Appendix to "Market Failure in Kidney Exchange"

Nikhil Agarwal, Itai Ashlagi, Eduardo Azevedo,
Clayton Featherstone, and Ömer Karaduman

March 25, 2019

A Extensions

A.1 Welfare gains from optimal point mechanisms

This section estimates the gain in welfare from moving to the point mechanism described above. This gain is equal to the deadweight loss that can be avoided by rewarding hospitals optimally as in Theorem 1. The following proposition derives equation (**??**).

Proposition A.1. *Consider an aggregate supply of pairs, $\boldsymbol{q}_0$, that results when hospitals choose supply optimally given rewards, $\boldsymbol{p}_0$. Further, consider strictly positive aggregate supply, $\boldsymbol{q}^*$, and rewards, $\boldsymbol{p}^*$, that maximize hospital welfare as in Theorem 1. Assume that the matrix $\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)$ is finite and non-singular and that the production function has constant returns to scale at $\boldsymbol{q}^*$. Then, the deadweight loss in hospital welfare at $\boldsymbol{q}_0$ can be approximated by either*

$$\frac{1}{2}\left[\boldsymbol{\nabla f}\left(\boldsymbol{q}_0\right) - \boldsymbol{p}_0\right] \cdot \left(\boldsymbol{q}^* - \boldsymbol{q}_0\right).$$

or

$$\frac{1}{2}\left[\boldsymbol{\nabla f}\left(\boldsymbol{q}_0\right) - \boldsymbol{p}_0\right] \cdot \left[\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)\right]^{-1} \left[\boldsymbol{\nabla f}\left(\boldsymbol{q}_0\right) - \boldsymbol{p}_0\right]'.$$

The error in both approximations is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$.

These formulas are a multidimensional version of the Harberger triangle formulas in one-dimensional linear commodity taxation. The first formula is the multidimensional version of the one half base times height formula. The second formula is the equivalent of the one half of the tax wedge squared times the inverse of the difference between the derivative of inverse supply and the derivative of inverse demand. The second formula shows that the deadweight loss is one half of a quadratic expression in the wedge $\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0$. The term $\boldsymbol{DP}_S(\boldsymbol{q}_0)$ accounts for the fact that a more elastic supply leads to larger deadweight losses. The term $\boldsymbol{D}^2\boldsymbol{f}$ accounts for the change in marginal products in response to a change in $\boldsymbol{q}$. For example, the deadweight loss is lower if increasing the supply of overdemanded pairs results in these pairs becoming less useful.

We use equation (**??**) to quantify the deadweight losses for a range of supply elasticities. We restrict attention to mechanisms that set reward vectors for the categories in the regression tree analysis above (Figure A1). The wedge $\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0$ and the curvature matrix $\boldsymbol{D}^2\boldsymbol{f}$ for these categories are estimated using our production function. To use equation (**??**), we need to specify supply elasticities through the matrix $\boldsymbol{DP}_S(\boldsymbol{q}_0)$. One challenge in directly specifying this quantity is that different submission types may respond differently to rewards. For example, the submission of hard-to-match types to the system may not substantially decrease when rewards are lowered because there are few other avenues for matching them. Our approach is to calculate the maximum deadweight loss under varying bounds on the maximum elasticity of any type of submission. This method allows us to be agnostic about the supply elasticities of different submission types. The deadweight loss is zero when we assume that the maximum elasticity is zero because the submissions will not respond to the

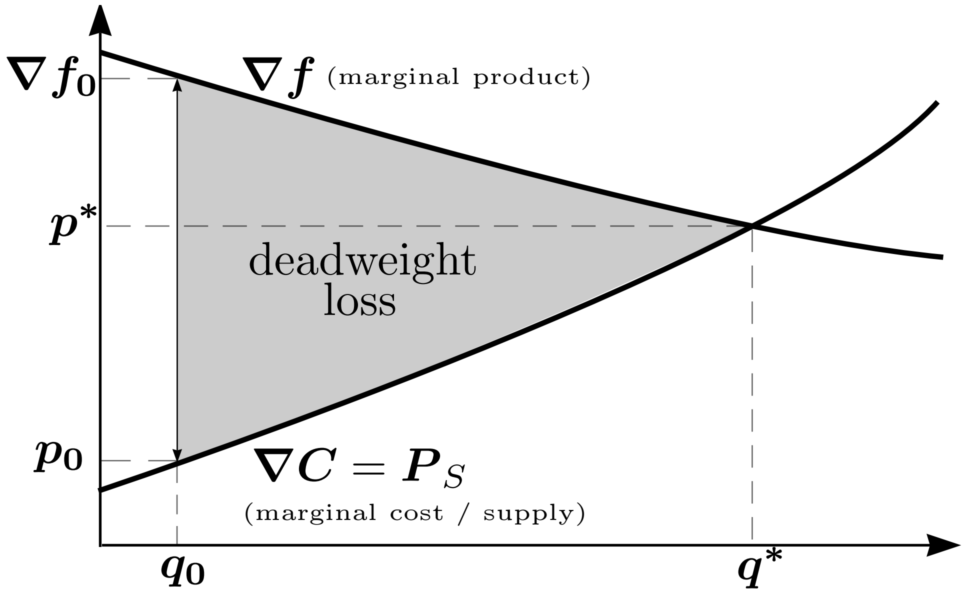


Figure A1: Hospital-Welfare Deadweight Loss from the Current Mechanism

Notes: The horizontal axis represents aggregate quantity and the vertical axis represents rewards vectors, marginal costs and marginal products, so both axes represent I-dimensional vectors. The deadweight loss from the current mechanism is the shaded area between marginal products and the supply curve (mathematically, the area is a path integral going from current rewards $\boldsymbol{p}_0$ to optimal rewards $\boldsymbol{p}^*$). Current rewards are $\boldsymbol{p}_0$, equal to the probability of matching each type of submission, while optimal rewards $\boldsymbol{p}^*$ equal marginal products. Current quantities $\boldsymbol{q}_0$ and rewards $\boldsymbol{p}_0$ are observed. Marginal products $\boldsymbol{\nabla f}$, including the current value $\boldsymbol{\nabla f}(\boldsymbol{q}_0)$, can be calculated from the production function. In contrast, the supply curve $\boldsymbol{P}_S = \boldsymbol{\nabla C}$ and optimal rewards $\boldsymbol{p}^*$ and quantities $\boldsymbol{q}^*$ are not observed and depend on the elasticity of supply.

