Welfare Consequences of Transplant Organ Allocation Policies*

Tayfun Sönmez† M. Utku Ünver‡

Boston College *Boston College*

An Outline Prepared for the NBER Market Design Working Group Meeting
Not to be distributed or posted online

1 Introduction

In order to analyze welfare and distributional consequences of different allocation policies for organ transplantation, we introduce a dynamic continuous arrival model of patients and donors, which is the first comprehensive model that incorporates various practices in transplantation, such as deceased donor allocation (for all organs), live donation (for kidneys and livers), and live donor exchange (for kidneys and also livers). The model not only gives us an explanation for some of the observed patterns in the data, but also helps us to analyze and quantify the welfare and distributional consequences of new allocation policies.

In particular, we propose a new exchange policy for live donors, which can substantially increase the number of pairs that can be matched through exchange while reducing inequality among blood types, we characterize the potential gains from this policy.

Currently compatible pairs generally do not participate in exchange, as the patient of the pair directly receives an organ from his donor. Only incompatible pairs participate in exchange. Incompatible pairs are either (a) blood-type incompatible (such as with an O blood-type patient and an

*M. Utku Ünver acknowledges the research support of NSF.

†Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467; sonmezt@bc.edu; www2.bc.edu/~ sonmezt

‡Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467; unver@bc.edu; www2.bc.edu/~ unver

A blood-type donor) or (b) blood-type compatible but tissue-type incompatible (such as the reciprocal of the above pair, with an A blood-type patient and O blood-type donor who is tissue-type incompatible with her patient). Because of this asymmetry, blood-type incompatible pairs are substantially more in number than blood-type compatible pairs participating in exchange. Moreover, for a blood-type incompatible pair to benefit from exchange with the exception of pairs with A and B blood types for the donor and the patient, a blood-type compatible pair is needed. However, the asymmetry in participation puts blood-type incompatible pairs at a high disadvantage and as a result not all pairs can benefit from exchange and the ones who can benefit have wait for their reciprocal blood-type compatible pairs to arrive at the pool. On the other hand, if compatible pairs can also participate in exchange, then the participation asymmetry will disappear, and exchange will benefit more than 90% of the pairs (cf. Roth, Sönmez, and Ünver (2005) and also Sönmez and Ünver (2010)).

However, it is not possible to force compatible pairs to participate in exchange. We propose to incentivize participation by linking deceased-donor wait list with the exchange pool. It is a common practice to give priority to live donors on the deceased-donor wait list in case they themselves get sick and need an organ transplant in the future. We propose giving similar incentives to the patients of compatible donors who give up their own compatible donor's organ in exchange for another pair's compatible organ. In this manner, the patient of a compatible donor receives a "guarantee" not to wait in the deceased-donor wait list by getting a "priority" in case the organ he receives in exchange fails in the future.

Another benefit of this policy can be seen in creating unified national programs for exchange. One of the biggest hurdles that needed to be overcome in kidney exchange in the US is unification of various kidney exchange programs. It is well established that running a single large program that encompasses all programs benefits more patients than running separate programs (cf. Roth, Sönmez, and Ünver (2004) and Roth, Sönmez, and Ünver (2007)). On the other hand, linking deceased-donor wait list with the live-donor exchange queue can only be done through the national kidney exchange program governed by the federal contractor, United Network for Organ Sharing (UNOS), which also directs the deceased-donor allocation program in the US. We show that in an environment which has multiple exchange programs, if compatible pairs participate exchange, they would participate through the national program of UNOS, which has the jurisdiction over the deceased-donor wait list, and in turn, this will attract other pairs to the national program rather than other programs. Hence, our proposed policy has the potential to unify various exchange programs to create a large exchange platform to exploit all benefits from exchange for the society.

1.1 Other Findings

We start our analysis by incorporating deceased-donor donation to the model, to predict the steady-state welfare consequences for different blood-type patients. Two types of deceased donation policies play an important role for many organs. The first commonly adopted policy is "within-blood-type" allocation where a patient can only receive a transplant from a deceased donor with the same blood type. The second policy is "compatible" allocation where a patient can receive a transplant from any compatible blood-type of deceased-donors. Both policies are governed through a priority allocation scheme which gives the greatest weight to waiting time in the queue. We model the priority allocation rule through a "first-come first-servedd" queue. Compatible allocation policy leads to a "pooling effect" by equalizing the waiting time of different blood types whose donors donate / patient receive to / from this group of patients / donors. On the other hand, "within-blood-type" allocation policy leads to separation of waiting times for different blood types with respect to the patient / donor inflow ratios of that blood type.[1]

Then, we consider live donation. Some patients have paired donors who would like to donate an organ, such as a kidney or part of the liver, to them. If they have blood- and tissue-type compatibility they donate to their patients and otherwise they are not utilized. Possibility of live donation helps unambiguously all patients, those with donors and without donors. We characterize the gains from live donation in our model. Patients without live paired donors benefit as patients with compatible live paired donors drop from competition for deceased donors. Among different blood types, O blood-type deceased donor list benefits the least while AB list benefits the most. O patients with a live donor have a lower chance to have a compatible live donor with respect to other types, as they can only receive from O donors. AB patients can receive from all types, and have the highest rate of compatible donors. On the other hand, A patients are more better off than B. As A is a more common blood type in the population, and hence for the paired donors of the patients with respect to B, and hence,a higher fraction of A blood-type patients benefit from live donation with respect to B. We quantify the amount of change in waiting time for a deceased donor in lists for different blood types when compatible live donation is feasible.

Next, we consider live-donor exchanges among incompatible pairs, for organs such as kidneys and livers. We characterize the welfare consequences of live-donor exchange on different blood type

[1]For example, for minorities where B blood type could be a dominant blood type unlike the majority of the population, deceased donation rates do not differ from the rest of the population, yet people are more prone to need transplant due to life-style choices are other factors. Hence, compatible and within-blood-type allocation policies are expected to lead to substantially different waiting times for B patients, who can receive organs from O deceased donors besides B deceased donors.

patients. A live-donor exchange involves the swap of paired live donor organs of two pairs when the donors are incompatible with their own patients but compatible with the patient of the other pair. This causes patients with blood-type compatible donors to be matched immediately either through direct donation (if compatible) or through exchange (if compatible). On the other hand the patients who have blood-type incompatible donors need to wait in the pool, and subject to survival, they get matched either through exchange with a live donor or with a deceased donor depending on the population characteristics.

