Kidney Exchange with Good Samaritan Donors: A Characterization

Tayfun Sönmez
Boston College

M. Utku Ünver
University of Pittsburgh
1 Introduction

- Transplantation is the preferred treatment for the most serious forms of kidney disease.

- More than 60,000 patients on the waitlist for deceased donor kidneys in the U.S., about 15,000 waiting more than 3 years. In 2004 about 3,800 patients died while on the waitlist while only 14,500 patients received a transplant form deceased (about 8,500) or live donors (about 6,000).

- Buying and selling a body part is illegal in many countries in the world including the U.S. Donation is the only source of kidneys in many countries.
Sources of Donation:

1. *Deceased Donors:* In the U.S. and Europe a centralized priority mechanism is used for the allocation of deceased donor kidneys, which are considered national treasure.

2. *Living Donors:* Live donations have been the increasing source of donations in the last decade. Two types:

   (a) *Directed donation:* Generally friends or relatives of a patient specifically want to donate their kidney to their loved ones.

   (b) *Undirected donation:* “Good Samaritans” (GS) who anonymously donate one of their kidneys. Usually GS kidney is treated as a deceased donor kidney and is transplanted to the highest priority patient in the deceased donor waiting list.
2 Donations and Live Donor Exchanges

• There are two tests that a donor should pass before she is deemed compatible with the patient:

  – Blood compatibility test: O type kidneys compatible with all patients; A type kidneys compatible with A and AB type patients; B type kidneys compatible with B and AB type patients; AB type kidneys compatible with AB type patients.

  – Tissue compatibility test (crossmatch test): HLA proteins play two roles (1) determine tissue rejection or compatibility and (2) how close the tissue match is.

• If either test fails, the patient remains on the deceased donor waiting list. If the donor is a directed donor, she goes home unutilized.

• Medical community has already come up with a way of utilizing these “unused” directed donors.
• A paired exchange involves two incompatible patient-donor couples such that the patient in each couple feasibly receives a transplant from the donor in the other couple. This pair of patients exchange donated kidneys.

![Diagram of a paired exchange]

Donor 1  Patient 1

Patient 2  Donor 2

• Larger exchanges can also be utilized (Two 3-way exchanges have been utilized in Johns Hopkins University Transplant Center)
3 Kidney Exchange Developments

- Kidney exchange mechanisms were proposed by Roth, Sönmez and Ünver *QJE* (2004), *JET* (2005) (also see *AER-P&$P* (2005), NBER wp (2005))

- New England Kidney Exchange (NEPKE) was established by the proposals of by Alvin Roth, Drs. Francis Delmonico Susan Saidman, and us in 2004

- A national exchange program is being proposed.
4 Integrating GS Donations with Paired Exchanges

In May 2005, surgeons at Johns Hopkins performed an exchange between a *Good Samaritan donor*, two incompatible patient-donor pairs, and a patient on the deceased-donor priority list.

- In the recent exchange at Johns Hopkins,
  - the kidney from the GS-donor is transplanted to the patient of the first incompatible pair,
  - the kidney from the first incompatible pair is transplanted to the patient of the second incompatible pair, and
  - the kidney from the second incompatible pair is transplanted to the highest priority patient on the deceased-donor priority list.

- What are plausible mechanisms to integrate GS donations with paired exchanges?
5 Other Related Literature

- Shapley and Scarf *JME* (1974) - housing market

- Roth *EL* (1982) - strategy-proofness of core as a mechanism in housing markets

- Ma *IJGT* (1994) - characterization of core in housing markets

- Svensson *SCW* (1999) - characterization of serial dictatorships in house allocation

- Abdulkadiroğlu and Sönmez *JET* (1999) - house allocation problem with existing tenants

- Ergin *JME* (2000) - another characterization of serial dictatorships in house allocation
6 The Model

• $\mathcal{I}$: a finite set of patients

• $\mathcal{D}$: a finite set of donors such that $|\mathcal{D}| \geq |\mathcal{I}|$.

• Each patient $i \in \mathcal{I}$ has a paired-donor $d_i \in \mathcal{D}$ and has strict preferences $P_i$ on all donors in $\mathcal{D}$.

