Sequential Kidney Exchange

Lawrence M. Ausubel and Thayer Morrill*

January 2013

Abstract

The literature on kidney exchange considers situations where two or more patients needing transplants have live donors volunteering to donate one of their kidneys, but the donated organs are incompatible with the respective patients. The traditional analysis assumes that all components of the live donor exchange must occur simultaneously. People cannot write enforceable contracts that commit them to donate their organs; consequently, incentive compatibility is obtained by trading simultaneously. Unfortunately, a two-way exchange then requires the simultaneous availability of four operating rooms and associated personnel, while a three-way exchange requires six operating rooms, etc. The requirement of four or more operating rooms for concurrent surgeries may pose a significant constraint on the beneficial exchanges that may be attained. The basic insight of this paper is that satisfaction of the incentive constraint does not require simultaneous exchange; rather, it requires that organ donation occurs no later than the associated organ receipt. Using sequential exchanges may relax the operating-room constraint and thereby increase the number of beneficial exchanges. We show that most benefits of sequential exchange can be accomplished with only two concurrent operating rooms.

1. Introduction

Kidney exchange provides a vivid illustration of the challenges and potential of market design. The idiosyncratic constraints of the problem are not mere technicalities to be abstracted away, but rather lie at the very heart of the market design problem. First and foremost, kidney exchange faces the constraint that a market, in the usual sense of the word, is illegal. This creates the obvious problem that we may not buy or sell kidneys but instead must exchange one kidney for another. However, it also creates a more subtle incentive constraint. An agent cannot write a contract compelling another agent to donate her kidney if she has already received a kidney in kind. As a result, the order in which kidneys are exchanged is crucial to the exchange being incentive compatible. For this reason, exchanges have been performed simultaneously so that neither party has the incentive to renege on the agreement.

* Department of Economics, University of Maryland, College Park, MD 20742, USA and Department of Economics, North Carolina State University, Raleigh, NC 27695, USA. Email: ausubel@econ.umd.edu and Thayer_Morrill@ncsu.edu. We gratefully acknowledge the support of National Science Foundation Grants SES-05-31254 and SES-09-24773.

However, this creates an additional constraint that the market designer must overcome. Exchanges must take place in close proximity and there is a limit to the number of organ transplants that can be performed simultaneously in the same hospital. We call this the hospital capacity constraint. Even a two-way exchange involves four simultaneous surgeries. Therefore, in most instances to date, kidney exchanges have been limited to two-way exchanges.

Roth, Sonmez, and Unver (2007), hereafter RSU, discuss the challenges and potential gains from an efficiently designed kidney exchange market.[2] They demonstrate that expanding the number of possible exchanges to include three-way as well as two-way exchanges would substantially increase the number of possible exchanges. In fact, recent research by Ashlagi et al. (2012) indicates that there is a substantial efficiency loss from limiting exchanges to even three-way or four-way exchanges. This is due to the fact that in practice there is a higher percentage of highly sensitized patient in an exchange pool than in the general patient population, and as a result, these patients are more difficult to match using only short cycles.

One obvious approach for making three-way or higher exchanges feasible is to replace simultaneous operations with appropriately sequenced operations. A requirement of simultaneous trade is more stringent than necessary. Rather, incentive compatibility continues to be satisfied if, for every donor-patient pair, the donation occurs *no later than* the associated receipt of a kidney.

In this paper, we exhibit theoretical environments where there are potential benefits to sequential kidney exchange and we take the insight to its logical conclusion. In particular, with a stationary population of agents, sequential kidney exchange allows us to achieve the maximal number of transplants while preserving incentive compatibility and yet never requiring more than two simultaneous operations. With a population consisting partly of recurring agent types and partly of unique agent types, sequential kidney exchange can be utilized to ease the hospital capacity constraint, both for recurring types and for unique types.

In many ways, this is analogous to the classic treatment of retirement savings in Samuelson's overlapping generations (OLG) model. The basic problem of retirement savings is that each generation would like to produce goods in the first period of its life and to consume goods in the second period of its life. However, the goods are perishable, so any generation cannot save directly for its own future. The problem is resolved in the OLG model by having, in each period, the current working-age generation produce goods for the previous generation — with the expectation that, in their retirement, goods will be produced for them by the next generation. This arrangement is incentive compatible on account that each generation is required to give up goods before receiving goods. Sequencing in the opposite direction would not be incentive compatible.

[2] As an indicator of the magnitude of this problem, as of December 13, 2009, there were 81,678 patients on the cadaver kidney waiting list in the United States. In 2008, 32,587 patients were added to the waiting list while 29,207 were removed. Of the patients removed, 4,746 patients died and 1,600 were removed because they became too sick to receive a transplant.

Similarly, in the kidney exchange problem, it may not be feasible for each donor (of a set of donor-patient pairs) to give up a kidney and for each patient (of the same set of donor-patient pairs) to receive a kidney simultaneously. However, it may be feasible for each donor to give up a kidney in period t and for each of the associated patients to receive a kidney in period $t+1$, as fewer concurrent operations are required. Effectively, each donor donates a kidney to the previous "generation" and each patient receives a kidney from the next "generation". Moreover, there is no incentive barrier to this sequencing, provided that each donor gives before — *not* after — the associated patient receives.

We illustrate this with a simple example. Figure 1 is a simple illustration with three agents. Each agent consists of a patient needing a kidney and her incompatible donor. We represent each agent by the blood-type of the patient and the blood-type of the donor. See RSU for a detailed description of kidney compatibility, but a type-A (resp. B) patient is incompatible with a type-B (A) donor. In this example, only two patients may be accommodated if we limit exchanges to two parties. However, if we allow larger exchanges, then all three agents may receive a kidney.

INSERT FIGURE 1 HERE

Figure 2 illustrates the same example if we allow the agents to donate in different periods. Now, the (A,B) agent gives to the (B,A) agent in one period and receives a kidney from next period's (A,A) agent, etc. We are able to achieve full efficiency, and yet we never require more than two concurrent operating rooms.

INSERT FIGURE 2 HERE

In our formal model, an identical population of people enters every period. The relevant notion of an agent is a pair comprising a kidney patient and an associated live donor. We initially focus on sequential exchanges in which the donor contributes her kidney exactly one period before the associated patient receives a kidney from another donor. In our first proposition, we show that when exchanges are done sequentially, we can match the same number of patients as with any simultaneous exchange in the static model. In our second proposition, we limit attention to stationary sequential exchanges and we show, conversely, that the number of patients matched cannot exceed the upper bound provided by the static model. The difference is that optimality in the static model may require n-way exchanges (requiring $2n$ operating rooms), whereas the sequential exchanges never require using more than two operating rooms.

Next, we consider the robustness of a sequential exchange to a non-stationary population. We demonstrate that to the extent that there is a subpopulation with isomorphic characteristics in each period then many of the same benefits may be realized as with a stationary population. Specifically, a sequential exchange may easily be modified in each period to accommodate a dynamic population so long as the designer continues to include the isomorphic subpopulation. The advantage is that any exchange to a member of the subpopulation may be implemented as a sequential exchange thereby relaxing the hospital capacity constraint.