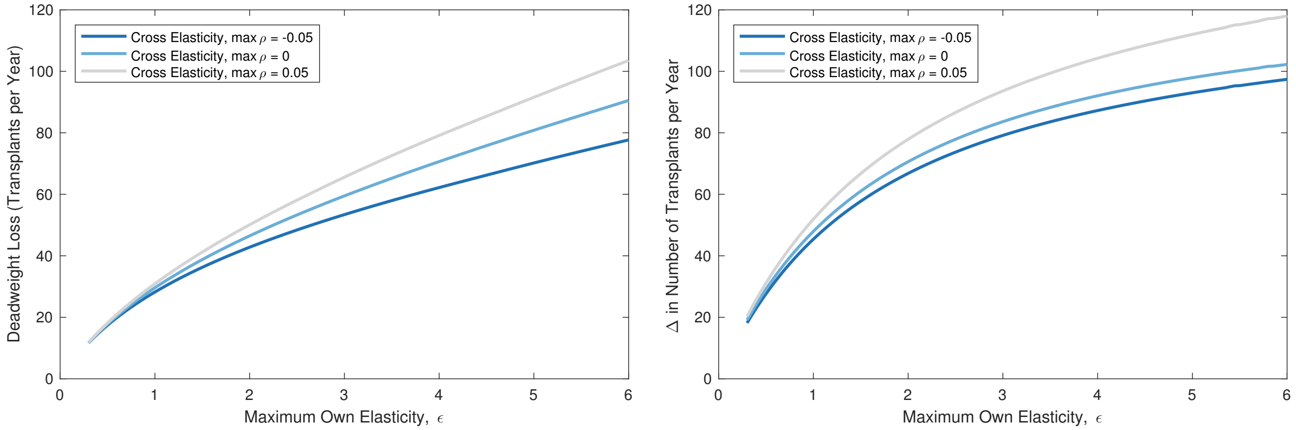
rewards system, resulting in $\boldsymbol{q}^* = \boldsymbol{q}_0$. As we increase the bound on the elasticity, submissions respond and the maximum implied deadweight loss increases. Further, we repeat this exercise for varying assumptions on cross-elasticities.[1]

Figure A2a plots the maximum hospital deadweight loss for bounds on the own-price elasticities ranging from 0 to 6. The curve in the middle describes the results for zero cross-price elasticities, and the other two curves present results for non-zero cross-price elasticities. The hospital deadweight loss is zero if supply is perfectly inelastic and is increasing in elasticity. The deadweight loss is above 40 transplants per year if the maximum elasticity is at

[1] Specifically, we solved the problem

$$\begin{aligned}
\max_{\boldsymbol{DP}_S(\boldsymbol{q}_0)} \quad & \tfrac{1}{2}(\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0)[\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)]^{-1}(\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0)' \\
\text{s.t.} \quad & \left(\frac{\partial P_{S,j}}{\partial q_j}\right)^{-1} \frac{p_{0,j}}{q_{0,j}} \in [0, \varepsilon]\,, \text{ for all } j \in 1 \ldots I, \\
& \left(\frac{\partial P_{S,k}}{\partial q_j}\right)^{-1} \frac{p_{0,k}}{q_{0,j}} = \rho \frac{\left(\frac{\partial P_{S,j}}{\partial q_j}\right)^{-1} \frac{p_{0,j}}{q_{0,j}} + \left(\frac{\partial P_{S,k}}{\partial q_k}\right)^{-1} \frac{p_{0,k}}{q_{0,k}}}{2}, \text{ for all } j,k \in 1 \ldots I \text{ with } k \neq j.
\end{aligned}$$

for each value of the bound on elasticities, ε.



(a) Hospital-Welfare Deadweight Loss

(b) Transplants Lost in the Platform

Figure A2: Losses Due to the Current Mechanism

Notes: Estimated losses from the current mechanism, using the approximation from Proposition A.1, as a function of the elasticity matrix of supply. Maximum own-elasticities are in the horizontal axis. The parameter ρ governs the cross-price elasticity of supply as formally described in footnote 1.

least 2. For very high elasticities, the deadweight loss increases at a slower rate because of production function curvature. The deadweight loss at an elasticity of 6 is only between 75 and 105 because the marginal products of the productive types that the optimal mechanism attracts decrease with supply. Although the results for large elasticities are subject to greater approximation error, it is unlikely that the deadweight losses come close to the efficiency loss relative to the first-best allocation, even for elasticities of about 6.

As discussed in Section 6.2, hospital deadweight losses will understate the loss in social welfare if hospitals undervalue transplants. Figure A2b shows the total increase in transplants facilitated by the NKR if it adopts the optimal points system. To do this, we added the area under $\boldsymbol{P}_S = \boldsymbol{\nabla C}$ to the hospital deadweight loss numbers calculated above (see Figure A1). Because a transplant increase at the NKR will come at the cost of fewer transplants at hospitals, this calculation overstates the loss in total welfare from the current mechanism. Not surprisingly, the estimated losses are higher than the previous figure. A little over 40 transplants are lost if the maximum elasticity is 1. This number is between 95 and 120 for an elasticity of 6. Therefore, social deadweight loss is higher than hospital deadweight loss, but the two are qualitatively similar.

A.2 Maximizing social welfare

Theorem 1 describes mechanisms that maximize hospital welfare. A natural alternative would be to use mechanisms that directly maximize social welfare. However, since hospitals consider private rather than social cost in response to a rewards vector, they won't necessarily choose submissions vectors that minimize the aggregate social cost. To account for this subtlety, define the **reward-moderated social cost** by

$$SC(\boldsymbol{q}) = \sum_{h=1}^{H} SC^h \left(\boldsymbol{S}^h \left(\boldsymbol{P}_S(\boldsymbol{q})\right)\right).$$

We assume that each hospital's supply is single-valued to ensure that this function is well-defined.

The reward-moderated social cost represents the aggregate social cost of inducing an aggregate supply $\boldsymbol{q}$ by using a linear and anonymous rewards scheme. Our next result describes the rewards in mechanisms that maximize social welfare.