  – Let $R_i$ denote the weak preference relation induced by $R_i$ and

  – For any $D \subset \mathcal{D}$, let $\mathcal{R}(D)$ denote the set of all strict preferences over $D$. 
A kidney exchange problem with good samaritan donors, or simply a problem, is a triple $\langle I, D, R \rangle$ where:

- $I \subseteq \mathcal{I}$ is any set of patients,
- $D \subseteq \mathcal{D}$ is any set of donors such that $d_i \in D$ for any $i \in I$, and,
- $R = (R_i)_{i \in I} \in [\mathcal{R}(D)]^{|I|}$ is a preference profile.

Given a problem $\langle I, D, R \rangle$, the set of “unattached” donors $D \setminus \{d_i\}_{i \in I}$ is referred as Good Samaritan donors (or in short GS-donors).

- Paired-donor $d_j$ of a patient $j$ is formally a GS-donor in a problem $\langle I, D, R \rangle$ if $d_j \in D$ although $j \notin I$. 
• Given $I \subseteq \mathcal{I}$ and $D \subseteq \mathcal{D}$, a matching is a mapping $\mu : I \to D$ such that

$$\forall i, j \in I, \ i \neq j \Rightarrow \mu(i) \neq \mu(j).$$

• We denote a problem $\langle I, D, R \rangle$ simply by its preference profile $R$

• A mechanism is a systematic procedure that selects a matching for each problem.
7 Axioms

7.1 Individual Rationality, Pareto Efficiency and Strategy Proofness

Fixed population axioms:

- A matching is \textit{individually rational} if no patient is assigned a donor worse than her paired-donor.
  - A mechanism is \textit{individually rational} if it always selects an individually rational matching.

- A matching is \textit{Pareto efficient} if there is no other matching that makes every patient weakly better off and some patient strictly better off.
  - A mechanism is \textit{Pareto efficient} if it always selects a Pareto efficient matching.
• A mechanism is *strategy-proof* if no patient can ever benefit by misrepresenting her preferences.

7.2 Weak Neutrality and Consistency

Variable population axioms:

• A mechanism is *weakly neutral* if labeling of GS-donors has no affect on the outcome of the mechanism.
Let for any \( i \in I \), \( R_i \in \mathcal{R}(D) \) for \( D \subset D \) and \( I \subset D \). For any \( J \subset I \) and \( C \subset D \), let \( R_{j}^{C} = (R_{i}^{C})_{i \in J} \) be the restriction of profile \( R \) to patients in \( J \) and donors in \( C \).

We refer \( \langle J, C, R_{j}^{C} \rangle \) as the restriction of problem \( \langle I, D, R \rangle \) to patients in \( J \) and donors in \( C \). The triple \( \langle J, C, R_{j}^{C} \rangle \) itself is a well-defined reduced problem if whenever a patient is in \( J \) then her paired-donor is in \( C \).

Given a problem \( \langle I, D, R \rangle \), the removal of a set of patients \( J \subset I \) together with their assignments \( \phi[R](J) \) under \( \phi \) and a set of unassigned donors \( C \subset D \) under \( \phi \) results in a well-defined reduced problem

\[
\langle I \setminus J, D \setminus (\phi[R](J) \cup C), R_{-j}^{-\phi[R](J)\cup C} \rangle
\]

if

\[
(\phi[R](J) \cup C) \cap \{d_i\}_{i \in I \setminus J} = \emptyset.
\]
• A mechanism is *consistent* if the removal of
  
  – a set of patients,

  – their assignments, and

  – some unassigned donors

  does not affect the assignments of remaining patients provided that the removal results in a well-defined reduced problem.

• Once a mechanism finds a matching, actual operations can be done months apart in different exchanges. Moreover, some unassigned donors (who are either GS-donors or donors of patients who already received a transplant) may be assigned to the deceased donor waiting list in the mean time. Therefore, *consistency* of the mechanism ensures that once the operations in an exchange are done and some unassigned donors become unavailable, there is no need to *renege* the determined matching, since the mechanism will determine the same matching in the reduced problem.
8 You Request My Donor-I Get Your Turn Mechanism

- Abdulkadiroğlu and Sönmez \textit{JET} (1999) introduced in the context of \textit{house allocation with existing tenants} (see also Chen and Sönmez \textit{JET} (2006) and Sönmez and Ünver \textit{GEB} (2005)

- A (priority) ordering \( f : f(1) \) indicates the patient with the highest priority in \( I \), \( f(2) \) indicates the patient with the second highest priority in \( I \), and so on.