We then consider "hybrid" exchanges in which transplants may occur both sequentially and simultaneously, and we consider longer waiting times than a single period. Although sequential exchanges ease the hospital capacity constraint, it comes at the cost of making patients wait an extra period to receive a kidney. A hybrid exchange serves as a compromise between these two tradeoffs. Suppose for example that a hospital has the capacity for a two-way exchange but not a three-way exchange. If the designer wishes to maximize the number of possible exchanges, subject to the capacity constraint, while minimizing the waiting time of each patient, then a hybrid exchange is superior to a sequential exchange. Compare Figure 2 to Figure 3. In the static exchange, at most two kidneys may be exchanged. In the sequential exchange, three kidneys are exchanged, only two hospital rooms are used at any given time, but all patients must wait a period to receive their kidney. In the hybrid exchange, three kidneys are exchanged, at most four hospital rooms are used at any given time, and only one patient must wait a period to receive her kidney. Our next proposition demonstrates that this tradeoff holds in general.

INSERT FIGURE 3 HERE

Observe that hybrid exchanges can only improve the number of feasible exchanges relative to the traditional analysis, as simultaneous exchanges are a special case. At the same time, it could be misleading to compare the number of exchanges possible in sequential exchanges with longer waiting times versus the number possible with only simultaneous exchanges, as we have effectively multiplied the population being matched by a factor related to the waiting time. This can be formalized by considering k-replicated economies in which there are k agents of each type. It turns out that there is a natural relationship between hybrid matching in which agents may be required to wait up to $2k - 1$ periods and static matching in the k-replicated economy. In our fourth proposition, we show that, for any static exchange in a k-replicated economy, there exists a corresponding hybrid exchange in the unreplicated but repeated economy with waiting times of up to $2k - 1$ periods, and in our fifth proposition, we establish the converse.

The most closely related literature on kidney exchange concerns non-simultaneous, extended, altruistic-donor (NEAD) chains. In a NEAD chain, an altruistic donor initiates a sequence of "domino transplantations" (Montgomery et al (2006), Roth et al. (2006), Rees et al. (2009)). Within the sequence, exchanges may be done simultaneously or sequentially. The principal difference from the sequential and hybrid kidney exchanges explored in the current paper is that the sequencing goes in the exact opposite direction. A NEAD chain creates a "bridge donor"—an agent who is asked to donate *after* the associated receipt of a kidney. While the medical literature typically does not use the language of incentive compatibility, it is concerned about "reneging risk". For example, in the *American Journal of Transplantation*, Gentry, Montgomery, Swihart and Segev (2009) write: "However, NEAD chains … run the risk of a bridge donor reneging, and add logistical complexity in that programs must maintain contact with bridge donors after a chain segment is completed." Empirically, they report that a NEAD chain at the Johns Hopkins Hospital was broken when a bridge donor reneged, and they conclude:

> A long wait between a donor's intended recipient getting a transplant and the donor's future nephrectomy could be a disadvantage if there is even a small chance that the donor will withdraw consent or become ineligible for health reasons. Additionally, it may be viewed as coercive to ask a donor's consent for his own nephrectomy many months after his intended recipient has been transplanted, especially if the recipient has had a poor outcome.

By way of contrast, there is no reneging risk in the sequential or hybrid kidney exchanges of our paper: an agent never receives a kidney before giving one. However, we retain the advantages of a NEAD chain. As exchanges are non-simultaneous, they reduce the logistical barriers to a many-agent exchange and may increase the number of agents that are able to be matched. For example, Rees et al. (2009) describes 10 kidney transplants initiated by a single altruistic donor.[3]

This paper is structured as follows. In Section 2, we develop sequential kidney exchange in a stationary population. In Section 3, we show robustness of our conclusions with a non-stationary population. Section 4 explores efficiency in a replicated economy. In Section 5, we conclude.

2. Efficiency in a stationary population

We begin by describing the static kidney exchange problem. Our primitive is the graph *G* representing the agents and their compatibilities for transplants. The graph has *N* nodes, representing the *N* agents. Each agent $a_i = (p_i, d_i)$ is a pair comprising a patient p_i and an associated donor d_i. We denote the set of agents by $X = \{a_1,...,a_N\}$. Edges of the graph are directional. There is an edge connecting agent a_i to agent a_j if and only if donor d_i is compatible with patient p_j; more formally:

$$e_{a_i,a_j} \in G \Leftrightarrow d_i \text{ and } p_j \text{ are compatible .} \tag{1}$$

Next, we describe the repeated version of the same problem. At every time $t \in \mathbb{Z}$, there is a set of agents $X_t \equiv \{a_{1t}, \dots, a_{Nt}\}$. Let $X^\infty \equiv \{X_t : t \in \mathbb{Z}\}$ be the set of all agents. Each agent $a_{it} = (p_{it}, d_{it})$ is a pair comprising a patient p_{it} and an associated donor d_{it}. The compatibilities of donors d_{is} and patients p_{jt} are exactly those induced by graph *G*:

[3] An alternative approach to a NEAD chain is a "domino paired donation" (DPD). A DPD is a chain of donations initiated by a non-directed donor. All exchanges are performed simultaneously and the donor in the last pair donates to a candidate on the waiting list. Ashlagi et al. (2010) run simulations using actual patient data from the Alliance for Paired Donation to compare the number of transplants that results from NEAD chains versus DPD. In particular, they compare relative performance for a range of renege rates for each bridge donor in a NEAD change. Even for relatively high renege rates, they find that NEAD chains outperform DPD when chains of length greater than four are allowed.

$$e_{a_i,a_j} \in G \Leftrightarrow d_{is} \text{ and } p_{jt} \text{ are compatible for all } s,t \in \mathbb{Z}. \tag{2}$$

We denote the graph in the repeated model by G^∞. We consider an infinitely repeated kidney exchange and compare the efficiency of exchanges that are constrained to be simultaneous with exchanges that may be across periods. In the next section, we relax the assumption that the population is isomorphic in each period.

If agent *j* is involved in an exchange, it must give a kidney to some agent *i* and receive a kidney from some agent *k*. Therefore, $\{e_{a_i,a_j}, e_{a_j,a_k}\} \subset G$. For the static model, this implies that a kidney exchange is a disjoint union of cycles.

Definition: Given a population *G*, a *cycle* is a set of agents $\{a_1,\ldots,a_n\} \subseteq X$ such that each agent appears only once, $e_{a_i,a_{i+1}} \in G, i \in \{1,\ldots,n-1\}$, and $e_{a_n,a_1} \in G$. A ***static kidney exchange*** is the disjoint union of cycles in *G*.

Remark: A static kidney exchange S induces a function $\mu : G \to G$ in a natural way. For each cycle $\{a_1,\ldots,a_n\} \in S$, define $\mu(a_i) = a_{i+1}$ for $1 \le i < n$ and $\mu(a_n) = a_1$. If an agent *a* is not part of any cycle in S, define $\mu(a) = \varnothing$. An agent *a* is said to be ***satisfied*** in a static kidney exchange if *a* receives a donated kidney, $\mu(a) \neq \varnothing$.

As mentioned earlier, institutional constraints may limit the length of an allowable kidney exchange. Therefore, define an ***n-way static kidney exchange*** to be the disjoint union of cycles of length no greater than *n*. Our objective is to maximize the number of agents that receive a kidney subject to the incentive and capacity constraints.