Proposition A.2 (Optimal Rewards for Maximizing Social Welfare)**.** *Consider a vector of rewards $\boldsymbol{p}$ and an allocation $\left(\boldsymbol{q}^h\right)_{h=1}^{H}$ with strictly positive aggregate quantity $\boldsymbol{q} = \sum_h \boldsymbol{q}^h$ that maximize social welfare subject to all hospitals choosing $\boldsymbol{q}^h \in \boldsymbol{S}^h(\boldsymbol{p})$ and subject to the total rewards allocated being the same as the number of transplants produced, that is, $f(\boldsymbol{q}) = \boldsymbol{p}\cdot\boldsymbol{q}$. Assume the production function has constant returns to scale at the optimal $\boldsymbol{q}$, and $SC(\boldsymbol{q})$ is differentiable. Then:*

1. *The platform rewards each type of submission with its marginal product plus an adjustment term.*
$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) + \boldsymbol{A}^{SW}(\boldsymbol{q}),$$
where
$$\boldsymbol{A}^{SW}(\boldsymbol{q}) = \frac{1}{1+\lambda^{SW}}\left[\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})\right] - \frac{\lambda^{SW}}{1+\lambda^{SW}}\boldsymbol{q'DP}_S(\boldsymbol{q}).$$
and
$$\lambda^{SW} = \frac{\left[\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})\right]\cdot\boldsymbol{q}}{\boldsymbol{q'DP}(\boldsymbol{q})\cdot\boldsymbol{q}}.$$

2. *The adjustment term $\boldsymbol{A}^{SW}(\boldsymbol{q})$ can be non-zero even with constant returns to scale at $\boldsymbol{q}$.*

3. *The allocation $(\boldsymbol{q}^h)_{h=1}^{H}$ maximizes $f(\boldsymbol{q}) - SC(\boldsymbol{q})$ if and only if, at the optimum, the average wedge between marginal cost and marginal reward-moderated social cost, $\left[\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})\right]\cdot\boldsymbol{q}$, is zero.*

Part 1 shows the optimal mechanism rewards submissions by their marginal products plus an adjustment term. The adjustment equals an externality term, which is greater for submissions whose marginal social costs are less than their marginal private costs, minus a shading term that depends on elasticities. At the optimum, hospitals are rewarded for their marginal contributions to the platform as well as to compensate them for any components of marginal private cost that aren't a part of marginal social cost. However, if there are not enough transplants to pay for these differences, the planner has to shade rewards. As in optimal linear commodity taxation, it is also better to shade rewards for submissions with more inelastic supply.

Part 2 shows that the key difference in this case, relative to Theorem 1, is that the adjustment term is not zero, even for constant returns to scale. Therefore, the optimal rewards depend on more information. To set optimal rewards, one must know, for each type of submission, the wedge between marginal private and social costs. Such knowledge requires identifying the submission types for which hospital objectives deviate most from social objectives. Moreover, one needs to know the elasticity matrix in order to measure how much shading must be done for each submission type. Elasticities matter so long as the average wedge between marginal private and social cost is non-zero because it results in the multiplier λ^{SW} being non-zero and an adjustment term that depends on elasticities. Finally, part 3 shows that the optimal reward vector does not attain first-best social welfare. Therefore, allocations that achieve first-best social welfare require non-linear and complex incentives for hospitals.

Taken together, using only the kidney exchange mechanism to maximize social welfare, as opposed to hospital welfare, runs into important challenges. Optimal rewards are more complex, depend on more information, and are sensitive to changes in the incentives facing hospitals that can affect overall externalities. These results suggest that solving agency problems is an important complement to improving the design of the kidney exchange mechanism.

A.3 Competing platforms

Two natural policy responses to the fragmentation are to mandate participation in a single platform or to merge platforms. These recommendations raise questions about the optimal strategy for competing platforms and the efficiency costs of imperfect competition. We now consider a platform that maximizes the number of transplants it facilitates.

Proposition A.3 (Oligopolistic Platforms)**.** *Consider a platform facing a smooth inverse residual supply curve of submissions $\boldsymbol{P}_{RS}(\cdot)$. Consider a vector of rewards $\boldsymbol{p}$ and a strictly positive aggregate quantity $\boldsymbol{q}$ that maximize the number of transplants in the platform subject to $\boldsymbol{p} = \boldsymbol{P}_{RS}(\boldsymbol{q})$ and subject to allocating the same number of transplants that are produced, that is, $\boldsymbol{p} \cdot \boldsymbol{q} = f(\boldsymbol{q})$. Assume the production function has constant returns to scale at the optimal $\boldsymbol{q}$. Then:*

1. *The platform rewards each type of submission with its marginal product, plus an adjustment term,*
$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) + \boldsymbol{A^C}(\boldsymbol{q}),$$
where
$$\boldsymbol{A^C}(\boldsymbol{q}) = \left(\frac{\boldsymbol{q'DP}_{RS}(\boldsymbol{q}) \cdot \boldsymbol{q}}{f(\boldsymbol{q})}\right) \boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{q'DP}_{RS}(\boldsymbol{q}).$$

2. *The adjustment term $\boldsymbol{A^C}(\boldsymbol{q})$ can be non-zero, even with constant returns to scale at $\boldsymbol{q}$. In particular, rewards are different from the rewards in Theorem 1.*

3. *If residual supply is perfectly elastic, so that the matrix $\boldsymbol{DP_{RS}}$ is zero, then rewards equal marginal products, as in Theorem 1.*

The proposition shows that a platform that maximizes the number of facilitated transplants does not set socially efficient rewards. Instead of setting rewards equal to marginal products, the platform subsidizes submissions that have elastic supply and are very productive. To see this, consider the simplest case, where $\boldsymbol{DP}_{RS}$ is a diagonal matrix (i.e. all cross-elasticities of residual supply are zero). Then, for each type i, the reward is marked down from marginal product according to an analogue of the Lerner index formula,

$$\frac{\frac{\partial f}{\partial q_i}\left(\boldsymbol{q}\right)-p_i}{p_i}=\frac{1}{\varepsilon_i^{RS}}-\frac{1}{\lambda}\cdot\frac{\frac{\partial f}{\partial q_i}\left(\boldsymbol{q}\right)}{p_i},$$

where ε_i^{RS} is the own-price residual supply elasticity and λ is the Lagrange multiplier on the constraint $f\left(\boldsymbol{q}\right)=\boldsymbol{P}_{RS}\left(\boldsymbol{q}\right)\cdot\boldsymbol{q}$. The expression shows the platform has incentives to skew the rewards: optimal markdowns are larger for submissions with low elasticities and submission categories that are less productive on the margin.

The proposition implies that competing, empire-building platforms exploit their market power and set rewards inefficiently. Additionally, the proposition implies that platforms set efficient rewards if the market is very competitive. Optimal rewards are close to marginal products if residual supply is very elastic, i.e. if ε_i^{RS} is close to infinity or, more generally, $\boldsymbol{DP}_{RS}$ is close to zero.

B Proofs

B.1 Preliminary results

We begin with a lemma showing that hospital supply can be aggregated, so that hospitals behave as a single hospital that takes the aggregate cost curve into account. This is similar to standard aggregation results in neoclassical firm theory (**?** p. 148).

