theta = (a, \mathrm{u})$, let

$$U^{i,\theta} := U^i \cap V_\theta, \qquad n_{i,\theta} := |U^{i,\theta}|,$$

and define

$$R_{i,\theta} := \sum_{c \in \mathcal{C}_3} |c \cap U^{i,\theta}| y_c.$$

Then

$$0 \leq R_{i,\theta} \leq n_{i,\theta}, \qquad \sum_{i \in N} R_{i,\theta} = R_\theta,$$

and

$$\sum_{\theta=(a,\mathrm{u})} R_{i,\theta} = \sum_{c \in \mathcal{C}_3} \gamma_c^i y_c \geq b_i.$$

For each utility-bearing class $\theta$, reduce the vector $(R_{i,\theta})_{i \in N}$ by total amount $R_\theta - S_\theta$, obtaining numbers $F_{i,\theta}$ such that

$$0 \leq F_{i,\theta} \leq R_{i,\theta}, \qquad \sum_{i \in N} F_{i,\theta} = S_\theta.$$

The total amount removed over all utility-bearing classes is $\delta_U$. Therefore,

$$\sum_{i \in N} \left( b_i - \sum_{\theta=(a,\mathrm{u})} F_{i,\theta} \right)^+ \leq \delta_U. \tag{2}$$

We now round this fractional assignment without increasing its total utility deficit. Consider the following maximum-cost flow network. There is a source $s$, one node for every utility-bearing class $\theta$, one node for every organization $i$, and a sink $t$. The arc $s \to \theta$ has lower and upper capacity both equal to the integer $S_\theta$, while the arc $\theta \to i$ has capacity $n_{i,\theta}$. Finally, from each organization node $i$ to $t$ add two parallel arcs: a unit-profit arc of capacity $b_i$, and a zero-profit arc of capacity $B := \sum_{\theta=(a,\mathrm{u})} S_\theta$ (which is large enough to be nonbinding). Thus a unit sent on $\theta \to i$ represents assigning one selected vertex of class $\theta$ to organization $i$, and the capacity $n_{i,\theta}$ enforces that no more such vertices are assigned than actually exist.

All supplies and arc capacities in this network are integral. Hence the feasible-flow polytope is integral: after the usual elimination of fixed lower bounds, its constraint matrix is the node–arc incidence matrix of a directed graph together with bound constraints, and is therefore totally unimodular. In particular, the maximum-cost flow problem has an integral optimum.

To see that the fractional assignment $F$ gives a feasible fractional flow, put $F_{i,\theta}$ units on the arc $\theta \to i$. For $x_i := \sum_{\theta=(a,\mathrm{u})} F_{i,\theta}$, send $\min\{b_i, x_i\}$ units from $i$ to $t$ on the unit-profit arc and the remaining $x_i - \min\{b_i, x_i\}$ units on the zero-profit arc. Its profit is therefore

$$\sum_{i \in N} \min\{b_i, x_i\} = \sum_{i \in N} b_i - \sum_{i \in N} (b_i - x_i)^+ \geq \sum_{i \in N} b_i - \delta_U,$$

where the inequality follows from (2). Let an integral optimum of the flow problem be fixed, and let $m_{i,\theta}$ be its flow on $\theta \to i$. Flow conservation and the arc capacities give

$$0 \le m_{i,\theta} \le n_{i,\theta}, \qquad \sum_{i \in N} m_{i,\theta} = S_\theta$$

for every utility-bearing class $\theta$. Moreover, at optimality the unit-profit arc out of organization $i$ carries $\min\{b_i, \sum_\theta m_{i,\theta}\}$ units; otherwise flow could be moved from the zero-profit arc to the unit-profit arc and strictly increase the objective. Since the integral optimum has profit at least that of the fractional flow above, we obtain

$$\sum_{i \in N} \left( b_i - \sum_{\theta=(a,\mathrm{u})} m_{i,\theta} \right)^+ \le \lceil \delta_U \rceil \le 3P. \tag{3}$$

Step 2: Selecting vertices for the rounded patterns. For each organization $i$ and each utility-bearing class $\theta$, choose exactly $m_{i,\theta}$ vertices from $U^{i,\theta}$. Thus, exactly $S_\theta$ selected utility-bearing vertices belong to class $\theta$.