Definition: Given a static kidney exchange problem *G*, define $\Delta_n(G)$ to be the maximum-cardinality *n*-way static kidney exchange of *G*. Define $\Delta(G)$ to be the maximum unbounded static kidney exchange. Equivalently:

$$\Delta(G) = \lim_{n\to\infty} \Delta_n(G).$$

In the repeated model, initially we restrict our attention to the case where an agent in period *t* may only donate to an agent in period $t-1$. We consider more generalized exchanges in Section 4. An agent may only receive a kidney if she has already donated a kidney, and a kidney exchange is a matching so that each agent who donates a kidney receives a kidney.

Definition: A ***sequential kidney exchange*** is a one-to-one function $f : X^\infty \to X^\infty$ such that for every $a_{it} = (p_{it}, d_{it}) \in X_t$:

(1) $f(a_{it}) \in X_{t-1} \cup \varnothing$

(2) if $f(a_i) \neq \varnothing$, then $\mathrm{e}_{a_i,f(a_i)} \in G$ and $\exists a^* \in X_{t+1}$ such that $f(a^*) = a_{it}$

The key advantage of a sequential kidney exchange is that the hospital capacity constraint is no longer binding. In a static exchange, the smallest possible exchange, a two-way exchange, requires four hospital operating rooms. In a sequential exchange, each exchange only requires two hospital rooms. This is the best-case scenario as any non-cadaver donation requires two operating rooms. Note that the incentive constraint is still satisfied as each agent gives a kidney before she receives one.

In a sequential kidney exchange, the exchanges no longer consist of disjoint cycles but instead are infinite chains. Since there are no binding contracts, an agent must believe she will receive a kidney in the next period in order to be willing to donate a kidney in the current period. An agent has a clone in each period. If a previous period's clone did not receive a kidney, she may be reasonably skeptical that donating a kidney in this period will result in her receiving a kidney in the next period. However, if her clone has received a kidney in every previous (and infinitely many) periods, then she should be confident in the exchange. As a result, we focus on exchanges where the same population receives a kidney in every period. We call this a ***stationary exchange***.

First, we show that a sequential exchange does at least as well as the unbounded static kidney exchange.

Proposition 1: $\Delta(G)$ many agents may be matched in a stationary sequential kidney exchange while never requiring more than two operating rooms for any exchange.

Proof: Each static exchange corresponds to a stationary sequential exchange in a natural way. Let μ be any static exchange, and let *S* be the set of agents that exchange a kidney. $S = \{a_i : \mu(a_i) \neq \varnothing\}$. Consider any $a_j, a_k, a_l \in S$ such that $\mu(a_j) = a_k$ and $\mu(a_l) = a_j$. For every time *t*, define $f(a_{jt}) = a_{k(t-1)}$. By construction, $f(a_{l(t+1)}) = a_{jt}$, so indeed every agent involved in the static match gives and receives a kidney. Since every static kidney exchange corresponds to a sequential exchange, the maximal, unbounded, static, kidney exchange corresponds to a sequential exchange.

Q.E.D.

One might think that there is enough flexibility in a dynamic exchange to improve on the number of agents that are matched. Unfortunately and rather interestingly, there is not.

Proposition 2: The maximum number of agents matched in any sequential, stationary exchange is $\Delta(X)$.

Proof: Look at any sequential, stationary exchange *f*. Fix any period *t* and let S_t be the set of agents that donate a kidney in period *t*. Since *f* is stationary, $S_t = S_{t-1}$. Start with any $a_{x_1 t} \in S_t$ and let $a_{x_2 (t-1)}$ be the agent $a_{x_1 t}$ donates its kidney to. In general, let x_i be the index such that $f(a_{x_{i-1} t}) = a_{x_i (t-1)}$. Consider the sequence $\{a_{x_1 t}, a_{x_2 t}, a_{x_3 t}, \ldots\}$. Since S_t

is finite, the sequence must repeat an agent. Let $a_{x_m t}$ be the first agent repeated. If $x_m = x_i$ where $i > 1$, then

$$f(a_{x_{m-1}t}) = a_{x_m t} = a_{x_i t} = f(a_{x_{i-1}t}).$$

This implies $a_{x_{i-1}t} = a_{x_{m-1}t}$ which contradicts the minimality of *m*. Therefore $i = 1$ and $\{a_{x_1 t}, \ldots, a_{x_{m-1}t}\}$ is a cycle in *G*. Continue this process with any agent $a_{lt} \in S_t$, $\{a_{x_1 t}, \ldots, a_{x_{m-1}t}\}$. This produces a disjoint union of cycles that corresponds to a static kidney exchange in G. Therefore, $|S| \leq \Delta(G)$ by the definition of $\Delta(G)$. Proposition 1 implies that the number of agents matched in a maximal sequential kidney exchange is at least $\Delta(G)$.

Q.E.D.

3. Robustness to a non-stationary population

In real populations, even though some donor-patient pairs are comparatively rare in that one cannot count on a qualitatively similar agent to enter the pool for a long time, some donor-patient pairs are undoubtedly very common. This immediately suggests that a market designer could utilize sequential exchanges for common agent types and simultaneous exchanges for rare agent types, easing the operating-room constraint. In this section, we show that the market designer can do better than this, by leveraging the existence of the more common types to facilitate the treatment of the rarer types.

We model the non-stationary population by assuming that, in each period, there is a *recurring* subpopulation and a *unique* subpopulation. "Recurring" agents are donor-patient pairs that occur sufficiently often that the designer can count on a similar pair to be present next month, allowing the agent's participation to be sequential, as in Section 2. "Unique" agents cannot be relied upon to create sequential trades, but they can still benefit from sequentiality involving the recurring agents.

Intuitively, a static cycle involving the overall population will often include both recurring types and unique types. Whenever the cycle passes through a recurring type, we can make the agent's participation sequential, using the same device as in Section 2: the donor is taken to be in one cohort, while the recipient is taken to be in the next cohort. This eases the operating-room constraint, allowing for exchanges that could not otherwise occur. In this way, even unique agents, whose participation cannot be made sequential, benefit from the sequential participation of their trading partners.

As before, our primitive is the graph G representing the agents and their compatibilities for transplants. An "agent" is a pair comprising a patient and a donor. At every time $t \in \mathbb{Z}$, a set of agents X_t enters the population. Each set X_t can be partitioned into a recurring population, Y_t, and a unique population, Z_t. We have $X_t = Y_t \cup Z_t$ and

$Y_t \cap Z_t = \varnothing$, for all $t \in \mathbb{Z}$. Note that the population X_t is no longer restricted to have the same cardinality in each period. Let $X^\infty \equiv \{X_t : t \in \mathbb{Z}\}$ be the set of all agents. The recurring population, Y_t, is isomorphic in each period. Specifically, for each agent $a_{it} \in Y_t$, and any time *s,* there exists an $a_{is} \in Y_s$ such that a_{it} and a_{is} are compatible with the same set of agents. In particular, for any $a_{it} \in Y_t$ and $x_i \in X_t$, if x_i's donor is compatible with a_{it}'s patient, then x_i's donor is also compatible with $a_{i(t-1)}$'s patient.