Lemma B.1. *Fix a vector of rewards* $\boldsymbol{p}$*. Aggregate supply* $\boldsymbol{S}(\boldsymbol{p})$ *is the Minkowski sum of individual supply* $\boldsymbol{S}^h(\boldsymbol{p})$ *for all hospitals. Moreover, if there is a set of individual supply vectors* $(\boldsymbol{q}^h)_{h=1}^H$ *with each* $\boldsymbol{q}^h\in\boldsymbol{S}^h(\boldsymbol{p})$*, then*

$$\sum_{h=1}^{H}C^h(\boldsymbol{q}^h)=C\left(\sum_{h=1}^{H}\boldsymbol{q}^h\right). \tag{B1}$$

Proof. Note that

$$\begin{aligned}\max_{(q^h)_{h=1}^H}\left[\boldsymbol{p}\cdot\left(\sum_{h=1}^H\boldsymbol{q}^h\right)-\sum_{h=1}^H C^h(\boldsymbol{q}^h)\right] &= \max_{\boldsymbol{q}\in\mathbb{R}_+^I}\ \max_{(\boldsymbol{q}^h)_{h=1}^H}\ \left[\boldsymbol{p}\cdot\left(\sum_{h=1}^H\boldsymbol{q}^h\right)-\sum_{h=1}^H C^h(\boldsymbol{q}^h)\right], \\ &\qquad\text{s.t.}\quad \sum_{h=1}^H\boldsymbol{q}^h=\boldsymbol{q}, \\ &= \max_{\boldsymbol{q}\in\mathbb{R}_+^I}\ \left[\boldsymbol{p}\cdot\boldsymbol{q}-C(\boldsymbol{q})\right].\end{aligned} \tag{B2}$$

Consider $\boldsymbol{q}_0 = \sum_{h=1}^{H} \boldsymbol{q}_0^h$ with each $\boldsymbol{q}_0^h$ in $\boldsymbol{S}^h(\boldsymbol{p})$. By the optimality of aggregate cost, we have that

$$\boldsymbol{p} \cdot \boldsymbol{q}_0 - C(\boldsymbol{q}_0) \geq \boldsymbol{p} \cdot \left(\sum_{h=1}^{H} \boldsymbol{q}_0^h \right) - \sum_{h=1}^{H} C^h(\boldsymbol{q}_0^h). \tag{B3}$$

Optimality of the $\boldsymbol{q}_0^h$ implies that the right-hand side of this inequality attains the maximum in the left-hand side of equation (B2). Hence,

$$\boldsymbol{p} \cdot \boldsymbol{q}_0 - C(\boldsymbol{q}_0) \geq \max_{\boldsymbol{q} \in \mathbb{R}_+^I} \left[\boldsymbol{p} \cdot \boldsymbol{q} - C(\boldsymbol{q}) \right],$$

so that $\boldsymbol{q}_0$ is in $\boldsymbol{S}(\boldsymbol{p})$. Inequality (B3) holds as an equality, which implies (B1) as desired.

Conversely, consider $\boldsymbol{q}_0$ in $\boldsymbol{S}(\boldsymbol{p})$. Equation (B2) implies that

$$\boldsymbol{p} \cdot \boldsymbol{q}_0 - C(\boldsymbol{q}_0) = \max_{(\boldsymbol{q}^h)_{h=1}^{H}} \left[\boldsymbol{p} \cdot \left(\sum_{h=1}^{H} \boldsymbol{q}^h \right) - \sum_{h=1}^{H} C^h(\boldsymbol{q}^h) \right].$$

Let $(\boldsymbol{q}_0^h)$ be supply vectors that minimize the sum of costs conditional on total supply being $\boldsymbol{q}_0$. Then

$$\boldsymbol{p} \cdot \boldsymbol{q}_0 - C(\boldsymbol{q}_0) = \boldsymbol{p} \cdot \left(\sum_{h=1}^{H} \boldsymbol{q}_0^h \right) - \sum_{h=1}^{H} C^h(\boldsymbol{q}_0^h).$$

Therefore, each $\boldsymbol{q}_0^h$ is in $\boldsymbol{S}^h(\boldsymbol{q}^h)$.

□

Given the assumptions in the body of the paper, we can treat aggregate supply and aggregate inverse supply as functions rather than correspondences.

Lemma B.2. *The aggregate supply* $\boldsymbol{S}(\boldsymbol{p})$ *is a single-valued correspondence.*

Proof. Since we assume that the maximum of each hospital's objective is attained for some quantity for every vector of rewards, Lemma B.1 implies that $\boldsymbol{S}(\boldsymbol{p})$ is non-empty for any rewards vector $\boldsymbol{p}$. Because $C(\boldsymbol{q})$ is strictly convex, $\boldsymbol{S}(\boldsymbol{p})$ is the unique maximizer of the function $\boldsymbol{p} \cdot \boldsymbol{q} - C(\boldsymbol{q})$. □

Lemma B.3. *For any strictly positive* $\boldsymbol{q}$, $\boldsymbol{P}_S(\boldsymbol{q})$ *is single-valued and* $\boldsymbol{P}_S(\boldsymbol{q}) = \boldsymbol{\nabla C}(\boldsymbol{q})$.

Proof. Since $\boldsymbol{q}$ is interior and $C(\boldsymbol{q})$ is smooth and strictly convex, the first-order necessary conditions are also sufficient for a maximum of the aggregate supply program. □

B.2 Proof of the Main Theorem

Proof of Theorem 1. Let $(\boldsymbol{p}^*, (\boldsymbol{q}^{h*})_{h=1,\ldots,H})$ maximize hospital welfare subject to all hospitals choosing $\boldsymbol{q}^h$ in $S^h(\boldsymbol{p})$ given $\boldsymbol{p}$ and subject to allocating the number of transplants that are produced. Mathematically, the tuple $(\boldsymbol{p}^*, \boldsymbol{q}^*, (\boldsymbol{q}^{h*})_{h=1,\ldots,H})$ in $\mathbb{R}^I \times \mathbb{R}^I_+ \times \mathbb{R}^{IH}_+$ maximizes

$$f(\boldsymbol{q}) - \sum_{h=1}^{h} C^h(\boldsymbol{q}^h)$$

subject to

$$\boldsymbol{q} = \sum_{h=1}^{H} \boldsymbol{q}^h, \tag{B4}$$

to each h

$$\boldsymbol{q}^h \in \boldsymbol{S}^h(\boldsymbol{p}), \tag{B5}$$

and to

$$\boldsymbol{p} \cdot \boldsymbol{q} = f(\boldsymbol{q}). \tag{B6}$$

Lemma B.1 implies that this maximization problem is equivalent to finding a pair $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ in $\mathbb{R}^I \times \mathbb{R}^I_+$ that maximizes

$$f(\boldsymbol{q}) - C(\boldsymbol{q}) \tag{B7}$$

subject to

$$\boldsymbol{q} \in \boldsymbol{S}(\boldsymbol{p}), \tag{B8}$$

and to (B6).