2 A Dynamic Model of Transplant Patients

We consider a comprehensive dynamic organ transplantation model (for organs such as heart, kidney, liver, and pancreas) to which the deceased-donor waiting list, live donation for kidneys and livers, and live-donor kidney and liver exchange can be incorporated. We consider a continuum flow model in analysis where the number of patients are in Lebesgue measures at a steady-state.

Consider patients who need a particular organ transplant. Each **patient** is represented by his blood type $X \in \mathcal{T} = \{A, B, AB, O\}$. Suppose p_X refers to the probability of having the probability of X blood type in the population distribution. We assume that there is an inflow π_X of blood-type X people getting sick per unit time. Suppose that in the population of new patients this expected life with the disease is distributed with a continuous distribution function $F(\cdot)$ on the interval $[0, T]$.[2] Thus, among the inflow of π_X measure of blood-type X patients at a given time, the measure of patients who are alive after t years on is given by $\pi_X[1 - F(t)]$. At the steady state, when transplantation option is not present, the total measure of blood-type X patients is $\int_0^T \pi_X[1 - F(t)]dt$.[3] (See Figure 2.)

[2]This expectancy is different for different organs due to disease progression and techniques that can be used to substitute for the deficiency in the body because of the failing organ. For example, kidney patients, who can live on dialysis, have in general longer survival times.

[3]Although we assume that inflow of patients is constant over time, we could easily make it a function of time as well. For example, population growth is a reason for increase of inflow. Increase in longevity is another reason, which not only affects π_X but also F, as older people have a higher tendency to need organ transplantation. These can be incorporated in our model easily. In that case a steady state does not exist. However, we can carry all of our analysis in this paper and draw similar results in that model as a function of time. For simplicity and transparency of our analysis, we will use a model with constant inflows.

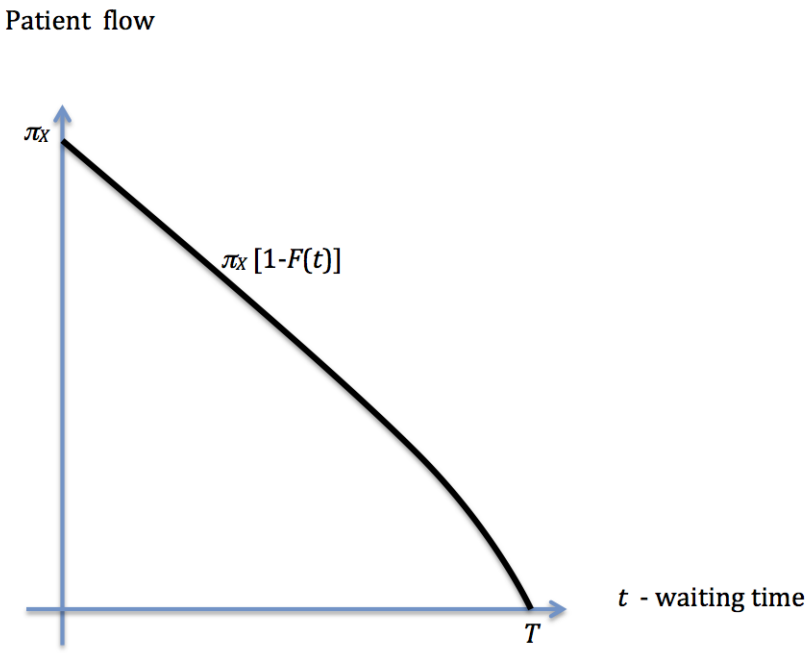


Figure 1: Steady-state blood-type X patient distribution over time without organ transplantation

3 Donation and Deceased-Donor Waiting List

The best remedy for organ failure is transplantation. A donor should be both blood-type and tissue-type compatible with a patient, before her organ(s) can be transplanted. O blood-type donors are blood-type compatible with all blood-type patients. A donors are blood-type compatible with A and AB patients and B donors are blood-type compatible with B and AB patients. On the other hand, AB donors are only blood-type compatible with AB patients. Once a donor is deemed compatible with a patient, she also has to be tissue-type compatible with the patient. Tissue-type compatibility requires that the patient's body does not form antibodies against a donor's DNA. The tissue rejection probability is $\chi < 1$ for each patient.

A common source of donation across organs is deceased donors. The deceased-donor wait list is governed by a central entity. For example in the US, for all organ types, United Network for Organ Sharing (UNOS) is the federal contractor, which is in charge of the wait list.

We denote the inflow of the X blood-type deceased donors, as $\delta_X < \pi_X$ per unit time. Among blood types, the ratio δ_X/π_X need not be constant even if donors are reserved exclusively for the same blood-type patients. For example, it is well known that among minority communities, organ failure is

more prominent than the Caucasian population while deceased donation rates are not that different.[4] As blood-type distribution of minorities are different from the Caucasian population (especially B blood type is observed in much higher frequency among Hispanics), the ratio δ_X/π_X is not constant among blood types even if donors are reserved for the same blood-type patients.

3.1 The Wait List Matching Protocols

The deceased donors are allocated through the points system of UNOS, which is a priority mechanism. When a deceased donor arrives, the point total for each compatible patient for the donor is determined. The organ is offered to the patient with the highest point total. If it is rejected by the patient or his doctor for any reason, then the organ is offered to the next patient, so on. In general, different point schemes are used for different organs. Deceased donor allocation policies usually differ across organs and across geographic transplant regions, although a centralized mechanism is used in allocation. For example for kidneys, strict "within-blood-type" allocation rules are applied. That is, kidneys of blood-type X are only offered to blood-type X patients.[5] On the other hand, livers are offered to any "compatible" patient under different by-laws. We inspect the welfare and dynamic consequences of these two policies separately.

Given a fixed blood-type allocation rule, waiting time of a patient is usually the biggest contributor to the points of a patient in deceased-donor allocation. Therefore, it is convenient to model deceased-donor allocation using "first-come first-served" queues for both "within-blood-type" and "compatible" blood-type allocation schemes.

We analyze these two first-come first-served matching protocols. We state the following lemma, which will help us model the steady state of the wait list.[6]

Lemma 1 (first-come first-served Wait List Matching Protocols) *Suppose that there is an ordered ω measure of blood-type X patients available in the queue and a $\sigma \leq \omega$ measure of donors who are blood-type compatible with X arrive; we match them with the first-come first-served matching protocol. Then, if $\sigma = \omega$, only a finite number of donors (and hence 0 measure of donors) are almost surely unmatched, and if $\sigma < \omega$, no donors are almost surely unmatched under this policy.*

Proof. We prove it by contradiction: If $\sigma = \omega$ then suppose an infinite or uncountable number of donors are unmatched, and if $\sigma < \omega$ then suppose a donor is unmatched with a positive probability

[4]For example see the US Department of Health and Human Services - The Office of Minority Health web page for organ donation http://minorityhealth.hhs.gov/templates/content.aspx?ID=3123As

[5]In the highly unlikely event that no X blood-type patient is available, then the organ is offered to any "compatible" patient.