In this section we continue to restrict a sequential exchange so that an agent receives a kidney no more than one period after donating a kidney. However, as the population varies each period, we will now allow a combination of simultaneous and sequential exchanges to occur.

Definition: A ***semi-sequential kidney exchange*** is a one-to-one function $f : X^\infty \to X^\infty$ such that for every $a_{it} = (d_{it}, p_{it}) \in X_t$:

(1) $f(a_{it}) \in X_{t-1} \cup X_t \cup \varnothing$

(2) if $f(a_{it}) \neq \varnothing$, then $e_{a_{it}, f(a_{it})} \in G$ and $\exists a^* \in X_t \cup X_{t+1}$ such that $f(a^*) = a_{it}$

Without loss of generality, we assume that, in every period *t,* there exists a static exchange of the agents X_t in which every agent in Y_t is satisfied ("satisfied" defined as in Section 2).[4] We add two more definitions:

Definition: $\{\mu_t\}_{t\in\mathbb{Z}}$ is a ***static exchange profile*** if each μ_t is a static exchange on X_t.

Definition: Let Δ_t^Y be the maximum-cardinality static kidney exchange on X_t subject to the constraint that all agents in Y_t are satisfied.

In the next proposition, we show that a static exchange profile in which every agent in Y_t is satisfied in each period *t* may be converted into a semi-sequential exchange in which each agent in Y_t participates sequentially. This shows that we are able to achieve Δ_t^Y many exchanges in each period while reducing the number of simultaneous operations that are required. Moreover, this is the reason why the recurring sub-population, Y, might be chosen to be less than maximal. If accommodating all the agents in *Y* imposes a significant constraint, then the social planner may improve efficiency by not including in *Y* the agents that are the most difficult to match. Note that omitting an agent from *Y* does not imply that the agent will not be matched; it just eliminates the constraint that the agent is satisfied in every period.

[4] If *Y* cannot be satisfied in every period, then redefine *Y* to be a maximal subset that can be satisfied in every period. For example, a static exchange consisting only of members in *Y* meets this criterion.

Figures 4 and 5 demonstrate how we may convert a variety of static exchanges into semi-sequential exchanges. If the static exchange has agent *a* giving to agent $y_{it} \in Y_t$, then the semi-sequential exchange simply has agent *a* give to $y_{i(t-1)} \in Y_{t-1}$, the copy of y_i that already donated a kidney in the $(t\text{-}1)^{\text{st}}$ period.

INSERT FIGURES 4 AND 5 HERE

Proposition 3: Consider any static kidney exchange profile in which all agents in Y are satisfied every period. Then the same trades can be accomplished in a semi-sequential kidney exchange in which any trade involving an agent in Y requires only two simultaneous operating rooms.

Proof: Consider any static kidney exchange profile $\{\mu_t\}_{t\in\mathbb{Z}}$ such that in every period *t*, each agent $y_{it} \in Y_t$ is satisfied. Specifically, for every period *t* and every $y_{it} \in Y_t$ there exists an agent $a_{it} \in X_t$ such that $\mu_t(a_{it}) = y_{it}$. We define a semi-sequential exchange as follows. Consider any agents $a_{it}, a_{jt} \in X_t$ such that $\mu_t(a_{it}) = a_{jt}$. If $a_j \in Y_t$, then set $f(a_{it}) = a_{i(t-1)}$. Otherwise, set $f(a_{it}) = a_{jt}$. This semi-sequential exchange modifies the static exchange profile in a natural way. Each agent in Y_t receives a kidney from an agent in period $t+1$. All other agents receive a kidney in the same period it donates. This exchange is valid as each agent in Y_t is satisfied in every period. As the exchanges involving agents in Y_t are done sequentially, they require only two simultaneous operating rooms.

Q.E.D.

Remark: Proposition 3 should not be misinterpreted to assert that every static kidney exchange profile can be restated as a semi-sequential kidney exchange requiring only two operating rooms. The static kidney exchange at time *t* could, for example, contain a cycle of *n* agents, all from the set Z_t. In that event, the corresponding semi-sequential exchange would still require 2*n* simultaneous operating rooms. However, to the extent that trades occur with agents in set Y, the operating-room constraint can be significantly relaxed. Specifically, if there are no more than *m* "consecutive" trades among members of Z_t, then 2*m* simultaneous operating rooms suffice. Figure 6(a) gives an example where the semi-sequential exchange does not reduce the operating constraint relative to a static exchange. Figure 6(b) gives an example of a static exchange that requires a 6-cycle while the corresponding semi-sequential exchange requires only two simultaneous rooms. The most beneficial case is one in which subset Y is sufficiently large within the set X_t that *all* trades involve at least one agent from set Y. In that event, we have the following corollary.

INSERT FIGURE 6 HERE

Corollary: Consider any static kidney exchange profile in which all agents in Y are satisfied every period and such that each agent outside of Y that receives a kidney donates to an agent in Y. Then the same trades can be accomplished in a semi-sequential kidney exchange profile with the following properties:

- No more than two simultaneous operating rooms are required; and
- No agent needs to wait more than one period.

When choosing the population Y to satisfy sequentially, the designer faces a trade-off. Each additional member of Y reduces the number of simultaneous operating rooms required. However, the population Y potentially constrains the number of exchanges that may be achieved in a given period. It should be noted that there are populations that may always be satisfied without imposing a constraint on the maximum number of matches. For example, consider an agent *a* where both the donor and the patient have blood type A but the donor's kidney is incompatible with the patient due to a positive crossmatch. As long as there is some patient with blood type A that receives a kidney in the maximal exchange, then *a* will always be included. For example, suppose a_i donates a kidney to a_j in an exchange and that a_j's patient has blood type A. Then if *a* where not included in the exchange, then we could increase the number of kidney's donated by having a_i donate to *a*, *a* donate to a_j, and otherwise leaving the exchange unchanged.

4. Efficiency in a replicated economy.

In this section, we examine a generalized sequential kidney exchange where the only restriction imposed is that an agent must give a kidney no later than when it receives a kidney. In particular, this allows an agent to participate in a simultaneous exchange, a sequential exchange where she donates a kidney but waits *multiple* periods before receiving a kidney, or a hybrid of the two. We show that there is a natural relationship between donors waiting up to $2k - 1$ periods to receive a kidney and a static population being replicated k times.

Definition: A general ***hybrid sequential kidney exchange*** is a one-to-one function $f : X^{\infty} \to X^{\infty}$ such that for every $a_{it} = (d_{it}, p_{it}) \in X_t$:

(1) $f(a_{it}) \in X_s \cup \varnothing, s \leq t$

(2) if $f(a_{it}) \neq \varnothing$, then $e_{a_i, f(a_i)} \in G$ and $\exists a^* \in X_r, r \geq t$, such that $f(a^*) = a_{it}$

In Section 2, we saw that a particularly simple form of sequential exchange, in which agents wait a single period between donating and receiving a kidney, relaxed the operating-room constraint. However, the cost of the sequential exchange is that all agents must wait a period to receive their kidney. Since we expect the hospital capacity constraint to be greater than 2, we first show how a hybrid exchange can be used to satisfy the capacity constraint yet reduce the number of agents that must wait a period. Figure 7 is an example where an efficient exchange requires a 6-cycle, but the hospital capacity constraint only allows for a maximum of three simultaneous exchanges. As the next proposition shows, this result is quite general.