By Lemma B.3, constraint (B8) is equivalent to $\boldsymbol{p} = \boldsymbol{\nabla C}(\boldsymbol{q})$. Thus, $\boldsymbol{q}^*$ maximizes (B7) in $\mathbb{R}^I_{++}$ subject to

$$\boldsymbol{\nabla C}(\boldsymbol{q}) \cdot \boldsymbol{q} = f(\boldsymbol{q}).$$

The production function and aggregate cost function are smooth, and $\mathbb{R}^I_{++}$ is an open set. Therefore, the Lagrange multiplier theorem implies that there exists λ such that $\boldsymbol{q}^*$ maximizes

$$f(\boldsymbol{q}) - C(\boldsymbol{q}) + \lambda \cdot \{f(\boldsymbol{q}) - \boldsymbol{\nabla C}(\boldsymbol{q}) \cdot \boldsymbol{q}\}.$$

Setting the derivative equal to zero, we have

$$\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{\nabla C}(\boldsymbol{q}) = \frac{\lambda}{1+\lambda} \boldsymbol{q'D^2C}(\boldsymbol{q}).$$

To solve for the Lagrange multiplier, we multiply by $\boldsymbol{q}$ on the right, and use the equality $\boldsymbol{\nabla C} \cdot \boldsymbol{q} = f$. We have

$$\boldsymbol{\nabla f}(\boldsymbol{q}) \cdot \boldsymbol{q} - f(\boldsymbol{q}) = \frac{\lambda}{1+\lambda} \boldsymbol{q'D^2C}(\boldsymbol{q})\boldsymbol{q}.$$

Therefore,

$$\frac{\boldsymbol{\nabla f}(\boldsymbol{q}) \cdot \boldsymbol{q} - f(\boldsymbol{q})}{\boldsymbol{q'D^2C}(\boldsymbol{q})\boldsymbol{q}} = \frac{\lambda}{1+\lambda}.$$

Substituting λ we get the final formula,

$$\boldsymbol{\nabla C}(\boldsymbol{q}) = \boldsymbol{\nabla f}(\boldsymbol{q}) - \left(\frac{\boldsymbol{\nabla f}(\boldsymbol{q}) \cdot \boldsymbol{q} - f(\boldsymbol{q})}{\boldsymbol{q}'\boldsymbol{D^2C}(\boldsymbol{q})\boldsymbol{q}}\right) \cdot (\boldsymbol{q}'\boldsymbol{D^2C}(\boldsymbol{q})).$$

Lemma B.3 allows us to replace $\boldsymbol{D}^2C(\boldsymbol{q})$ with $\boldsymbol{DP}_S(\boldsymbol{q})$.

Part 1 of the theorem follows by substituting $\boldsymbol{p} = \boldsymbol{\nabla C}(\boldsymbol{q})$. Part 2 follows because, with constant returns to scale, $\boldsymbol{\nabla f}(\boldsymbol{q}) \cdot \boldsymbol{q} - f(\boldsymbol{q}) = 0$ so that $\lambda = 0$. Part 3 follows because, with no agency problems, welfare and social welfare coincide.

□

B.3 Additional Proofs

Proof of Proposition A.1. In what follows, we are considering an asymptotic where $\boldsymbol{q}_0 \to \boldsymbol{q}^*$. Note that, by Lemma (B.1) and the fact that $\boldsymbol{q}^*$ is strictly positive, $\boldsymbol{p}^* = \boldsymbol{\nabla C}(\boldsymbol{q}^*)$. Further, since $\boldsymbol{q}_0$ is positive, we have that $\boldsymbol{p}_0 = \boldsymbol{\nabla C}(\boldsymbol{q}_0)$.

Part 1: Approximation of $\boldsymbol{q}^* - \boldsymbol{q}_0$:

Theorem (1) shows that

$$\boldsymbol{\nabla f}(\boldsymbol{q}^*) - \boldsymbol{\nabla C}(\boldsymbol{q}^*) = \boldsymbol{0}.$$

Taking a Taylor expansion of the left-hand side about $\boldsymbol{q_0}$, we have

$$\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{\nabla C}(\boldsymbol{q}_0) + (\boldsymbol{q}^* - \boldsymbol{q}_0)' \left[\boldsymbol{D^2f}(\boldsymbol{q}_0) - \boldsymbol{D^2C}(\boldsymbol{q}_0)\right] + \boldsymbol{\varepsilon_1} = \boldsymbol{0}, \tag{B9}$$

where $\|\boldsymbol{\varepsilon_1}\|$ is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|)$. Therefore,

$$\boldsymbol{q}^* - \boldsymbol{q}_0 = -\left[\boldsymbol{D^2f}(\boldsymbol{q}_0) - \boldsymbol{D^2C}(\boldsymbol{q}_0)\right]^{-1} \cdot \left[\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0\right]' + \boldsymbol{\varepsilon}_2, \tag{B10}$$

where the error term

$$\boldsymbol{\varepsilon}_2 = -\left[\boldsymbol{D^2f}(\boldsymbol{q}_0) - \boldsymbol{D^2C}(\boldsymbol{q}_0)\right]^{-1} \cdot \boldsymbol{\varepsilon_1'}$$

has a magnitude that is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|)$.