For every altruistic class $\theta = (a, \mathrm{a})$,

$$S_\theta \le R_\theta \le |V_\theta|.$$

Choose an arbitrary set of exactly $S_\theta$ existing altruistic vertices from $V_\theta$. These vertices are needed as cycle positions, but their coverage contributes no utility.

For every augmented class $\theta$,

$$S_\theta = \sum_{\pi \in \mathcal{P}} \mu_\theta^\pi q_\pi.$$

We can therefore partition the selected vertices of class $\theta$ among all occurrences of $\theta$ in the pattern copies so that every pattern $\pi$ receives $q_\pi$ pairwise disjoint vertex sets for each of its slots, each of size $q_\pi$. This remains valid when the same augmented class occurs more than once in a pattern because distinct slot occurrences receive disjoint sets.

Step 3: Realizing the rounded patterns. Apply Lemma C.3 separately to every pattern $\pi$. At most

$$2h|\pi| \le 6h$$

selected vertices assigned to one pattern remain uncovered. Hence, this gives an integral exchange $\mathcal{E}$ in which at most $6Ph$ selected vertices remain uncovered in total.

Only uncovered utility-bearing vertices can reduce organizational utility. Consequently, the total utility deficit created in the realization step is at most $6Ph$. Combining this with (3) gives

$$\sum_{i \in N} \big(b_i - u_i(\mathcal{E})\big)^+ \le 3P + 6Ph \le 9Ph,$$

where the last inequality uses $h \ge 1$. □ □

*Proof.* Proof of Lemma 4.2. Let

$$h := \lceil K \log(m+2) \rceil,$$

where $K = K(g,p)$ is chosen as in Lemma C.1. By that lemma,

$$\Pr(\mathcal{R}_h) \ge 1 - (m+2)^{-4}.$$

Conditional on $\mathcal{R}_h$, Lemma C.4 applies to every fractional exchange $y$ and gives an integral exchange $\mathcal{E}$ satisfying

$$\sum_{i\in N}\left(\left\lfloor\sum_{c\in\mathcal{C}_3}\gamma_c^i y_c\right\rfloor - u_i(\mathcal{E})\right)^+ \le 9Ph.$$

Because

$$P \le (2g)^2 + (2g)^3$$

and $h = \mathcal{O}_{g,p}(\log(m+2))$, there is a constant $C_0 = C_0(g,p)$ such that

$$9Ph \le C_0 \log(m+2).$$

The event $\mathcal{R}_h$ is defined entirely in terms of the realized compatibility graph, so the conclusion holds simultaneously for every fractional exchange $y$. □ □

C.2. **Extending the logarithmic bound to constant $\Delta$.** In this section, we show that the argument of Theorem 4 extends to every fixed constant $\Delta$.

THEOREM 7. *Consider a partition exchange economy with a constant exhange bound $\Delta$, where the compatibility graph on the original vertices $V$, including the existing altruistic donors in $V \setminus U$, is generated according to the random model above.*

*Then, there is a constant $C = C(g,p,\Delta)$ such that, if $d(\mathcal{G})$ denotes the minimum number of* additional *altruistic donors required for a supplemented-core exchange, then*

$$\mathbb{E}[d(\mathcal{G})] \le C\log(|V|+2)$$

*and*

$$\Pr\left(d(\mathcal{G}) \le C\log(|V|+2)\right) \ge 1 - (|V|+2)^{-4}.$$

*The expectation and probability are taken over the realization of the random compatibility arcs. The supplemented-core exchange can be chosen Pareto-optimal.*

*Proof.* Proof. For $2 \le k \le \Delta$, define an augmented pattern of length $k$ as a cyclic tuple $\pi = (\theta_1,\dots,\theta_k)$ of augmented classes with admissible consecutive transitions. Let $\mathcal{P}_\Delta$ denote the set of such patterns and let $P_\Delta := |\mathcal{P}_\Delta|$. Since $g$ and $\Delta$ are fixed,

$$P_\Delta \le \sum_{k=2}^{\Delta}(2g)^k = \mathcal{O}_{g,\Delta}(1).$$

The uniform random event $\mathcal{R}_h$ is strengthened as follows. For every pattern of length $k$, every family $M$ of $s$ pairwise vertex-disjoint directed partial paths occupying the first $k-1$ slots, and every set $Z$ of $t$ available vertices for the final slot, whenever $s,t \ge h$, some path in $M$ can be completed with a vertex of $Z$. The two required final arcs are either deterministic or appear independently with probability at least $p$. Thus, for fixed $M$ and $Z$, the probability that no completion exists is at most $\exp(-p^2st)$.