INSERT FIGURE 7 HERE

Proposition 4: For any α, a simultaneous exchange of length $k > \alpha$ can be converted to a hybrid exchange where no more than α many simultaneous exchanges occur and at most $\left\lfloor \frac{k}{\alpha} \right\rfloor + 1$ agents must wait a period to receive their kidney.

Proof: Figure 7 captures the intuition for the argument. Consider any cycle of agents $\{a_1, a_2, \ldots, a_k\}$ in a simultaneous exchange. Now consider the following assignment in the replicated economy:

$$f(a_{jt}) = \begin{cases} a_{j+1,t} & j \neq 0 \bmod(\alpha) \\ a_{j+1,t-1} & j = 0 \bmod(\alpha) \\ a_{1,t-1} & j = k \end{cases}$$

This exchange is well defined, no set of simultaneous exchanges involves more than α many agents, and at most $\left\lfloor \frac{k}{\alpha} \right\rfloor + 1$ agents must wait a period to receive their kidney.

Q.E.D.

A natural question to ask is whether any additional gains may be realized if we sometimes require agents to wait longer than a single period. However, it would be misleading to compare the number of exchanges possible with such a sequential exchange to the exchanges possible with only simultaneous exchanges, as we have effectively multiplied the population being matched by a factor related to the waiting time. In order to make a more reasonable comparison, we utilize the concept of a replicated economy and we compare the efficient number of matches in the generalized exchange with the efficient number of simultaneous matches in the replicated economy.

Definition: Given a graph *G* and its corresponding set of agents *X*, for any integer *k,* define the *k*-replicated economy $G^{(k)}$ as follows. The set of agents is $X^{(k)} = \{a_i^j : i \in \{1,...,N\}, j \in \{0,...,k-1\}\}$. Moreover, $e_{a_i^j, a_l^m} \in G^{(k)}$ if and only if $e_{a_i, a_k} \in G$.

We find a positive result. The hybrid sequential exchange does at least as well as simultaneous exchange in a replicated economy. Moreover, there is a natural relationship between a sequential exchange where agents may be required to wait $2k - 1$ periods and a static matching in the k-replicated economy.

Proposition 5: Consider any static exchange in a k-replicated economy $X^{(k)}$. Then there exists a corresponding hybrid sequential exchange in X^{∞}, the unreplicated but repeated economy, where:

1. The average number of trades in each period is $\dfrac{\Delta(X^{(k)})}{k}$; and
2. No agent waits more than $2k - 1$ periods after giving a kidney to receive a kidney.

We give a formal proof below, but the intuition is captured in the following figures. Suppose we have a static exchange in a k-replicated economy. Figure 8 gives an example of a 3-way static exchange in a 3-replicated economy.

INSERT FIGURE 8 HERE

In the repeated economy, we have only one population each period. However, k consecutive economies are isomorphic to a k-replicated economy. Therefore, for any period t, we associate agent a_{it} in the repeated economy with agent $a_i^{t \bmod (k)}$ in the k-replicated economy.

INSERT FIGURE 9 HERE

Now way we can follow the same procedure as in Proposition 1 to transform a static exchange into a sequential exchange. In this generalized procedure, k consecutive populations exchange with the k previous consecutive populations.

INSERT FIGURE 10 HERE

Proof: Consider any k-replicated economy $X^{(k)}$ and any maximal match μ. Let $\mu(X^{(k)})$ denote the set of agents that exchange kidneys. Now consider the unreplicated but repeated economy. For convenience, relabel each X_t as X_{kr+s} where:

$$r = \left\lfloor \frac{t}{k} \right\rfloor,\ 0 \le s \le k-1 \text{ and } t = rk + s$$

Define a hybrid sequential exchange as follows. If $\mu(a_i^m) = a_j^n$, then let $f(a_{i,kr+m}) = a_{j,k(r-1)+n}$. Let $f(x) = \varnothing$ otherwise. Since μ is a well defined exchange, f must be a well defined hybrid exchange. By construction, the number of agents that exchange kidneys among periods $\{rk, rk+1, \ldots, rk+k-1\}$ is equal to the number of agents that exchange kidneys in μ. Therefore, the average number of kidneys matched in any given period is $\frac{\Delta(X^{(k)})}{k}$. Also, note that in this exchange, a donor gives to an agent at most $2k-1$ periods prior to her own period.

Q.E.D.

Section 2 considered exchanges where no two agents in the same period are matched. Here we consider both inter and intra-period exchanges. We find that such exchanges, on average, do no better.

Proposition 6: Consider a steady state exchange in which an agent in period *t* may only donate to an agent in either period *t* or *t* – 1. If Δ many agents are matched each period, then there exists a static exchange in a replicated economy where proportion $\frac{\Delta}{N}$ of the agents receive kidneys.

Proof: Look at any hybrid exchange *f*. Let $f(X_t) \in \wp(X_t)$ denote the set of agents that receive a kidney in any period *t*. As there are infinitely many periods and the cardinality of $\wp(X)$ is finite, by the pigeonhole principle there must exist two periods *i* and $i+j$ such that $f(X_i) = f(X_{i+j})$. Now, consider the *j*-replicated economy $X^{(j)}$ and define a static match as follows:

$$\mu(a_k^l) = \begin{cases} a_m^n & \text{if } f(a_{k,i+l}) = a_{m,i+n}, 1 \le l \le j, 1 \le n \le j \\ a_m^j & \text{if } l = 1, \text{ and } f(a_{k,i+1}) = a_{m,i} \\ \varnothing & \text{if } f(a_{k,i+l}) = \varnothing \end{cases}$$

This is a well-defined static exchange as *f* is well defined and $f(X_i) = f(X_{i+j})$.

Q.E.D.

5. Conclusion

In this short paper, we have explored the implications of relaxing the simultaneous-exchange constraint that has been imposed in all of the previous literature on kidney exchange. While there are evident incentive reasons to require the donor to give up her kidney *no later than* the associated patient receives his transplant, the need is less compelling for the two operations to occur at exactly the same time. If we permit sequential exchanges in which the donor gives up her kidney in one period and the designated recipient receives a donation in a later period, the constraint posed by a limited number of concurrent operating rooms is relaxed and a greater number of beneficial transplants is possible.

For a practical implementation of this market design innovation to be successful, the critical ingredient is to assure donors that this is not a "Ponzi scheme" and to give them confidence that their designated recipients will be served.[5] There are three aspects to the needed confidence:

- Confidence that a compatible donor for the designated recipient will enter the pool with high probability;
- Confidence that this compatible donor will also be willing to participate in a sequential exchange; and
- Confidence that this compatible donor will be matched with the designated recipient.

With a stationary population, all three requirements may be easily achieved with a sequential exchange. In reality, there is never a truly stationary population as both the size and the characteristics of the donor-patient pool change in each period. Therefore, it is reasonable to dichotomize agents into two subpopulations: agents who recur in each period with high probability; and agents who are comparatively rare. With a semi-sequential exchange, the agents that recur with a high enough probability may be accommodated sequentially while exchanges involving comparatively rare agents are done simultaneously.