Part 2: First approximation of the deadweight loss:

The deadweight loss at $\boldsymbol{q}_0$ is given by

$$DWL = f(\boldsymbol{q}^*) - C(\boldsymbol{q}^*) - [f(\boldsymbol{q}_0) - C(\boldsymbol{q}_0)].$$

A second-order Taylor expansion of $f(\boldsymbol{q}^*) - C(\boldsymbol{q}^*)$ around $\boldsymbol{q}_0$ yields that

$$\begin{aligned} DWL = {} & [\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{\nabla C}(\boldsymbol{q}_0)] \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) \\ & + \frac{1}{2} \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0)' \cdot \left[\boldsymbol{D^2f}(\boldsymbol{q}_0) - \boldsymbol{D^2C}(\boldsymbol{q}_0)\right] \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) + \boldsymbol{\varepsilon}_3, \end{aligned}$$

where ε_3 is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$. We can substitute

$$(\boldsymbol{q}^* - \boldsymbol{q}_0)' \left[\boldsymbol{D^2 f}(\boldsymbol{q}_0) - \boldsymbol{D^2 C}(\boldsymbol{q}_0)\right] = - \left[\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0\right] - \boldsymbol{\varepsilon_1}$$

using equation (B9). This yields

$$DWL = \frac{1}{2} \cdot [\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0] \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) + \varepsilon_4,$$

where $\varepsilon_4 = -\boldsymbol{\varepsilon_1} \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) + \varepsilon_3$ is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$. This establishes the first approximation formula, as Lemma B.3 lets us replace $\boldsymbol{D^2 C}(\boldsymbol{q}_0)$ with $\boldsymbol{DP_S}(\boldsymbol{q}_0)$.

Part 3: Second approximation of the deadweight loss:

To establish the second approximation formula we substitute $\boldsymbol{q}^* - \boldsymbol{q}_0$ from equation (B10) into the first approximation to get

$$DWL = -\frac{1}{2} \cdot [\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0] \cdot \left[\boldsymbol{D^2 f}(\boldsymbol{q}_0) - \boldsymbol{D^2 C}(\boldsymbol{q}_0)\right]^{-1} \cdot [\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0]' \\ + \frac{1}{2} \cdot [\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0] \cdot \boldsymbol{\varepsilon_2} + \varepsilon_4, \quad \text{(B11)}$$

where ε_4 is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$ and $\|\boldsymbol{\varepsilon_2}\|$ is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|)$. Since equation (B9) shows that $\|\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0\|$ is $\mathcal{O}\left(\|\boldsymbol{q^*} - \boldsymbol{q_0}\|\right)$, and the product of a $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|)$ term and a $\mathcal{O}\left(\|\boldsymbol{q^*} - \boldsymbol{q_0}\|\right)$ term is $o\left(\|\boldsymbol{q^*} - \boldsymbol{q_0}\|^2\right)$, we conclude that $\varepsilon_5 = \frac{1}{2} \cdot [\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0] \cdot \boldsymbol{\varepsilon_2} + \varepsilon_4$ is $o\left(\|\boldsymbol{q^*} - \boldsymbol{q_0}\|^2\right)$. Lemma B.3 lets us replace $\boldsymbol{D^2 C}(\boldsymbol{q}_0)$ with $\boldsymbol{DP_S}(\boldsymbol{q}_0)$.

□

Proposition A.2. Let $(\boldsymbol{p}^*, (\boldsymbol{q}^{h*})_{h=1,\ldots,H})$ maximize social welfare subject to all hospitals choosing supply optimally given $\boldsymbol{p}$ (breaking ties by looking at social cost) and subject to not promising more transplants than are produced. Mathematically, the tuple $(\boldsymbol{p}^*, (\boldsymbol{q}^{h*})_{h=1,\ldots,H})$, in $\mathbb{R}^I \times \mathbb{R}_+^{IH}$ maximizes

$$f(\boldsymbol{q}) - \sum_{h=1}^{H} SC^h\left(\boldsymbol{q}^h\right)$$

subject to constraints (B4), (B5), and (B6). By Lemma B.1 , the assumption that the maximum is strictly positive, and the assumption that C is smooth and convex, the maximization problem is equivalent to finding $\boldsymbol{q}^*$ in $\mathbb{R}_{++}^I$ that maximizes

$$f(\boldsymbol{q}) - SC(\boldsymbol{q})$$

subject to

$$\boldsymbol{P}_S(\boldsymbol{q}) \cdot \boldsymbol{q} = f(\boldsymbol{q}).$$

By the Lagrange multiplier theorem, there exists a multiplier λ^{SW} such that, at the optimum,

$$\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q}) + \lambda^{SW}(\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{P}_S(\boldsymbol{q}) - \boldsymbol{q}'\boldsymbol{DP_S}(\boldsymbol{q})) = \boldsymbol{0}.$$

If we use the fact that $\boldsymbol{\nabla C}(\boldsymbol{q}) = \boldsymbol{P}_S(\boldsymbol{q})$, this becomes

$$\mathbf{0} = \boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{P}_S(\boldsymbol{q}) + (\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})) + \lambda^{SW}(\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{P}_S(\boldsymbol{q}) - \boldsymbol{q}'\boldsymbol{DP_S}(\boldsymbol{q})),$$

$$\boldsymbol{P}_S(\boldsymbol{q}) - \boldsymbol{\nabla f}(\boldsymbol{q}) = \frac{1}{1+\lambda^{SW}}(\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})) - \frac{\lambda^{SW}}{1+\lambda^{SW}}\boldsymbol{q}'\boldsymbol{DP_S}(\boldsymbol{q}),$$

The right-hand side of the last expression is the adjustment term $\boldsymbol{A}$. To obtain the formula for the Lagrange multiplier we multiply on the right by $\boldsymbol{q}^*$ and use the fact that $\boldsymbol{\nabla f}(\boldsymbol{q}^*)\cdot\boldsymbol{q}^* = \boldsymbol{p}^*\cdot\boldsymbol{q}^* = f(\boldsymbol{q}^*)$:

$$\lambda^{SW} = \frac{(\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q}))\,\boldsymbol{q}}{\boldsymbol{q}'\boldsymbol{DP_S}(\boldsymbol{q})\boldsymbol{q}}.$$

The second and third parts of the proposition follow from the formula for the Lagrange multiplier.

□

Proposition A.3. The platform chooses $\boldsymbol{q}$ in $\mathbb{R}^I_+$ to maximize

$$f(\boldsymbol{q})$$

subject to

$$f(\boldsymbol{q}) = \boldsymbol{P}_{RS}(\boldsymbol{q})\cdot\boldsymbol{q}.$$

Because the solution is interior, there exists a Lagrange multiplier λ such that

$$\boldsymbol{\nabla f}(\boldsymbol{q}) + \lambda(\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{P}_{RS}(\boldsymbol{q}) - \boldsymbol{q}'\boldsymbol{DP}_{RS}(\boldsymbol{q})) = \mathbf{0}.$$

Substituting that the optimal rewards $\boldsymbol{p} = \boldsymbol{P}_{RS}(\boldsymbol{q})$, we obtain

$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) + \frac{1}{\lambda}\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{q}'\boldsymbol{DP}_{RS}(\boldsymbol{q}).$$

To calculate the Lagrange multiplier, we right multiply by $\boldsymbol{q}$ and use $\boldsymbol{p}\cdot\boldsymbol{q} = \boldsymbol{\nabla f}(\boldsymbol{q})\cdot\boldsymbol{q} = f(\boldsymbol{q})$ to obtain

$$\lambda = \frac{f(\boldsymbol{q})}{\boldsymbol{q}'\boldsymbol{DP}_{RS}(\boldsymbol{q})\boldsymbol{q}}.$$

These two formulas imply part 1 of the proposition statement. The observation in part 2 follows directly from the formula in part 1.