There are at most $m^{(k-1)s+t}$ choices for $M$ and $Z$. Since $s,t \ge h$, we have $st \ge \frac{h}{k}\big((k-1)s+t\big)$. Therefore,

$$\big((k-1)s+t\big)\log m - p^2st \le -\left(\frac{p^2h}{k} - \log(m+2)\right)\big((k-1)s+t\big).$$

Choosing $h = \lceil K_\Delta \log(m+2) \rceil$ for a sufficiently large constant $K_\Delta = K(g,p,\Delta)$ makes the failure probability summable. Taking a union bound over all lengths $k \le \Delta$ and all augmented patterns gives a failure probability of at most $(m+2)^{-4}$. To realize $q$ copies of a pattern of length $k$, first match the first slot to the second slot, then extend the resulting disjoint paths one slot at a time through slots $3, \dots, k-1$, and finally match the partial paths to the last slot using the two-edge completion relation. Every matching step loses fewer than $h$ paths. Hence, at least $q - (k-1)h$ cycles are obtained. At most $(k-1)h$ supplied vertices remain uncovered in each slot, and thus at most $k(k-1)h \le \Delta(\Delta-1)h$ vertices are lost for one pattern.

Flooring the aggregate pattern weights causes utility loss less than

$$\sum_{\pi \in \mathcal{P}_\Delta} w(\pi) \le \Delta P_\Delta.$$

The realization step causes utility loss at most $\Delta(\Delta-1)P_\Delta h$. Therefore, for every fixed $\Delta$,

$$d(\mathcal{G}) \le \Delta P_\Delta + \Delta(\Delta-1)P_\Delta h = \mathcal{O}_{g,p,\Delta}(\log(|V|+2))$$

on the high-probability event. Using the trivial bound

$$d(\mathcal{G}) \le |U| \le |V|$$

on the complementary event also gives

$$\mathbb{E}[d(\mathcal{G})] = \mathcal{O}_{g,p,\Delta}(\log(|V|+2)).$$

Existing altruistic donors cause no difficulty in this extension. Their pattern slots are included in all capacity calculations and in the realization procedure, but uncovered altruistic slots are never charged as utility losses because vertices in $V^i \setminus U^i$ do not contribute to $u_i$. □ □

## Appendix D. Deterministic Type Representation and Related Results

D.1. **Type-representation of graphs:** Another related but conceptually distinct approach to model KEP compatibility graphs is the deterministic type-based model (Dickerson et al. 2017). In contrast to the probabilistic model, compatibility here is deterministic: whether one pair is compatible with another is fully determined by their assigned types. However, unlike in the random graph model above where types reflect observable donor and patient characteristics (e.g., blood type, cPRA), the "type" in this setting is an abstract construct introduced solely to parametrize the compatibility graph. It does not necessarily correspond to any specific medical or demographic attributes. Specifically, we say that a (di)graph $\mathcal{G} = (V, E)$ *can be represented by a set of types* $T$, if it holds that there exists a labeling $f : V \to T$, such that for any two types $t_1, t_2 \in T$, either $\{(u,v) \mid u \in f^{-1}(t_1), v \in f^{-1}(t_2)\} \subseteq E$ or $\{(u,v) \mid u \in f^{-1}(t_1), v \in f^{-1}(t_2)\} \cap E = \emptyset$. That is, the existence of a (directed) edge only depends on the types of the endpoints. Dickerson et al. (2017) conducted simulations showing that kidney exchange graphs can usually be represented by only a few types. We may assume that for any type $t_i$, there is no edge $(u,v)$ with $f(u) = f(v) = t_i$, as donor-recipient pairs are not added to KEPs, if they are compatible with each other.

D.2. **Pairwise Exchanges.** In this section, we show that if the compatibility graph can be represented using types, the maximum number of altruists needed is at most one-third the number of types.

For this, we first show that the number of independent odd cycles is bounded.

LEMMA D.1. *If a graph $\mathcal{G}$ can be represented with $t$ types, then $\nu(\mathcal{G}) \leq \frac{t}{3}$.*

*Proof.* Proof. We show that there exists an independent set of odd cycles of size $\nu(\mathcal{G})$ that contains each type at most once. As odd cycles have at least 3 vertices, the statement will follow.