This also suggests how we can transition from a system of purely simultaneous exchanges to one utilizing both sequential and semi-sequential exchanges. Initially, a very small group of agents Y could be processed sequentially where Y consists of the agent-types that are essentially guaranteed to reoccur in the next month. Over time, Y can be expanded until the benefits from relaxing the hospital capacity constraint no longer exceeds the costs associated with including an agent whose type may not reoccur in the next period.

The variance of the characteristics of the donor-patient pool in each period reduces the number of exchanges that may be handled sequentially. This suggests another advantage of converting regional exchange programs into a national exchange program. Several papers have quantified the increase in the number of simultaneous exchanges that are possible when the population being matched is expanded (see Toulis and Parkes, 2010, and Ashlagi et al. 2012). A separate advantage is that a large, national exchange program reduces the variability of the population in each period.

It is unreasonable to assume that agents will only participate in an exchange program if they are guaranteed to receive a kidney. After all, even in a simultaneous exchange, agents are never guaranteed a successful transplant. Similarly, agents will participate in a sequential exchange so long as the probability they receive a kidney is high enough that

[5] A Ponzi scheme is an investment fraud that involves the payment of purported returns to existing investors from funds contributed by new investors.

the expected benefit from a successful transplant outweighs the cost associated with donation.

All of these confidence issues will be easier to satisfy as a greater flow of donor-patient pairs enter the kidney exchange and as a longer history of trades develops. As this occurs, and as a historical database becomes available, it will be possible to provide donors with reliable, individualized information such as: "With 93% probability, a live donor compatible with your designated recipient will be offered within one month." Effective fallback options can also be developed: for example, if no compatible donor emerges within one month, the patient can be offered the option of jumping to the front of the cadaver queue.[6] Finally, people should have no concerns that the next generation of donor-patient pairs will decline to participate (the usual downfall of Ponzi schemes and asset bubbles) as, unfortunately, the population in need of kidney transplants will not be declining any time soon.

Sequential kidney exchanges come at some cost; while more beneficial trades occur than with simultaneous exchanges, they occur with some amount of added delay. If the added delays are felt to be excessive, policymakers would do best to increase the number of concurrent operating rooms. Our analysis includes consideration of hybrid exchanges (where some exchanges occur simultaneously, and some occur sequentially) and, as the operating-room constraint is eased, it is evident that the optimal solution among hybrid exchanges would exhibit a shift away from sequential exchanges and toward simultaneous exchanges. Still, at any reasonable cost of delay, it seems likely that a social planner would want some fraction of the exchanges to be sequential rather than simultaneous.

At the same time, the reader should recognize that "delay" is not necessarily costly in this context. When the donor and her designated recipient are members of the same household, it may be extremely onerous for both to be recovering from surgery at exactly the same time. For members of the same household, a little sequentiality may be viewed as a good thing.

Sequential kidney exchange holds some promise as an improvement upon the current solution to the market design problem. It does not violate incentive compatibility; nor does it violate the legal constraint against payment of valuable consideration (other than in-kind directed donations) for organ transplants. In short, for kidney exchanges, it may be better first to give and then to receive; rather than always to give and receive simultaneously.

[6] The cadaver queue is the list of patients waiting to receive a cadaver kidney. See Roth et al. (2006) for a detailed description of how a system of paired exchanges might interact with the cadaver queue.

References

Ashlagi, I., Gilchrist, D. S., Roth, A. E., and Rees, M. A. (2010) "Nonsimultaneous Chains and Dominos in Kidney Paired Donation – Revisited," working paper.

Ashlagi, I., Gamarnik, D., Rees, M.A., and Roth, A.E. (2012) "The Need for (long) Chains in Kidney Exchange," working paper.

Montgomery, R.A., Gentry, S. E., Marks, W. H., et al. (2006) "Domino Paired Kidney Donation: A Strategy to Make Best Use of Live Non-Directed Donation," *Lancet* 368: 419-421.

Rees et al. (2009), "A Non-simultaneous Extended Altruistic Donor Chain," *New England Journal of Medicine*, 360;11, March 12, 2009, 1096-1101.

Roth et al. (2006), "Utilizing List Exchange and Undirected Donation through "Chain" Paired Kidney Donations," *American Journal of Transplantation*, 6, 11, November 2006, 2694-2705.

Roth, A. E., Sonmez, T., and Unver, U. (2004) "Kidney Exchange," *Quarterly Journal of Economics* 119(2): 457-488.

Roth, A. E., Sonmez, T., and Unver, U. (2005a) "Pairwise Kidney Exchange," *Journal of Economic Theory* 125: 151-188.

Roth, A. E, Sonmez, T., and Unver, U. (2005b) "A Kidney Exchange Clearinghouse in New England," *American Economic Review Paper and Proceedings* 95(2): 376-380.

Roth, A. E., Sonmez, T., and Unver, U. (2007) "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review* 97(3): 828-851.

Toulis, P. and Parkes, D.C. (2011) "A Random Graph Model of Kidney Exchanges: Optimality and Incentives," *Proc of the 11th ACM Conference on Electronic Commerce*: 232-332.

Figures

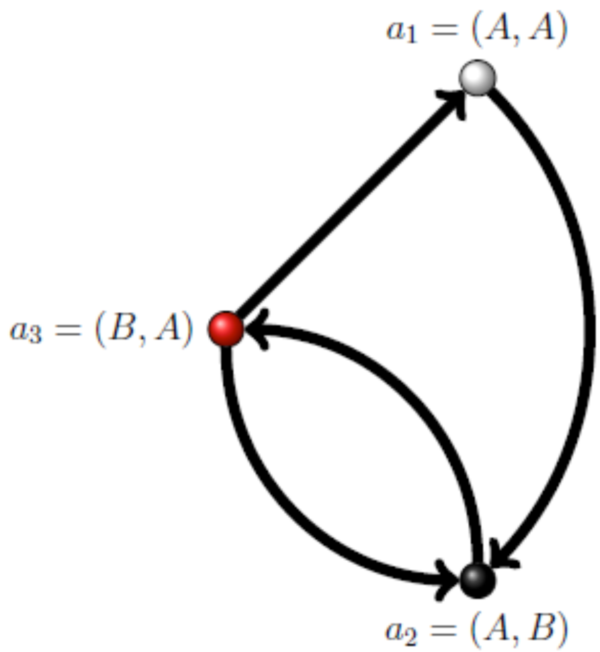


Figure 1 – Each node represents a donor-patient pair. An A-blood-type patient is incompatible with a B-blood-type donor, but compatible with an A-blood-type donor.