- Given a set of patients \( J \in I \), the restriction of \( f \) to \( J \) is an ordering \( f_J \) of the patients in \( J \) which orders them as they are ordered in \( f \).

- Each ordering \( f \in \mathcal{F} \) defines a YRMD-IGYT mechanism.
– For any problem \( \langle I, D, R \rangle \), let \( \psi^f[R] \) denote the outcome of YRMD-IGYT mechanism induced by ordering \( f \).

– Let \( \psi^f[R_J^C] \) denote the outcome of the YRMD-IGYT mechanism induced by ordering \( f_J \) for problem \( \langle J, C, R_J^C \rangle \).
For any problem \( \langle I, D, R \rangle \), matching \( \psi^f[R] \) is obtained with the following YRMD-IGYT algorithm in several rounds.

**Round 1(a):** Construct a graph in which each patient and each donor is a node. In this graph:

- each patient “points to” her top choice donor (i.e. there is a directed link from each patient to her top choice donor),

- each paired-donor \( d_i \in D \) points to her paired-patient \( i \) in case \( i \in I \), and to the highest priority patient in \( I \) otherwise,

- and each GS-donor points to the patient with the highest priority in \( I \).
Define: a cycle is an ordered list \((c_1, j_1, \ldots, c_k, j_k)\) of donors and patients where donor \(c_1\) points to patient \(j_1\), patient \(j_1\) points to donor \(c_2\), donor \(c_2\) points to patient \(j_2\), \ldots, donor \(c_k\) points to patient \(j_k\), and patient \(j_k\) points to donor \(c_1\).
Since there is a finite number of patients and donors, there is at least one cycle. If there is no cycle without a GS-donor then skip to Round 1(b). Otherwise consider each cycle without a GS-donor. (Observe that if there is more than one such cycle, they do not intersect.) Assign each patient in such a cycle the donor she points to and remove each such cycle from the graph. Construct a new graph with the remaining patients and donors such that

- each remaining patient points to her first choice among the remaining donors,

- each remaining paired-donor $d_i \in D$ points to her paired-patient $i$ in case her paired patient $i$ remains in the problem, and to the highest priority remaining patient otherwise,

- and each GS-donor points to the highest priority remaining patient.
There is a cycle. If there is no cycle without a GS-donor then skip to Round 1(b); otherwise carry out the implied exchange in each such cycle and proceed similarly until either no patient is left or there exists no cycle without a GS-donor.

**Round 1(b):** There is a unique cycle in the graph, and it includes both the highest priority patient among remaining patients and a GS-donor. Assign each patient in such a cycle the donor she points to and remove each such cycle from the graph. Proceed with Round 2.
In general, at

*Round $t(a)$*: Construct a new graph with the remaining patients and donors such that

- each remaining patient points to her first choice among the remaining donors,

- each remaining paired-donor $d_i \in D$ points to her paired-patient $i$ in case her paired patient $i$ remains in the problem, and to the highest priority remaining patient otherwise,

- and each remaining GS-donor points to the highest priority remaining patient.
There is a cycle. If the only remaining cycle includes either a GS-donor or a paired-donor whose paired-patient has left, then skip to Round $t(b)$; otherwise carry out the implied exchange in each such cycle and proceed similarly until either no patient is left or the only remaining cycle includes either a GS-donor or a paired-donor whose paired-patient has left.

Round $t(b)$: There is a unique cycle in the graph, and it includes the highest priority patient among remaining patients and either a GS-donor or a paired-donor whose paired-patient has left. Assign each patient in such a cycle the donor she points to and remove each such cycle from the graph. Proceed with Round $t+1$.