[6]This is in spirit similar to the Erdös and Rényi (1960) random graph convergence result.

under the first-come first-served policy. Then, in either case, an infinite or uncountable number of patients are unmatched as well. But then, take a donor who is unmatched, then there exists almost surely a compatible unmatched patient, as the probability of finding no tissue-type compatible patient is $\lim_{n\to\infty} \chi^n = 0$. ■

3.2 Steady State of The Wait List

We are ready to characterize the steady state of the wait list under the two first-come first-served matching protocols .

3.2.1 "Within-Blood-Type" Deceased Donation

Consider any blood type X. In the steady state, as $\delta_X < \pi_X$, there will always be a positive measure of X blood type patients available in the wait list. Moreover, as first-come first-served protocol is used, this δ_X measure will be transplanted to the longest X-blood-type standers in the wait list who are still alive. Thus, by Lemma 1, these donors will be almost surely matched to the longest waiting cohort of δ_X measure of patients. Let this cohort have arrived t_X years before the current point in time. As there are $\pi_X[1 - F(t_X)]$ measure of patients, we should have $\pi_X[1 - F(t_X)] = \delta_X$. Hence, at steady state, the entry time of the longest standers in the X wait list can be found through $t_X = F^{-1}(1-\delta_X/\pi_X) < T = F^{-1}(1)$. This is also the waiting time for X blood-type patients subject to survival. We state the following characterization of the wait list at steady state: (See also Figure 3.2.1.)

Theorem 1 (Within-Blood-Type Deceased Donation) *Under the within-blood-type first-come first-served deceased donor allocation policy, at steady state, the (expected) waiting time for X blood-type patients in the wait list is $t_X^w = F^{-1}(1-\delta_X/\pi_X)$, and the measure of their wait list is $\int_0^{t_X} \pi_X[1 - F(t)]dt$.*

3.2.2 "Compatible" Deceased Donation

The following lemmata relate the role of blood-type compatibility relationship to the waiting times of different blood types under the compatible deceased donor allocation policy.

Lemma 2 *For two blood types X and Y, if X blood-type organs are blood-type-compatible with Y blood-type patients, then the compatible deceased donor allocation times of X and Y blood-type patients at steady state satisfy $t_Y^c \leq t_X^c$.*

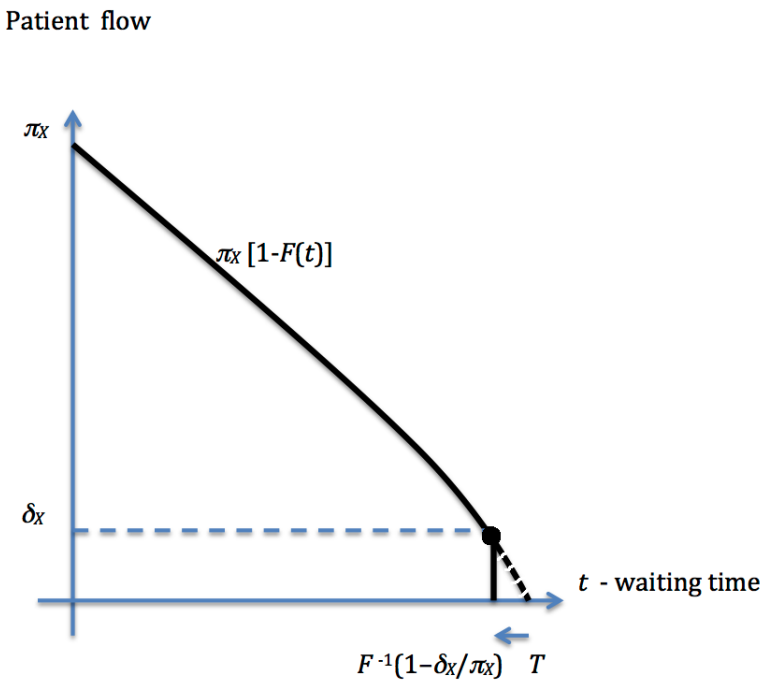


Figure 2: **Deceased donor waiting list** for X blood-type patients at steady state: incoming deceased donors, of δ_X measure, at each time are matched with δ_X measure of patients who have been waiting. Waiting time decreases from T to $F^{-1}(1-\delta_X/\pi_X)$.

Proof. Because of the partial order structure of the blood-type compatibility relationship, since X organs are blood-type-compatible with Y patients, Y organs are not blood-type-compatible with X patients. Moreover, Y patients are blood-type-compatible with all blood types that can feasibly donate to X patients.

Suppose to the contrary of the claim, $t_Y^c > t_X^c$. Then the longest-waiting Y patients would receive maximum number of organs that would otherwise go to X patients, as they are waiting longer than the longest-waiting X patients. Hence, either Y patients do not wait at all, i.e. $t_Y^c = 0$ or X patients never receive transplant $t_X^c = T$. Either case contradicts the assumption. ■

If blood types in some $\mathcal{S} \subseteq \mathcal{T}$ donate organs only to the blood types in $\mathcal{S}$ and they receive organs only from blood types in $\mathcal{S}$ at steady state, and there is no proper subset of $\mathcal{S}$ with this property, then we say that blood types in $\mathcal{S}$ are **pooled**. For example if O blood-type organs are transplanted to A and B blood-type patients besides O, and A and B blood-type organs are only transplanted to A and B blood-type patients, respectively, then $\{O, A, B\}$ is a pooled set. On the other hand neither $\{O, A\}$ is pooled (as O blood-type organs are also transplanted to B blood-type pairs) or $\{A, B\}$ is pooled (as both A and B blood-type patients also receive O blood-type organs). The whole blood type set $\mathcal{T} = \{O, A, B, AB\}$ is not pooled, either, as its proper subset $\{O, A, B\}$ is pooled.