□

C Data Appendix

This study used five main anonymized data sets: a database of all kidney exchange transplants done in the US from January 1, 2008 through December 4, 2014 (the OPTN transplant data), databases of all kidney exchange transplants organized by each of the three largest multi-hospital platforms in the US (the NKR, APD, and UNOS transplant data), and a database of all patient and donor registrations to the largest of those platforms (the NKR registration data).

C.1 Transplant data

In order to document the kidney exchange market, we merged the OPTN transplant data with the transplant data from NKR, APD, and UNOS. In what follows, we will describe these data and the merge procedure we used.

Obtaining the datasets: The OPTN provided us with a dataset on all transplants conducted in the US, known as the Standard Transplant Analysis and Research (STAR) dataset. The STAR dataset by itself lacks two key pieces of information: the transplant hospitals where the kidney was put into the patient and removed from the donor (which we use to determine whether a transplant is internal or external) and the unacceptable antigens for the patient (which we need to measure sensitization). These supplemental pieces of information are also available from the OPTN on request. Merging is done by using OPTN identifiers. The OPTN database contains records on 4377 kidney exchange transplants.

We obtained each of the platform datasets directly from the platform. The platform datasets contain records on 1400 kidney exchange transplants in total: 1193 from NKR, 100 from APD, and 107 from UNOS.

Dataset merge algorithm: In order to identify which transplants in the comprehensive OPTN database were organized by each of the three platforms, we matched records in the platform files to records in the OPTN file. However, because all datasets are anonymized, the merge must be done on the biological characteristics of each transplant's recipient and donor and logistical information on the transplant itself.

To do this, we use a match quality criterion. We start by matching all pairs of records that are uniquely the highest quality match (above a threshold) in the other dataset. However, things are less clear when a record has multiple highest quality matches. In these situations, we take the conservative approach of considering none of the matches genuine with one exception. Namely, we are willing to rule out a match because it involves a record that is part of an even better, unambiguous match.

To illustrate how this exception affects our matching procedure, consider a situation in which there are two records in the platform data, r_p and r'_p, that are the highest quality matches for record r_o in the OPTN data. Further, r_p and r'_p are of equally high quality for r_o. Without further information, a conservative match approach would avoid deciding whether r_o should be matched with r_p or r'_p. However, if there were another OPTN record r'_o such that the match r'_p–r'_o is unambiguous, then we determine that r'_p–r_o is not a match and instead match r_p–r_o. The remainder of this section formalizes this intuitive discussion.

Let the set of records in the platform data be R_p and the set in the OPTN data be R_o. A **match** between some platform record $r_p \in R_p$ and some OPTN record $r_o \in R_o$ can be represented by the ordered pair $(r_p, r_o) \in R_p \times R_o$. Some matches can be immediately rejected as unacceptable; we denote the **universe** of acceptable matches, $U \subseteq R_p \times R_o$. For our specific dataset and application, we deem a match unacceptable if the transplant dates are more than 31 days apart, if the recorded ages of the donor or the recipient are more than 10 years apart, or if either the receipient or the donor fails to match on both blood type or sex.[2]

We say that two matches **collide** if they share either an OPTN or a platform record. The disjunction is exclusive, that is, a match does not collide with itself. Then, a **merge** of platform records to OPTN records is a subset of the universe, $M \subseteq U$, that contains no two matches that collide. Intuitively, any OPTN record should be matched to at most one platform record, and vice-versa. Note that this definition allows for there to be no ordered pair in M that contains a given record from either database, which codifies the idea that the record remains unmatched.

The quality of a given match is represented by the **match rank function**, $\rho : U \mapsto \{1, \ldots, n\}$. The specific ranked criterion used for our application is defined as follows.

- Rank 1: The set of all acceptable matches where the donor and recipient each match exactly on blood type, sex, the hospital where the transplant was conducted, and all six major HLA alleles (two each on the HLA-A, HLA-B, and HLA-DR loci).
- Rank $n \in \{2, 3, 4, 5\}$: The set of all acceptable matches where the donor and recipient each match on at least $7 - n$ out of the six major HLA alleles.

We will treat the desirability of these ranks lexicographically, that is, we do not avoid a rank 1 match to enable any number of rank 2 matches, etc. In constructing a merge, we will often be faced with the choice of including one or the other of two colliding matches. We say that the match m' **weakly supercedes** the match m if m collides with m' and $\rho(m') \leq \rho(m)$. Similarly, m' **strictly supercedes** the match m if that rank inequality is strict.

We include a match in the merge if and only if any weakly superceding match is itself strictly superceded by a match in the merge. Formally, a merge $M \subseteq U$ is **rank stable** when $m \in M$ if and only if, for any $m' \in U$ that weakly supercedes m, there exists an $m'' \in M$ that strictly supercedes m'.[3,4]

[2] The theoretical properties of the merge algorithm described below do not depend on the specific criteria used to determine which matches are acceptable.

[3] In our definition, the requirement that M be a merge is redundant:

Claim C.1. Assume that a match $m \in U$ is in some set $M \subseteq U$ if and only if, for any $m' \in U$ that weakly supercedes m, there exists an $m'' \in M$ that strictly supercedes m'. Then, M must be a merge.

Proof. By way of contradiction, assume there are collisions in M. Let m' be the highest ranked match in M that collides with another match $m \in M$. Clearly, m' is not strictly superceded by another match in M, but it does weakly supercede m, a contradiction. □

[4] A few examples clarify how this criterion formalizes the intuition provided earlier in this section. First, consider a universe that contains only two matches of rank one; further, let those matches collide. More concretely, $\{m \in U \mid \rho(m) = 1\} = \left\{(r_p, r_o), (r_p', r_o)\right\}$.One could randomly choose either (r_p, r_o) or (r_p', r_o) to be in the merge; however, the choice would be incorrect half the time. A more conservative approach would keep both matches out of the merge by insisting that a match be in the merge if and only if it isn't weakly superceded by some other match.