Take any maximum size set of independent odd cycles $\mathcal{C}$. Suppose for the contrary that $t_i$ appears in cycles $c_{j_1}, c_{j_2}$. Then the neighbors of the type $t_i$ vertex in $c_{j_1}$ are connected to the type $t_i$ vertex in $c_{j_2}$ too with an edge, which contradicts the independence of the cycles. Similarly, if a type $t_i$ appears twice in the same odd cycle $c_j$, then there exists an edge between two non-adjacent (as a type is never compatible with itself) vertices in the cycle. Hence, we can obtain a smaller odd cycle $c'_j$ and replace $c_j$ with $c'_j$ in $\mathcal{C}$. This still gives an independent set of odd cycles, and after a finite number of steps, no type will appear twice within an odd cycle either. □ □

Thus, we obtain the following corollary establishing the existence of the $\frac{t}{3}$-supplemented core.

COROLLARY D.1. *In a partition exchange economy representable by $t$ types, the $\frac{t}{3}$-supplemented core is always nonempty.*

D.3. **Cyclic Exchanges.** We next give a deterministic guarantee for arbitrary bounded cycle length. If the compatibility graph admits a representation by a constant number of types, then the number of additional donors required by each organization is bounded independently of the market size.

THEOREM 8. *Given a partition exchange economy with a compatibility graph that can be represented by $t$ types and a bound $\Delta \geq 2$ on the exchange cycle lengths, the $(3t+1)(\Delta-1)n$ -supplemented core is always nonempty. Furthermore, it contains a Pareto-optimal solution.*

*If the original endowments contain no altruistic vertices, i.e., $V^i = U^i$ for every $i \in N$, the bound improves to*

$$(2t+1)(\Delta-1)n.$$

*Proof.* Proof. Take a type representation $f : V \to T$ of the original compatibility graph, where $|T| = t$. We first make one harmless refinement that is important for keeping track of utility. For every $v \in V^i$, define its *refined type*

$$\widehat{f}(v) := \begin{cases} (f(v), \mathrm{u}), & v \in U^i, \\ (f(v), \mathrm{a}), & v \in V^i \setminus U^i. \end{cases}$$

and let $\widehat{T}$ be the set of nonempty refined types. Splitting a type class into subclasses cannot destroy a type representation, because compatibility was already constant between every pair of original type classes. Hence compatibility also depends only on the refined types. We have $|\widehat{T}| \leq 2t$, and, crucially, every refined type consists entirely of utility-bearing vertices or entirely of altruistic vertices.

By Lemma 2.2, there exists a fractional exchange $y^*$ with the following property. Put

$$b_i := \left\lfloor \sum_{c \in \mathcal{C}_\Delta} \gamma_c^i y_c^* \right\rfloor, \qquad \gamma_c^i := |c \cap U^i|.$$

If an exchange $\mathcal{E}$ in an extension by additional altruistic donors satisfies $u_i(\mathcal{E}) \geq b_i$ for every $i \in N$, then $\mathcal{E}$ belongs to the corresponding supplemented core.

For a refined type $\tau \in \widehat{T}$ and organization $i$, let

$$\widehat{\beta}_c^{\tau,i} := |\{v \in c \cap V^i : \widehat{f}(v) = \tau\}|, \qquad \widehat{a}_{\tau,i} := |\{v \in V^i : \widehat{f}(v) = \tau\}|.$$

Consider the polyhedron

$$\mathcal{P} = \left\{ z \in [0,1]^{\mathcal{C}_\Delta} : \sum_c \widehat{\beta}_c^{\tau,i} z_c \leq \widehat{a}_{\tau,i} \ \forall (\tau, i), \quad \sum_c \gamma_c^i z_c \geq b_i \ \forall i \right\}.$$

Rows of the former type are denoted by $\widehat{A}_{\tau,i}$ and rows of the latter type are denoted by $B_i$. The fractional exchange $y^*$ belongs to $\mathcal{P}$: the first family of constraints follows from the vertex-capacity constraints of a fractional exchange, after summing them over each refined type class, and the second family follows from the definition of $b_i$.