Lemma 3 *For two blood types X and Y, if Y blood-type patients receive X blood-type organs at steady state under the compatible deceased donor allocation policy then $t_X^c = t_Y^c$.*

Moreover, if blood types in $\mathcal{S}$ are pooled together for some $\mathcal{S} \subseteq \mathcal{T}$ then their waiting times are given by

$$t_X^c = t_{\mathcal{S}} \equiv F^{-1}\left(1 - \frac{\sum_{X \in \mathcal{S}} \delta_X}{\sum_{X \in \mathcal{S}} \pi_X}\right) \quad \forall\, X \in \mathcal{S}. \tag{1}$$

Proof. Suppose Y patients receive X organs at steady state under the compatible allocation policy. By Lemma 2, $t_Y^c \leq t_X^c$. Suppose the inequality is strict. Then either all X organs would go to longest-waiting X patients, which would contradict the fact that X organs are transplanted to Y patients, or X patients would not be waiting at all in the waiting list, which would contradict the assumption that $t_Y^c < t_X^c$. Hence, $t_Y^c = t_X^c$.

Next, suppose that blood types in some $\mathcal{S} \subseteq \mathcal{T}$ are pooled together. Then there is a chain of blood types $\{X_1, ..., X_k\} = \mathcal{S}$ such that X_1 receives from X_1 and X_2, ..., X_{k-1} receives from X_{k-1} and X_k. By the previous paragraph, all types in $\mathcal{S}$ have the same waiting time under the compatible allocation scheme. Moreover, the supply demand equations for these types are given as for all $X \in \mathcal{S}$,

$$\sigma_X = \pi_X[1 - F(t_{\mathcal{S}})]$$

where $t_{\mathcal{S}}$ is the common waiting time and σ_X is the measure of organs supplied to X patients. Moreover, $\sum_{X\in\mathcal{S}} \sigma_X = \sum_{X\in\mathcal{S}} \delta_X$. Hence summing up left-hand-sides and right-hand-sides of these equations, respectively, we get $\sum_{X\in\mathcal{S}} \delta_X = (\sum_{X\in\mathcal{S}} \pi_X)[1 - F(t_{\mathcal{S}})]$. Solution for $t_{\mathcal{S}}$ is given as in Equation 1. ■

Observe that $t^w_X = t_{\{X\}}$ as defined in Equation 1 for all blood types X.

Using Lemmata 2 and 3 together with the first-come first-served feature of the wait list policy and the partial order structure of the blood-type compatibility relationship, we can determine which types will be pooled together under compatible deceased donor allocation:

Theorem 2 (Compatible Deceased Donation) *Under the compatible first-come first-served deceased donor allocation policy, at steady state, for two blood types X and Y, if Y has the longest within-blood-type allocation time and X has the shortest within-blood-type allocation time among all blood types that are compatible with Y patients, then X and Y patients will be pooled together (possibly with other types). Moreover, we can treat X and Y as a composite blood type with deceased donor inflow $\delta_X + \delta_Y$ and patient inflow $\pi_X + \pi_Y$ such that its patients are blood-type-compatible with all blood types Y patients can receive from and its donors are blood-type-compatible with all blood types X donors can donate to.*

Proof. Suppose Y has longest t^w, and t^w_X is lowest among all blood types that can donate to Y. If $Y = X$ then the theorem trivially follows. Hence, suppose that $Y \neq X$. We have $t^w_Y > t^w_X$. By Lemma 2, $t^c_Y \leq t^c_X$. As t^w_X is the shortest among t^w for types that Y can receive from, the only way $t^c_Y \leq t^c_X$ can happen is that Y patients receive X organs at steady state or X pools with another type which has a higher t^w than Y. However, the latter is not correct by assumption. Therefore, Y and X patients are pooled (possibly together with other types). By Lemma 3, $t^c_Y = t^c_X$. Moreover, by transferring some of the X organs Y and X patients receive to other compatible-type patients, the waiting time of Y and X patients can be adjusted above $t_{\{X,Y\}}$ but no higher than t^w_Y. Similarly, by transferring some of the X organs that Y patients are receiving to X patients, and substituting those with other compatible organs for Y, the waiting time of Y and X patients can be adjusted below $t_{\{X,AB\}}$ but no lower than t^w_X. Observe that no blood types waiting time can be made shorter than t^w_X or longer than t^w_Y, at steady state, under the constraint of Lemma 3, which says that all donating blood types to Y patients will have the same waiting time. Hence, the composite type of X and Y behaves like Y when it is receiving organs and behaves like X when it is donating organs with deceased donor inflow $\delta_X + \delta_Y$ and patient inflow $\pi_X + \pi_Y$, by Lemma 3. ■

The above theorem can be iteratively used, once we determine whether the longest waiting blood type X is pooled with another type. If it is not, then we can use the above theorem with the rest of

the blood types, as $t^c_X = t^w_X \geq t^c_Y$ for all Y. If it is, on the other hand, then, we can treat X and its pooling type Y as a composite type and use $t^w_{X,Y} = t_{\{X,Y\}}$ (as defined in Equation 1) and apply the above theorem again. By iterating the above analysis, we can determine which types will be pooled with each other.

We state how this theorem can be used to determine the waiting time for different blood types. First suppose that AB patients have the shortest within-blood-type deceased donor allocation time. We discuss what happens when this does not hold later in this subsection.

Proposition 1 *Under the compatible first-come first-served deceased donor allocation policy, at steady state, the (expected) waiting time for X blood-type patients t^c_X can be found as follows when the inflow of deceased donor to patient ratio $\frac{\delta_X}{\pi_X}$ is the largest for AB: AB patients are only served AB organs, and $t^c_X \geq t^c_{AB} = t^w_{AB}$ for all X. Moreover,*

- *If $\frac{\delta_O}{\pi_O} \leq \frac{\delta_A}{\pi_A}, \frac{\delta_B}{\pi_B}$ then O organs are only served to O patients, and hence, A and B patients only receive A and B organs, respectively. Thus, $t^c_O = t^w_O \geq t^c_A = t^w_A, t^c_B = t^w_B \geq t^c_{AB}$.*
- *If $\frac{\delta_B}{\pi_B} \leq \frac{\delta_O}{\pi_O} \leq \frac{\delta_A}{\pi_A}$ then B patients receive O organs, i.e., B and O waiting lists are pooled together, while A patients only receive A organs. Hence, $t^w_B \geq t^c_B = t^c_O = t_{\{O,B\}} \geq t^w_O \geq t^w_A \geq t^c_{AB}$.*[7]
- *If $\frac{\delta_B}{\pi_B} \leq \frac{\delta_A}{\pi_A} \leq \frac{\delta_O}{\pi_O}$ then two subcases are possible:*[8]
  - *If $\frac{\delta_A}{\pi_A} \leq \frac{\delta_B+\delta_O}{\pi_B+\pi_O}$ then O organs are served to O, A, and B patients, i.e., O, A, B types are pooled together. Hence, $t^w_B \geq t^w_B \geq t^c_O = t^c_A = t^c_B = t_{\{O,A,B\}} \geq t^w_O \geq t^c_{AB}$.*
  - *Otherwise, then O organs are only served to O and B patients, i.e., O and B types are pooled and A patients only receive A organs. Hence, $t^w_B \geq t^c_O = t^c_B = t_{\{O,B\}} \geq t^c_A = t^w_A \geq t^w_O \geq t^c_{AB}$.*