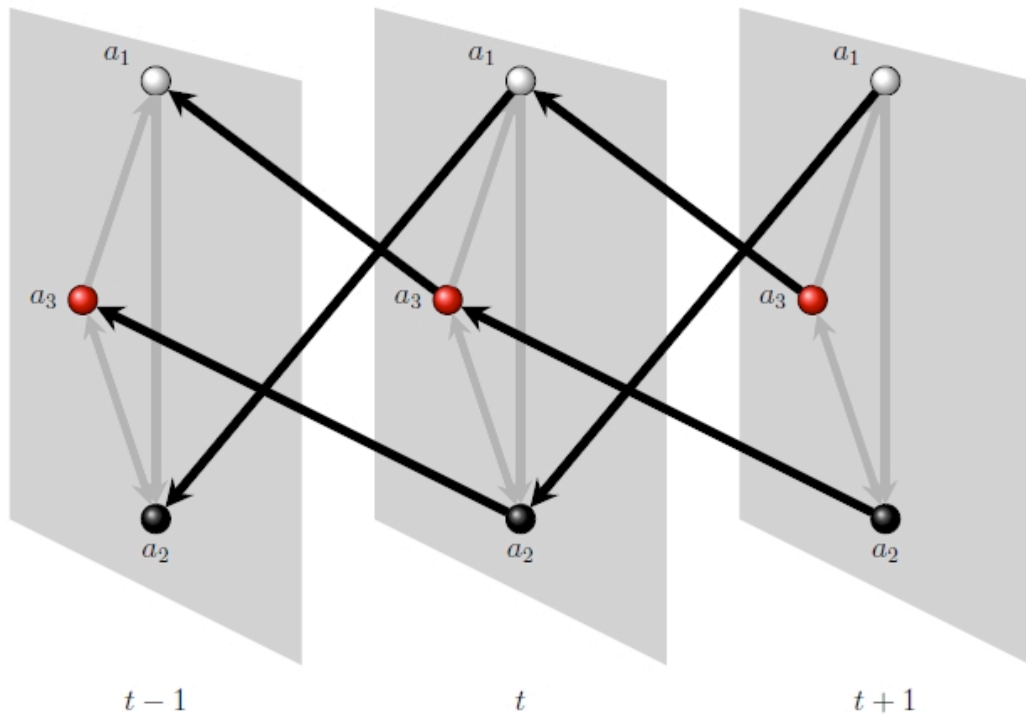


Figure 2 – The incentive and hospital capacity constraints are satisfied if agents donate a kidney before they receive a kidney.

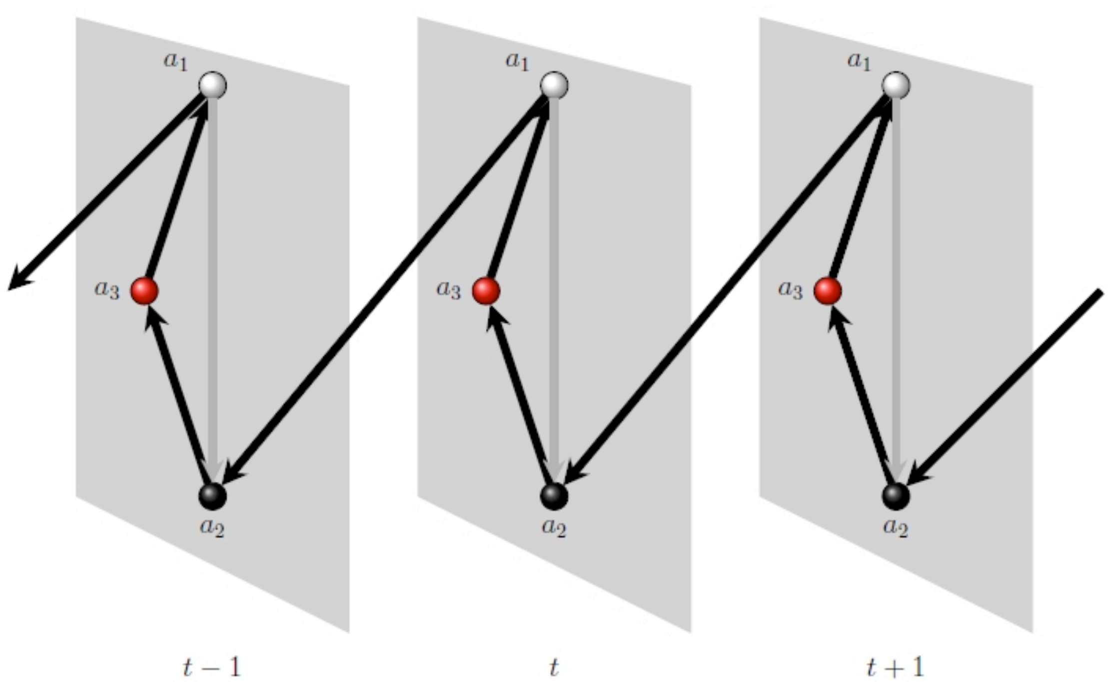


Figure 3 – A hybrid exchange. All possible exchanges can occur while using at most four concurrent operating rooms and requiring only one patient each period to wait to receive a kidney.

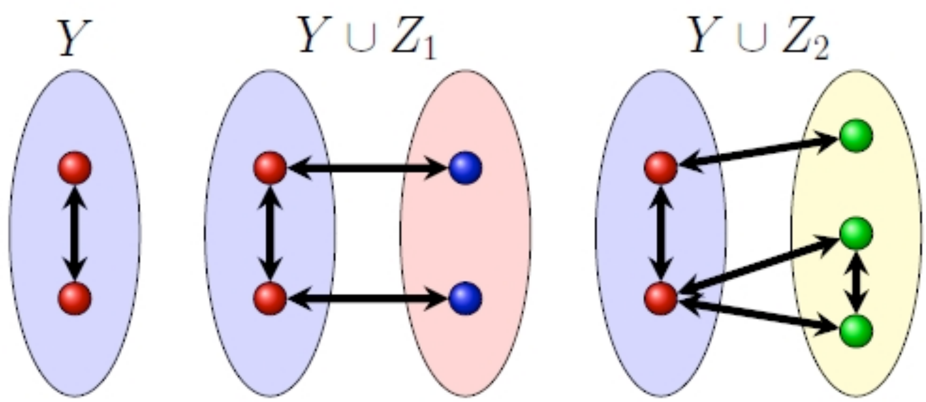


Figure 4 – The populations used in Figure 5.