We apply an iterative-rounding procedure using the refined-type rows. At every iteration we take an extreme point $z^*$ that maximizes

$$\sum_{c \in \mathcal{C}_\Delta} \left( \sum_{i \in N} \gamma_c^i \right) z_c^*.$$

Integral variables are fixed and deleted, with all right-hand sides updated. For the residual system, we delete a refined-type upper-bound row $(\tau, i)$ whenever

$$\widehat{A}_{\tau,i} \mathbf{1} \leq \widehat{a}_{\tau,i} + \Delta - 1,$$

and delete a utility lower-bound row $i$ whenever

$$B_i z^* \leq \Delta - 1.$$

We then recompute an objective-maximizing extreme point of the residual polyhedron.

Suppose that all remaining variables are strictly fractional and no row is eligible for deletion. Give one token to each fractional variable $z_c^*$. Variable $z_c^*$ gives $\gamma_c^i \frac{z_c^*}{\Delta}$ tokens to row $B_i$, and $\widehat{\beta}_c^{\tau,i} \frac{1 - z_c^*}{\Delta}$ tokens to row $\widehat{A}_{\tau,i}$. Since every cycle has length at most $\Delta$,

$$\sum_i \gamma_c^i \leq |c| \leq \Delta, \qquad \sum_{i,\tau} \widehat{\beta}_c^{\tau,i} = |c| \leq \Delta,$$

so $z_c^*$ distributes at most one token in total.

Every tight remaining $B_i$ row receives at least one token: indeed, its updated right-hand side is integral and, since the row is not deletable,

$$B_i z^* = b_i \geq \Delta.$$

Likewise, every tight remaining refined-type row receives at least one token. If $\widehat{A}_{\tau,i}z^* = \widehat{a}_{\tau,i}$, then nondeletability and integrality imply

$$\widehat{A}_{\tau,i}\mathbf{1} \geq \widehat{a}_{\tau,i} + \Delta,$$

and therefore

$$\widehat{A}_{\tau,i}(\mathbf{1} - z^*) \geq \Delta.$$

By the extreme-point characterization, the number of variables is at most the number of linearly independent tight rows. Since the variables jointly own only one token each while every tight row needs at least one, equality must hold everywhere. In particular every nonzero row touching a fractional column is present and tight, and each fractional column $c$ must satisfy

$$\sum_i \gamma_c^i = \Delta, \qquad \sum_{i,\tau} \widehat{\beta}_c^{\tau,i} = \Delta.$$

Consequently, summing all tight refined-type rows gives the same row vector, namely the vector having entry $\Delta$ in every remaining column, as summing all tight utility rows. Thus the tight rows are linearly dependent, a contradiction. Hence the iterative procedure terminates with an integral vector $z^*$.

It remains to turn this integral vector into an actual vertex-disjoint exchange. This is the point at which we explicitly distinguish two effects of rounding: a violation of a refined-type capacity, and a loss of utility. For every organization $i$ and refined type $\tau$, write

$$q_{\tau,i} := \sum_{c \in \mathcal{C}_\Delta} \widehat{\beta}_c^{\tau,i} z_c^*, \qquad r_{\tau,i} := (q_{\tau,i} - \widehat{a}_{\tau,i})^+.$$

Also put

$$s_i := (b_i - B_i z^*)^+.$$

By the rule under which rows were deleted,

$$r_{\tau,i} \leq \Delta - 1 \quad \text{and} \quad s_i \leq \Delta - 1. \tag{4}$$

Regard every occurrence of a vertex in every selected cycle $c$ with $z_c^* = 1$ as a *slot*; the slot remembers the organization and refined type of that vertex. For each pair $(\tau, i)$, assign distinct original vertices of refined type $\tau$ from organization $i$ to $\min\{q_{\tau,i}, \widehat{a}_{\tau,i}\}$ of its slots. For each of the remaining $r_{\tau,i}$ slots introduce a distinct additional altruistic donor. We charge this donor to organization $i$ only for counting purposes; it still belongs to the market designer and is not a player. As in the standard existential interpretation of a supplemented-core upper bound, the set of additional donors may be chosen with the compatibilities required by the construction. Choose this donor's kidney to be compatible with the patient in the successor slot of that cycle. Equivalently, one may choose it to have the same outgoing compatibility type as the original type underlying $\tau$. The incoming arc into the new altruistic vertex is automatic because its patient is dummy. Hence replacing a missing slot by this donor preserves both arcs needed at that position of the cycle.