The algorithm terminates when there is no patient left in the graph.
9 Characterization of the YRMD-IGYT Mechanisms

Our main result is a characterization of the YRMD-IGYT mechanism:

Theorem 1: A mechanism is Pareto efficient, individually rational, strategy-proof, weakly neutral, and consistent if and only if it is a YRMD-IGYT mechanism.
We present our main result through two propositions:

Proposition 1: For any ordering $f \in \mathcal{F}$, the induced YRMD-IGYT mechanism $\psi^f$ is Pareto efficient, individually rational, strategy-proof, weakly neutral and consistent.

Proposition 2: Let $\phi$ be a Pareto efficient, individually rational, strategy-proof, weakly neutral, and consistent mechanism. Then $\phi = \psi^f$ for some $f \in \mathcal{F}$. 
Sketch of Proof of Proposition 2:

• Construct $f$ as follows: Let $d_{gs} \in D$ be a GS-donor.

  - Construct $R^1$ as follows
    \[
    \begin{array}{cccc}
    R^1_1 & R^1_2 & \cdots & R^1_n \\
    d_{gs} & d_{gs} & \cdots & d_{gs} \\
    d_1 & d_2 & \cdots & d_n \\
    \vdots & \vdots & \ddots & \vdots
    \end{array}
    \]

    Pareto efficiency of $\phi \Rightarrow$ for some $i$, $\phi[R^1](i) = d_{gs}$. Let $f(1) = i$.

  - Construct $R^2$ as follows:
    \[
    \begin{array}{cccc}
    R^2_{f(1)} & R^2_1 & R^2_2 & \cdots & R^2_n \\
    d_{f(1)} & d_{gs} & d_{gs} & \cdots & d_{gs} \\
    \vdots & d_1 & d_2 & \cdots & d_n \\
    \vdots & \vdots & \ddots & \vdots
    \end{array}
    \]

    Individual rationality of $\phi \Rightarrow \phi[R^2](f(1)) = d_{gs}$.

    Pareto efficiency of $\phi \Rightarrow$ for some $i \neq f(1)$, $\phi[R^2](i) = d_{gs}$. Let $f(2) = i$. 
– similarly construct $R^3$ by changing $f(2)$’s preferences so that only $d_{f(2)}$ is acceptable. We continue similarly... This gives a unique ordering $f$. 
• Let $R \in \mathcal{R}(D)^{|I|}$ for $I \subseteq \mathcal{I}$ and $D \subseteq D$. We will prove that $\psi^f[R] = \phi[R]$.

• To prove this result we construct an interim preference profile $R'$ using $R$. Use YRMD-IGYT algorithm to construct $\psi^f[R]$.

  – Let $A^t$ be the patients removed in round $t(a)$ for any $t$.

  – Let $B^t$ be the patients removed in round $t(b)$ for any $t$.

• $R'_i$ is constructed in two different ways for a patient $i \in I$ depending on how she leaves the algorithm. Suppose she leaves the algorithm in round $t$. Two cases are possible: She leaves

  1. (i) in round $t(a)$ or (ii) in round $t(b)$ and she is not the highest priority patient in this cycle.

  2. in round $t(b)$ and she is the highest priority patient in this cycle
Figure 1: Construction of Preference $R'_i$ for Case 1

Figure 2: Construction of Preference $R'_i$ for Case 2 when $\psi^f[R](B^t) = \{\psi^f[R](i), c, c'\}$
By construction, $\psi_f [R'] = \psi_f [R]$. We will prove four claims that will facilitate the proof of Proposition 2.

We consider the patients in $A^1$ in the first two claims.

**Claim 1:** For any $\hat{R}_{-A^1} \in R^{|I\backslash A^1|}$ and $i \in A^1$, we have $\phi \left[ R'_{A^1}, \hat{R}_{-A^1} \right] (i) = \psi_f [R] (i)$.

The proof uses *individual rationality* and *Pareto efficiency* of $\phi$.

**Claim 2:** For any $\hat{R}_{-A^1} \in R^{|I\backslash A^1|}$, and any $i \in A^1$, we have $\phi \left[ R_{A^1}, \hat{R}_{-A^1} \right] (i) = \psi_f [R] (i)$.

The proof uses Claim 1, *strategy-proofness* in addition to *individual rationality* and *Pareto efficiency* of $\phi$. 
We consider the patients in $B^1$ in the next two claims.