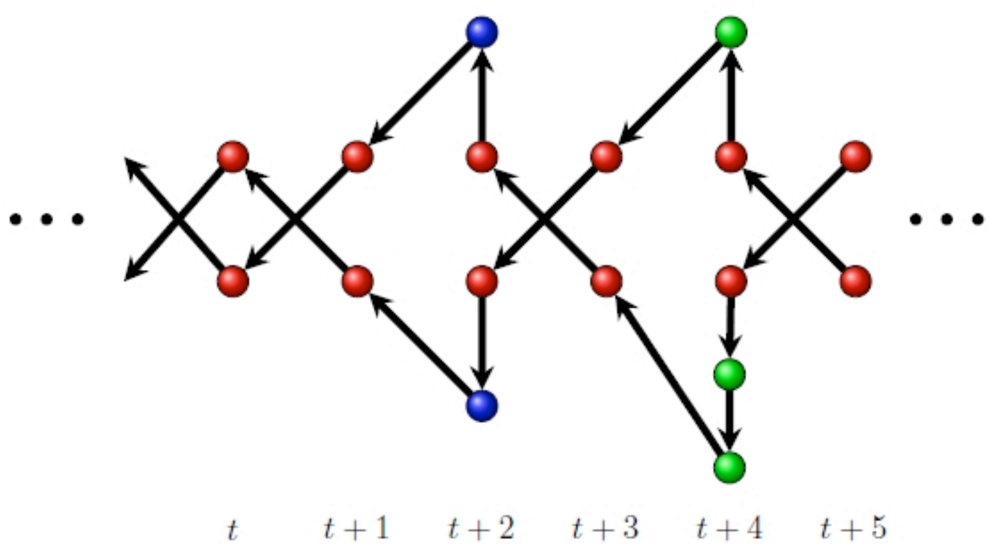
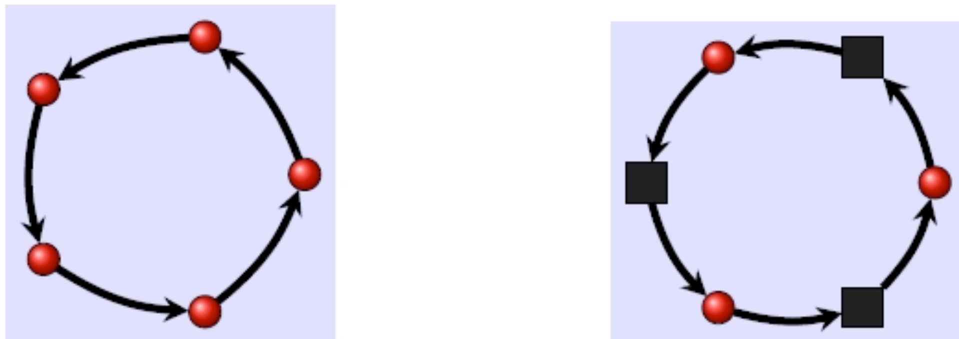


Figure 5 –This example demonstrates how the populations from Figure 4 can be incorporated into a hybrid exchange. Instead of donating to a member of Y in the same period, the agent donates to the corresponding member of Y from one period earlier.



Figures 6(a) and 6(b) – In these two figures, circles represent members of Z_t while squares represent members of Y. The cycle on the left is requires 10 simultaneous operating rooms in both the static and hybrid exchanges. The cycle on the right requires 12 simultaneous rooms in a static exchange but only two simultaneous rooms in a hybrid exchange.

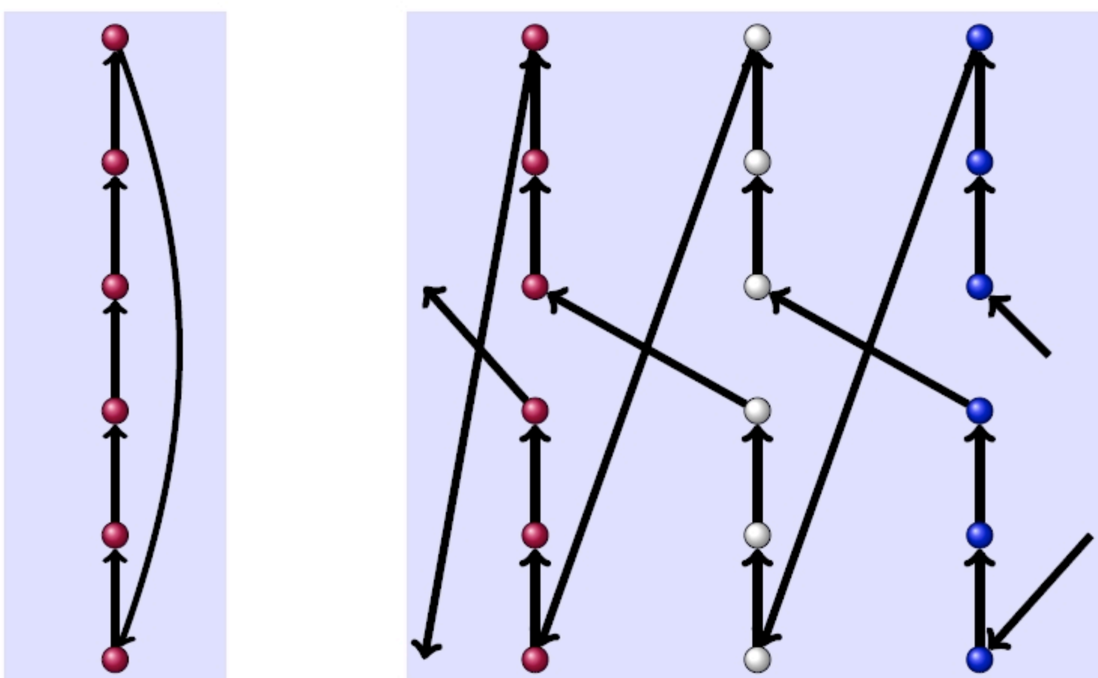


Figure 7 – Suppose the six-cycle of the left panel is required to attain full efficiency, but the hospital capacity constraint allows only three simultaneous transplants. The

hybrid exchange of the right panel attains full efficiency while requiring fewer agents to wait than a sequential exchange.

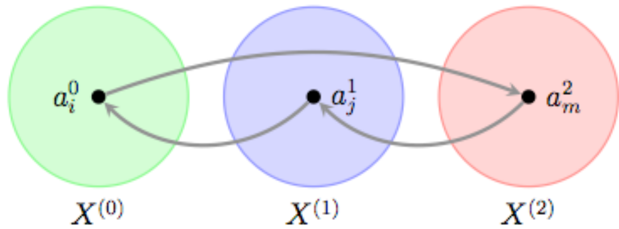


Figure 8 – A three-way exchange in a 3-replicated economy.

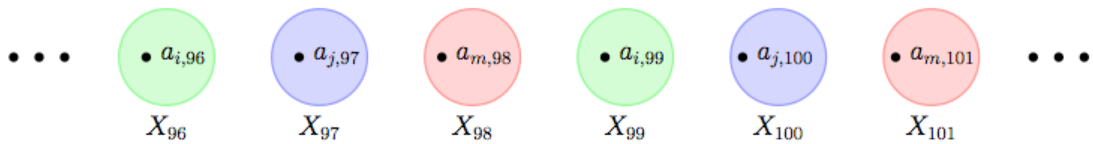


Figure 9 – A 3-replicated economy is isomorphic to three consecutive periods in the repeated economy.

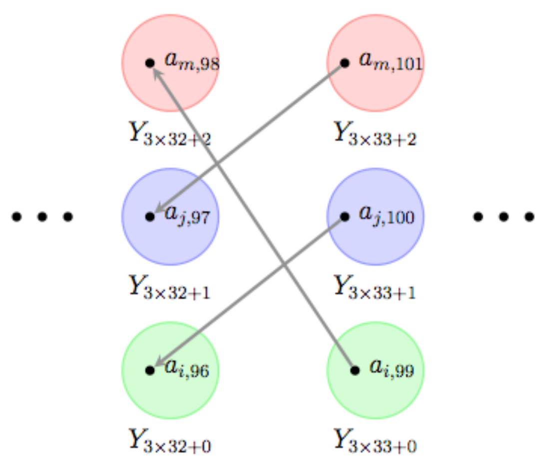


Figure 10 – The same three-way exchange in the static economy recreated as a hybrid exchange in the repeated economy.