Because distinct original vertices were assigned to all nonreplaced slots and all newly introduced donors are distinct, the resulting cycles are vertex-disjoint. Replacing an original vertex by a donor does not change the number of positions in its cycle, so every resulting cycle still has length at

most $\Delta$. Compatibility of two neighboring original vertices is preserved because vertices of the same refined type are interchangeable; if a neighbor is a new altruist, feasibility follows from the preceding paragraph. Thus we obtain a genuine exchange, denoted by $\mathcal{E}_0$.

The preceding construction repairs every type-capacity violation, but a donor used in a slot of a utility-bearing refined type has a dummy patient and hence cannot contribute the transplant that the abstract slot contributed to $B_i z^*$. This is exactly the loss identified in the referee's objection. Let

$$R_i^{\mathrm{u}} := \sum_{\tau=(t,\mathrm{u})} r_{\tau,i}, \qquad R_i^{\mathrm{a}} := \sum_{\tau=(t,\mathrm{a})} r_{\tau,i}.$$

Since every utility-bearing slot that is filled by an original vertex gives one transplant and precisely $R_i^{\mathrm{u}}$ utility-bearing slots are filled by dummy-patient donors, we have the exact identity

$$u_i(\mathcal{E}_0) = B_i z^* - R_i^{\mathrm{u}}. \tag{5}$$

Therefore the remaining utility deficit of organization $i$ is

$$d_i := (b_i - u_i(\mathcal{E}_0))^+ \le s_i + R_i^{\mathrm{u}}.$$

There are at least $d_i$ unmatched utility-bearing vertices of organization $i$: indeed $b_i \le |U^i|$, while $\mathcal{E}_0$ uses exactly $u_i(\mathcal{E}_0)$ distinct patients of $i$. Choose any $d_i$ such unmatched vertices $v_1, \dots, v_{d_i}$. For each $v_j$, introduce one further altruistic donor $a_j$ whose kidney is compatible with the patient of $v_j$. Then

$$(a_j, v_j)$$

is a feasible directed 2-cycle: $a_j \to v_j$ is feasible by the choice of the donor, while $v_j \to a_j$ is automatic because $a_j$ has a dummy patient. These cycles are mutually disjoint and disjoint from $\mathcal{E}_0$. Adding them to $\mathcal{E}_0$ gives an exchange $\mathcal{E}$ with

$$u_i(\mathcal{E}) \ge b_i \qquad \forall i \in N. \tag{6}$$

This second family of donors is necessary in the argument: a donor that replaces a missing utility-bearing slot repairs the donor position of that cycle, but its dummy patient means that the corresponding patient transplant has disappeared. The additional 2-cycle above restores precisely that lost unit of utility (and also covers any deficit caused by deletion of a $B_i$ row).

We finally count the donors. The slot-repair stage uses $R_i^{\mathrm{u}} + R_i^{\mathrm{a}}$ donors for organization $i$, and the utility-restoration stage uses $d_i \le s_i + R_i^{\mathrm{u}}$ donors. Thus

$$D_i \le R_i^{\mathrm{a}} + 2R_i^{\mathrm{u}} + s_i.$$

For each original type there is at most one nonempty utility-bearing refined type and at most one nonempty altruistic refined type. Hence, by (4),

$$R_i^{\mathrm{u}} \le t(\Delta - 1), \qquad R_i^{\mathrm{a}} \le t(\Delta - 1), \qquad s_i \le \Delta - 1,$$

and consequently

$$D_i \le (3t + 1)(\Delta - 1). \tag{7}$$

Summing over all organizations gives at most $(3t+1)(\Delta-1)n$ additional altruistic donors. If $V^i = U^i$ for all organizations, no altruistic refined type occurs, so $R_i^{\mathrm{a}} = 0$ and the sharper bound $D_i \leq (2t+1)(\Delta-1)$ follows.

By (6) and Lemma 2.2, $\mathcal{E}$ belongs to the claimed supplemented core. Finally, if $\mathcal{E}$ is not Pareto-optimal in the same extended graph $\mathcal{G} + V^0$, choose an exchange $\mathcal{E}'$ that Pareto-dominates $\mathcal{E}$ and maximizes $\sum_i u_i(\mathcal{E}')$. Then $\mathcal{E}'$ is Pareto-optimal and still satisfies $u_i(\mathcal{E}') \geq u_i(\mathcal{E}) \geq b_i$ for every $i$. Another application of Lemma 2.2 therefore shows that $\mathcal{E}'$ is in the same supplemented core. □ □