Proof. It follows from Theorem 2 and the fact that $t_{\mathcal{S}}$ is decreasing in $\frac{\sum_{X\in\mathcal{S}}\delta_X}{\sum_{X\in\mathcal{S}}\pi_X}$ for all non-empty $\mathcal{S} \subseteq \mathcal{T}$. ■

When AB blood type does not have the shortest within-blood-type deceased donor allocation time, AB patients will be pooled with some other blood type patients. Determining which types will be pooled together with AB types requires more detailed analysis. They can be pooled together with all types or exclusively with A, B, or both without O. We can use Theorem 2 for this analysis.

[7] The case $\frac{\delta_A}{\pi_A} \leq \frac{\delta_O}{\pi_O} \leq \frac{\delta_B}{\pi_B}$ is the symmetric version of this case with respect to A and B.

[8] The case $\frac{\delta_A}{\pi_A} \leq \frac{\delta_B}{\pi_B} \leq \frac{\delta_O}{\pi_O}$ is the symmetric version of this case with respect to A and B.

In reality AB blood-type patients have a very small ratio among all patients, and regardless of which blood types they are pooled together this may not change how the other blood types will be treated with respect to each other. The following assumption ensures this: For all $X, Y \in \{O, A, B\}$, if $\frac{\delta_X}{\pi_X} < \frac{\delta_Y}{\pi_Y}$ then $\frac{\delta_X+\delta_{AB}}{\pi_X+\pi_{AB}} < \frac{\delta_Y}{\pi_Y}$. If this holds, AB types will be pooled with the blood type X that has the shortest within-blood-type allocation time. Then, we can treat those two types as one and figure out whether they will be pooled with other types using the above proposition. In this case, we need to use $\delta_X + \delta_{AB}$ instead of δ_X and $\pi_X + \pi_{AB}$ instead of π_X, and all our analysis will go through.

3.3 Policy Discussion: Theory, Empirics, and Their Implications

4 Live Donation

Organs such as kidney and liver have live donation possibilities. Especially live kidney donation is very common and $PP\%$ of all donation has been by live donors in 2011.[9]

We will refer to a live donor as a **paired donor**. We will assume that each patient can have at most one paired donor. We assume λ fraction of incoming patients have a paired donor (such as a spouse). We also assume that the blood types of the patient and donor are independent and uncorrelated.[10] The patient and his paired donor are represented as a **pair**. The blood types of the pair, $X - Y \in \mathcal{T} \times \mathcal{T}$, X being the patient's and Y being the donor's blood type, determines the type of the pair.

If the paired-donor of a patient is both blood-type and tissue-type compatible then we refer to the pair as a **compatible** pair, otherwise it is an **incompatible** pair. Recall that there is a χ probability chance that a blood-type compatible donor being tissue-type incompatible with a patient. Let p_Y be the probability of a patient having a Y blood-type donor, given that he has a donor. [11]

We assume in the rest of the paper that

[9] Each human has two kidneys and can have a perfectly healthy life with a single kidney. Also the risk associated with live donation surgery is very small. There is $PP\%$ chance that something will go wrong for the donor, and $PP\%$ chance that the donor will die complications due to surgery. On the other hand, live-donor liver donation is done through donation of part of a liver, and it is much riskier (there is $PP\%$ chance that he donor will die due to complications associated with donation). The ratio of live donation is much smaller, $PP\%$ for liver.

[10] In reality, if the paired donor is a blood relative of the patient, the blood types of the patient and donor are correlated through degree of relation and genetic laws. Hence, potentially figuring out the exact correlation can be complicated. For our purposes, we simply assume the blood types of the patient and his paired donor are uncorrelated to make our arguments.

[11] We assume that population blood-type distribution and donor blood-type distribution are identical.

- Fraction λ is not large enough so that deceased donors are not enough for all patients without paired donors; i.e., $\pi_X(1-\lambda) \leq \delta_X$ for all X; and

- incompatible pairs with different donor and patient blood types are less in measure then their reciprocal type pairs; i.e., for all $X \neq Y$ such that donor blood type Y can feasibly donate to X, we have $\chi p_Y \pi_X < p_X \pi_Y$.[12]

These are both supported by data. The first condition ensures that – as in real life – availability of live donation does not solve the organ shortage problem. The second condition is related to live-donor exchange, which we start inspecting in the the next section. It ensures that incompatible pairs that are blood-type compatible are less in measure than their reciprocal type pairs.

We can calculate the inflow measures of different compatible and incompatible pair types:

- An O blood-type patient needs an O blood-type donor. Thus, $\pi_O p_O \lambda(1-\chi)$ is the inflow measure of O blood-type patients with a compatible live donor. On the other hand, $\pi_O p_O \lambda \chi$ is the measure of incompatible $O-O$ pairs, $\pi_O p_Y \lambda$ is the measure of $O-Y$ pairs with $Y \in \{A, B, AB\}$, who are all incompatible.

- An $X \in \{A, B\}$ blood-type patient can get an organ from O or X blood-type donor. Thus, given $Y \in \{X, O\}$, $\pi_X p_Y \lambda(1-\chi)$ is the inflow measure of X blood-type patients with a compatible Y live donor who become sick per unit time; on the other hand, $\pi_X p_Y \lambda \chi$ is the measure of incompatible $X-Y$ pairs. We have $\pi_X p_Y \lambda$ as the inflow measure of $X-Y$ pairs with $Y \in \{A, B, AB\} \setminus \{X\}$. The latter are incompatible pairs.

- An AB blood-type patient can get an organ from all blood-type donors. Thus, $\pi_{AB} p_Y \lambda(1-\chi)$ is the inflow measure of compatible $AB-Y$ pairs, and $\pi_{AB} p_Y \lambda \chi$ is the inflow measure of incompatible $AB-Y$ pairs for all $Y \in \mathcal{T} = \{O, A, B, AB\}$.