**Claim 3:** $\phi \left[ R'_{B^1}, R_{-B^1} \right] (i) = \psi^f [R] (i)$ for all $i \in B^1$.

The proof uses Claim 2, consistency and weak neutrality in addition to strategy-proofness, individual rationality and Pareto efficiency of $\phi$.

**Claim 4:** $\phi [R] (i) = \psi^f [R] (i)$ for all $i \in B^1$.

The proof uses Claims 2 and 3, strategy-proofness, consistency, and individual rationality of $\phi$. 
For the rest of the patients, we use \textit{consistency} of $\phi$ and the above 4 claims.

By Claim 2 and Claim 4,

$$\phi[R](i) = \psi^f[R](i) \quad \text{for all } i \in A^1 \cup B^1.$$  

By invoking \textit{consistency}, we can remove patients in $A^1 \cup B^1$ and their assigned donors and we can similarly prove

$$\phi[R](i) = \psi^f[R](i) \quad \text{for all } i \in A^2 \cup B^2.$$  

Iteratively we continue to prove that

$$\phi[R] = \psi^f[R].$$
10 Independence of the Axioms

The following examples establish the independence of the axioms.

Example 1: Individually rational, strategy-proof, weakly neutral and consistent but not Pareto efficient mechanism: Let mechanism \( \phi \) assign each patient \( i \in I \) her paired-donor \( d_i \) for each problem \( \langle I, D, R \rangle \).

Example 2: Pareto efficient, strategy-proof, weakly neutral and consistent but not individually rational mechanism: Fix an ordering \( f \in F \) and let mechanism \( \phi \) be the serial dictatorship induced by \( f \).
Example 3: Pareto efficient, individually rational, weakly neutral and consistent but not strategy-proof mechanism: Fix an ordering $f \in \mathcal{F}$. Let $g \in \mathcal{F}$ be constructed from $f$ by demoting patient $f(1)$ to the very end of the ordering. For any problem $\langle I, D, R \rangle$, let

$$\phi[R] = \begin{cases} 
\psi^g[R] & \text{if } dR_id_f(1) \text{ for all } i \in I \text{ and } d \in D, \\
\psi^f[R] & \text{if otherwise.}
\end{cases}$$
Example 4: Pareto efficient, individually rational, strategy-proof, and consistent but not weakly neutral mechanism:

Let $\mathcal{I}, \mathcal{D}$ be such that $|\mathcal{I}| \geq 2$ and $|\mathcal{D}| \geq |\mathcal{I}| + 2$. Let $i_1, i_2 \in \mathcal{I}$ and $d^* \in \mathcal{D} \setminus \{d_i\}_{i \in \mathcal{I}}$. Let $f, g \in \mathcal{F}$ be such that $f(1) = g(2) = i_1$, $f(2) = g(1) = i_2$ and $f(i) = g(i)$ for all $i \in \mathcal{I} \setminus \{i_1, i_2\}$. For any problem $\langle \mathcal{I}, \mathcal{D}, R \rangle$, let

$$
\phi[R] = \begin{cases} 
\psi_f[R] & \text{if } i_1 \in \mathcal{I}, d^* \in \mathcal{D} \text{ and } d^* R_i d \text{ for all } d \in \mathcal{D} \setminus \{d_i\}_{i \in \mathcal{I}} \\
\psi_g[R] & \text{if otherwise.}
\end{cases}
$$
Example 5: Pareto efficient, individually rational, strategy-proof, and weakly neutral but not consistent mechanism:

Let $f, g \in \mathcal{F}$ be such that $f \neq g$. For any problem $\langle I, D, R \rangle$, let

$$\phi[R] = \begin{cases} 
\psi^f[R] & \text{if there are odd number of GS-donors}, \\
\psi^g[R] & \text{if there are even number of GS-donors}.
\end{cases}$$
11 Conclusions

- The result can be generalized to a setting in which the deceased donor waiting patients (without any paired donors) are also explicitly modeled. (A similar domain with house allocation existing tenants problem).

- New England Program for Kidney Exchange (NEPKE) has started to integrate GS donations with paired exchanges.