Although, live donation is a great tool to serve more organ patients, due to the partial order structure of blood-type compatibility among blood types, not all blood types will be affected equally when live donation is possible. For example, O blood type patients are at a disadvantage with respect to other types in finding a compatible paired donor. Although deceased donor waiting times will fall for all types, this tool will be the least advantageous to the O patients in general. Hence, the deceased donor allocation times for O patients will decrease the least. On the other hand, seemingly A and B

[12] A simple requirement that would make the second condition of the assumption hold is that donor and patient inflow rates across blood types have a similar ratio i.e., $\pi_X/\pi_Y \approx p_X/p_Y$ for all blood types X, Y. This would be ensured if live donation and getting sick rates are not too different for different blood types.

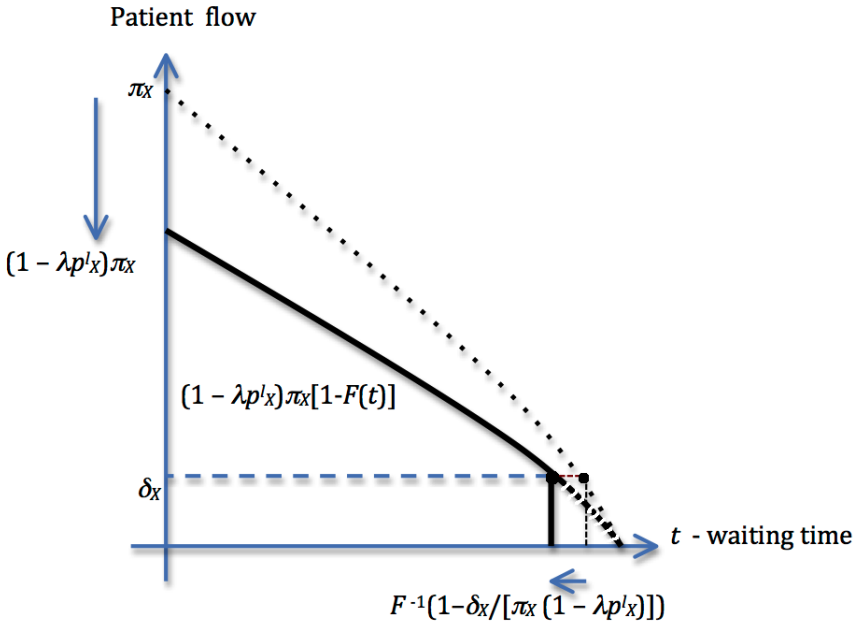


Figure 3: **Deceased donor waiting list under live donation** for X blood-type patients at steady state: Inflow π_X of patients decreases by $\lambda p^l_X \pi_X$. As a result waiting time decreases from $F^{-1}(1 - \delta_X/\pi_X)$ to $F^{-1}\left(1 - \frac{\delta_X}{\pi_X(1-\lambda p^l_X)}\right)$.

patients are symmetric. However, in general A blood type is more prominent in the population than B. Therefore, at random A patients will have a higher chance of finding a compatible paired donor then B types, given that they can all receive from O donors as well as their own types. Finally, AB patients will have the highest reduction in deceased donor allocation time among all blood types with the help of live donation. [13] We can state this conclusion as our main result of this section.

Theorem 3 (Direct Live Donation) *Live donation will unambiguously decrease the steady state deceased donor allocation waiting times for all patients. Patients with compatible live donors will not wait at all. If we have equal deceased donation ratios, $\frac{\delta_X}{\pi_X}$, for all blood types X, then regardless of the deceased-donor allocation protocol*

- *O patients will benefit with the lowest reduction in waiting time;*
- *AB patients will benefit with the highest reduction in waiting time; and*

[13] Although these conclusions seem to have been reached with the help of our assumption that blood types of patients are uncorrelated with their paired donors, a version of this result will also hold true even if there is positive correlation in a pair's blood types; however the magnitude of the difference in eventual waiting times will not be as extreme.

- *if additionally $p_A > p_B$, A patients will benefit from live donation more than B patients in terms of reduction in waiting time.*

Proof. We derive the waiting times under live donation for within-blood-type deceased-donor allocation policy. Let p^l_X be compatibility probability of an X blood type patient with his paired live donor. Observe that we have $p^l_O = p_O(1-\chi)$, $p^l_A = (p_O + p_A)(1-\chi)$, $p^l_B = (p_O + p_B)(1-\chi)$, and $p^l_{AB} = 1-\chi$. Then the inflow measure of of X blood type patients to the wait list decrease by $\pi_X \lambda p^l_X$ (see Figure 4). Hence, the waiting time for X blood-type patients in the wait list under within-blood-type deceased donor allocation policy is given by

$$t^{l/w}_X = F^{-1}\left(1 - \frac{\delta_X}{\pi_X(1-\lambda\, p^l_X)}\right). \tag{2}$$

Observe that $p^l_O < p^l_A, p^l_B < p_{AB}$. Hence AB has the highest reduction while O has the lowest reduction in waiting time provided that $\frac{\delta_X}{\pi_X}$ is equal for all blood types X. If $p_A > p_B$ then $p^l_A > p^l_B$, and hence A has a higher reduction in waiting time than B. Hence, in this case, $t^{l/w}_O > t^{l/w}_B > t^{l/w}_A > t^{l/w}_{AB}$. Replacing $\frac{\delta_X}{\pi_X}$ with $\frac{\delta_X}{\pi_X(1-\lambda\, p^l_X}$ for all X in Proposition 1, deceased-donor kidneys are only served to their own blood-type patients under compatible deceased-donor allocation policy. Hence, this policy also results with the same waiting times as within-blood-type allocation, i.e., $t^{l/c}_X = t^{l/w}_X$ for all X. ■

5 Live-Donor Exchange

For the compatible pairs, there is no reason for waiting for a deceased-donor organ. The paired-donor can immediately donate to the patient. Hence, compatible paired donors immediately donate to their patients, and these patients do not enter the wait list queue.

On the other hand, there is a parallel market for donor exchange for incompatible pairs. While waiting for a deceased-donor organ in the waiting list, they also wait for a paired-exchange to be conducted with another incompatible pair. A paired-exchange matches two pairs where the patient of the first pair is compatible with the second pair and the patient of the second pair is compatible with the donor of the first pair. We refer to such pairs as **mutually compatible** pairs. [14]

There can be different policies determining which types of mutually compatible pairs are matched with each other, as a pair type can be mutually compatible with several other types. We will

[14]We can also think of exchanges that can match more than two pairs, such as 3-way, 4-way etc. For simplicity we will focus on 2-way exchanges in our analysis, however, our results can easily be extended to cover 3-way and 4-way exchanges as in Roth, Sönmez, and Ünver (2007). Any sizes of exchanges greater than 4 will not change the results as reported in this paper.

assume that the donor exchange is conducted in an optimal manner and inspect two different policies regarding compatible pairs. However while selecting among a particular pair from a given type $X-Y$, organ exchange is also operated in a first-come first-served basis.

We assume in modeling that for pairs who arrive at the same time, a maximal matching is determined among the tissue-type compatible patients of the same blood-type and they are immediately matched. It turns out that this is possible almost surely as long as there are more or equal measure of patients than donors. We state a slightly different version of Lemma 1 in this case, which immediately follows from Erdös and Rényi (1960) random graph convergence theorem:

Lemma 4 (Live-Donor Exchange Matching Protocol) *Suppose that there is an ω measure of blood-type X patients available in the queue and a $\sigma \leq \omega$ measure of blood-type X donors arrive. Then, almost surely there exists a matching that matches all donors with tissue-type compatible patients.*

Using the terminology in Ünver (2010), we classify the pairs into several categories, based on their desirability in exchange.

We refer to type $X-Y$ as the **reciprocal** of pair type $Y-X$. **Overdemanded pair types** are the ones with a blood type donor which can donate to her patient's blood type yet they are not of the same blood type. There are $A-O, B-O, AB-A, AB-B, AB-O$ pair types. **Underdemanded pair types** are those with a blood type donor that cannot feasibly donate to her patient's blood type, excluding types A-B and B-A. That is, underdemanded types are reciprocals of overdemanded types, and they include $O-A, B-O, A-AB, B-AB, O-AB$. **Reciprocally demanded pair types** are $A-B$ and $B-A$, as they can be matched with each other in a donor exchange, when tissue incompatibility does not exist. Finally **Self-demanded pair types** are the ones with the same blood-type donor and patient: $O-O, A-A, B-B, AB-AB$.

The names associated with these classes will be more meaningful after our analysis. The following lemma shows the role of overdemanded types in exchange (similar results were also reported in Roth, Sönmez, and Ünver (2007) and Ünver (2010)):

Lemma 5 (Live-donor Exchange Blood-Type Feasibility) *An underdemanded type pair can only be matched with an overdemanded type pair in an exchange, and overdemanded types can be used to match other overdemanded, underdemanded, reciprocally demanded, or self-demanded type pairs. Additionally, reciprocally demanded type pairs can only be used to match the other reciprocally demanded type pairs and self-demanded type pairs can only be used to match their own type pairs. In particular:*

- *An underdemanded $O-A$ ($O-B$, respectively) type pair can only be matched in an exchange with a pair from overdemanded types $A-O$ ($B-O$, respectively) or $AB-O$. An underdemanded $A-AB$ ($B-AB$, respectively) type pair can only be matched in an exchange with a pair from overdemanded types $AB-A$ ($AB-B$, respectively) or $AB-O$. An underdemanded $O-AB$ type pair can only be matched in an exchange with an overdemanded $AB-O$ type pair.*

- *A reciprocally demanded $A-B$ ($B-A$, respectively) type pair can only be matched in an exchange with a pair from the other reciprocally demanded $B-A$ ($A-B$, respectively) or from overdemanded types $AB-A$ ($AB-B$, respectively) or $AB-O$.*

- *A self-demanded $X-X$ type pair can be matched in an exchange with a pair from the same type. Additionally, an $O-O$ type pair can only be matched with a pair from overdemanded types $A-O, B-O$, or $AB-O$; an $A-A$ ($B-B$, respectively) type pair can only be matched with a pair from overdemanded types $AB-A$ ($AB-B$, respectively) and $AB-O$; and an $AB-AB$ type pair can only be matched with a pair from overdemanded types $AB-A, AB-B$, or $AB-O$.*

5.1 Live-Donor Exchange, Live-Donor Donation, and Deceased-Donor Wait List

In this subsection, we model how the live-donor exchange pool and deceased-donor wait list would evolve under live-donor donation and optimal exchange technologies. Recall that only incompatible patient-live-donor pairs participate in exchange.

Theorem 4 (Optimal Live-donor Exchange Rule) *A policy that dictates matching the longest-waiting pairs of a type with their longest-waiting reciprocal type pairs constitutes an optimal live-donor exchange policy.*

Proof. Suppose for all $X \neq Y$ such that donor blood type Y is compatible with patient blood type X, we have $\chi p_Y \pi_X < p_X \pi_Y$ by assumption. Recall that the inflow of each $X-Y$ type pair into the exchange pool was calculated as follows: For an overdemanded or self-demanded type $X-Y$, $\chi\lambda p_Y \pi_X$ is the inflow measure, and for an underdemanded or reciprocally demanded type $X-Y$, $\lambda p_Y \pi_X$ is the inflow measure. Hence, by Lemma 5, as long as the inflow measure of an underdemanded $X-Y$ type pairs is not smaller than the number of reciprocal overdemanded $Y-X$ type pairs, which is ensured by assumption, then the only way of matching most number of patients through live-donor exchange is to match underdemanded pairs with their reciprocals, as self-demanded pairs can be matched

with their own type pairs, and reciprocally-demanded-type $A-B$ pairs can be matched with $B-A$ pairs. Then under the assumptions, the measure of overdemanded type pairs and self-demanded pairs waiting in the pool will be zero as they will immediately be matched to their reciprocal types by Lemma 4. ■

Theorem 5 (Steady State under Live-Donor Exchange I) *Consider the within-blood-type deceased-donor allocation policy and optimal live-donor exchange policy. Consider a blood type X. If $\frac{\delta_X}{1-\lambda} \leq \frac{\chi p_X \pi_Y}{p_Y}$ for all blood types Y, to which X can feasibly donate, and $\frac{\delta_X}{1-\lambda} \leq \frac{p_X \pi_Y}{p_Y}$ if there is Y such that $\{X,Y\} = \{A,B\}$, then*

- *no pair with an X blood-type patient receives deceased-donor organ;*
- *for any Y, the following hold for $X-Y$ pairs:*
  - *if such a pair is compatible, then the patient immediately receives his donor's organ;*
  - *if such a pair is incompatible and yet blood-type Y can feasibly donate to X, then the patient receives immediately an organ through exchange with a $Y-X$ pair;*
  - *if blood-type Y cannot feasibly donate to X, then subject to survival, such a pair is matched with a $Y-X$ pair through exchange after waiting time*

$$t^e_{X-Y} \;=\; F^{-1}\left(\max\left\{1-\frac{\rho p_X \pi_Y}{p_Y \pi_X}, 0\right\}\right), \tag{3}$$

  where $\rho = \begin{cases} 1 & \text{if } \{X,Y\} = \{A,B\} \\ \chi & \text{otherwise} \end{cases}$ *; and*

- *X blood-type patients without paired donors are matched with deceased donors subject to survival after waiting time*

$$t^{w/e}_X = F^{-1}\left(1-\frac{\delta_X}{\pi_X(1-\lambda)}\right). \tag{4}$$

Proof. ■

Theorem 5's conditions ensure that the chance for exchange for underdemanded $X-Y$ type pairs (thorough reciprocal incompatible overdemanded $Y-X$) pairs is at least as large as receiving deceased donation for blood-type X patients without a paired-donor (and similarly for reciprocally demanded $A-B$ and $B-A$ type pairs in case $X-Y$ is one of them). However, this may not always be true in reality. In this case, if for some type $X-Y$

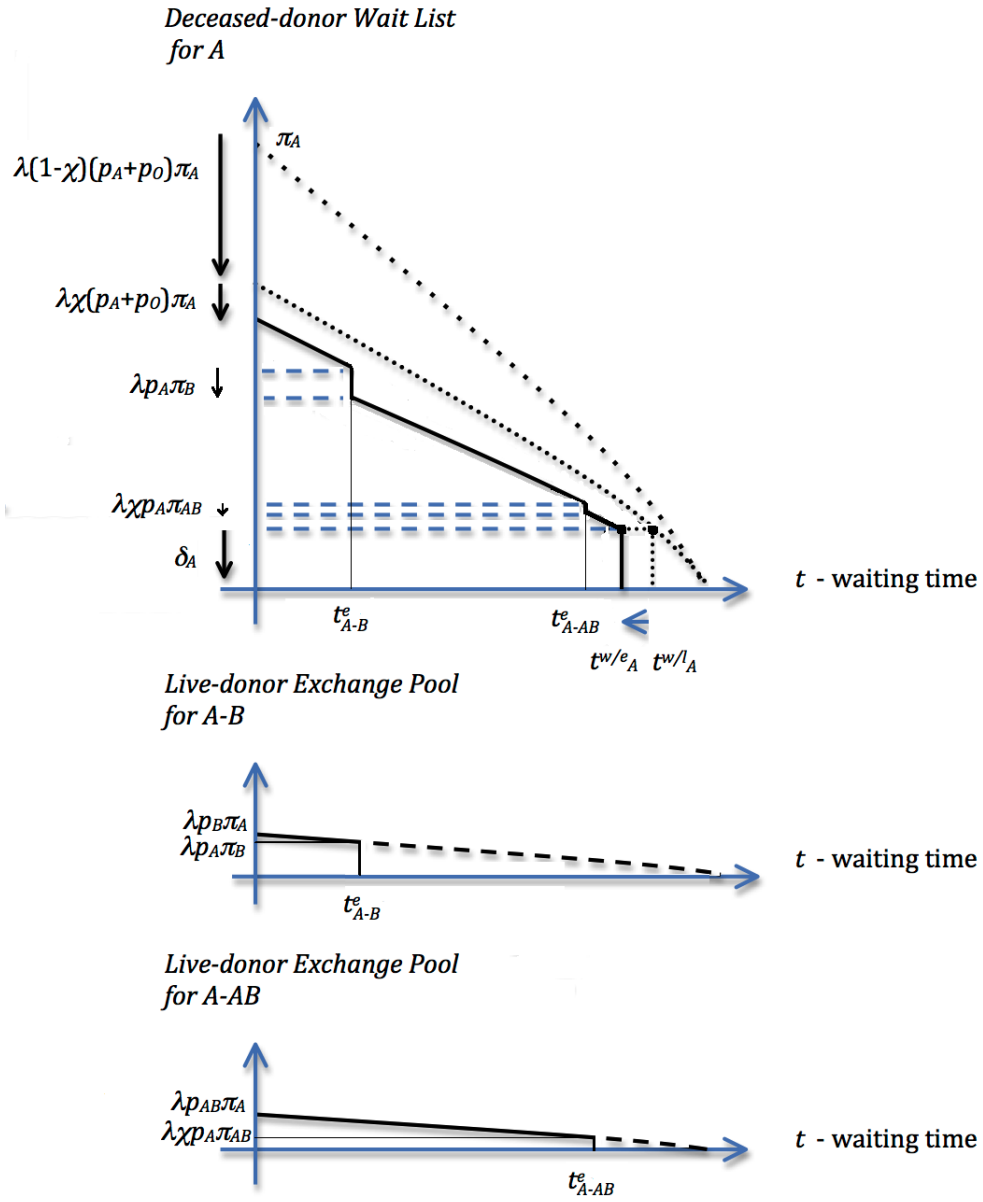


Figure 4: **Deceased-donor wait list under direct live donation and exchange of incompatible live-donors** for A blood-type patients at steady state: Inflow π_A of patients decreases by $\lambda(1-\chi)(p_A+p_O)\pi_A = \lambda p^l_A \pi_A$ as a result of immediate compatible direct live donation and a further $\lambda\chi(p_A+p_O)\pi_A$ as a result of incompatible live donor exchange for $A-O$ and $A-A$ types who do not wait in the wait list. Assuming that $p_A\pi_B \geq p_A\pi_B$ at time t^e_{A-B}, a measure of $\lambda p_A\pi_B$ of $A-B$ pairs are matched through exchange. At time t^e_{A-AB} a measure of $\lambda\chi p_A\pi_{AB}$ of $A-AB$ pairs are matched through exchange. As a result waiting time in the deceased-donor wait list decreases from $t^{w/l}_A$ to $t^{w/e}_A$.

6 A New Proposal: How to Engage Compatible Pairs in Exchange

7 The Competition between Multiple Exchange Programs and Compatible Pairs

8 Conclusion

References

Erdös, P., and A. Rényi (1960): "The Evolution of Random Graphs," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5, 17–61.

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

——— (2005): "A Kidney Exchange Clearinghouse in New England," *American Economic Review Papers and Proceedings*, 95, 376–380.

——— (2007): "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review*, 97, 828–851.

Sönmez, T., and M. U. Ünver (2010): "Altruistically Unbalanced Kidney Exchange," Working paper.
ΩÜnver

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