Proposition C.1 in Section C.3 below shows that there can be only one merge that satisfies rank stability. Moreover, Algorithm 1 finds this unique merge, a result proved in Proposition C.2. This simple greedy algorithm puts all matches of rank 1 that aren't weakly superceded in the merge, removes from the universe all matches that collide with the matches in the merge, and then repeats the process for all matches of rank 2, etc.

```
Input: $U$ and $\rho(\cdot)$
Initialize $M \leftarrow \emptyset$ and $X_0 \leftarrow U$;
for $i \leftarrow 1$ to $n$ do
    $M_i \leftarrow \emptyset$;
    $Y_i \leftarrow \{m \in X_{i-1} \mid \rho(m) \leq i\}$;
    foreach $m \in Y_i$ do
        if $m$ doesn't collide with some other element of $Y_i$ then
            $M_i \leftarrow M_i + \{m\}$;
        end
    end
    $X_i \leftarrow X_{i-1}$;
    foreach $m \in X_{i-1}$ do
        if $m$ collides with some element of $M_i$ then
            $X_i \leftarrow X_i \setminus \{m\}$;
        end
    end
    $X_i \leftarrow X_i \setminus M_i$;
    $M \leftarrow M + M_i$;
end
Output: $M$
```

Algorithm 1: After each iteration i of the main loop, M_i contains all rank-i matches in the unique rank stable merge, and X_i contains all elements of the universe U that don't collide with a match in any M_j for $j \in \{1, \ldots i\}$.

Quality of the merge We now describe the output of the merge to show that the algorithm performs well in practice. The percentage of platform records matched to an OPTN record

This approach, however, is inadequate. Consider a universe that has a unique rank-one match and exactly two rank-two matches. Further, assume that the rank-one match collides with only one of the rank-two matches and that the rank-two matches collide with each other. More concretely, $\{m \in U \mid \rho(m) = 1\} = \{(r'_p, r'_o)\}$, $\{m \in U \mid \rho(m) = 2\} = \{(r_p, r_o), (r'_p, r_o)\}$. Intuitively, (r'_p, r'_o) should be in the merge, and indeed, it isn't weakly superceded. From this, we should logically conclude that (r'_p, r_o) isn't genuine, and hence include (r_p, r_o) in the merge. However, since (r_p, r_o) and (r'_p, r_o) weakly supercede each other, neither could be in the merge under the previous paragraph's solution concept. What that concept misses is that while (r'_p, r_o) weakly supercedes (r_p, r_o), it is itself strictly superceded by (r'_p, r'_o), which should definitely be in the merge.

Table C1: Agreement Between Matched Records in the Transplant Merge

Platform	Age within 5 years	Transplant date within 1 day	5 or more HLA alleles match	Blood type and sex match	Transplant hospital matches	At most one criterion is violated	No criterion is violated
NKR	97.8%	97.4%	95.0%	95.1%	97.2%	97.6%	90.8%
APD	92.8%	95.9%	93.8%	96.9%	95.9%	96.9%	87.6%
UNOS	87.1%	99.0%	96.0%	99.0%	99.0%	99.0%	83.2%
Overall	96.6%	97.4%	95.0%	95.5%	97.3%	97.6%	90.0%

was 94% overall (94% for NKR, 97% for APD, and 94% for UNOS). Moreover, the matches seem to be high quality: Table C1 reports the percentage of matches that meet various criteria. At least one criteria violated for only about 9% of the matches at the NKR.

In order to assess whether measurement error may be significant, we restricted attention to a stricter match criterion. Specifically, we use the same procedure but exclude any match where i) either donor or recipient fail to match within 5 years on age, ii) the transplant date fails to match within one day, or iii) the transplant hospital fails to match. This leaves us with an OPTN match for 90% of NKR records. Of those matches, 94.3% fail to violate any of conditions in Table B1. The results in Section 3 of the paper are nearly identical, suggesting that potential error in the matching procedure does not drive our results. These results are available upon request.

To further confirm the quality of the match procedure, we use the fact that the UNOS and OPTN databases share common identifiers, since UNOS is a contractor for the OPTN. This allows us to see the true merge for the UNOS data (but not for the APD or NKR data, since those datasets are anonymized to different identifiers). Comparing true matches to the matches selected by our algorithm, we find that only one UNOS record was incorrectly matched. That is, our algorithm chose the correct match in the OPTN dataset for 99% of the UNOS records it matched. Even in this case, where the exact matches can be verified, at least one criterion is violated for about 17% of the cases. The similarity of this statistic to those for the NKR transplants provides confidence in the fact that record-keeping differences, rather than errors is the likely cause. Overall, these results give confidence in our merge algorithm.

C.2 Registration data

In this subsection, we describe how the list of registrations to the NKR was assembled. The NKR provided us with snapshot files of the patient and donor pool between April 1, 2012 and December 4, 2014. These files are typically daily snapshots except for some missing periods, each of which is up to a month in length. Each snapshot corresponds to a different date and includes basic medical records for each patient and donor in the pool, their listing dates, the related patient for each donor (if any), and whether a patient is unpaired. From these snapshots we recover patient and donors departures, which may be due to being transplanted or other undocumented reasons. A small number of patients and donors depart without a transplant during the period of a missing snapshot; for these, we use bounds on the departure time using the two closest available snapshots, before and after the real departure date. These snapshots also include each patient's set of donors that may not be matched (i.e. are blocked) despite being virtually compatible. Some of these blocked donors are due to patient preference (not to match with these donors) and others are due to match failures.

C.3 Proofs concerning Algorithm 1 and stable merges

Proposition C.1. *Given any two rank stable merges M and M', it must be that $M = M'$.*

Proof. We will prove the proposition by induction on rank.

Base case: We show that M and M' contain the same set of rank-one matches. By way of contradiction, assume there exists a rank-one match m that is in M but not in M'. For m not to be in M', it must be weakly superceded by some other match m' that is not strictly superceded in M'. Then, for m to be in M, there must be some match $m'' \in M$ that strictly supercedes m'. But, it is impossible to strictly supercede a rank-one match, a contradiction.

Induction step: We show that if M and M' contain the same set of rank-k or better matches, then they must also contain the same set of rank-$(k+1)$ matches. By way of contradiction, assume there exists a rank-$(k+1)$ match m that is in M but not in M'. For m not to be in M', it must be weakly superceded by some other match m' that is not strictly superceded in M'. But then for m to be in M, there must be some match $m'' \in M$ that strictly supercedes m'. Since m'' must be of strictly higher rank than m, by assumption, $m'' \in M'$ as well. But then, m'' strictly supercedes m' in M', a contradiction.

□

Proposition C.2. *Algorithm 1 produces the unique rank stable merge.*

Proof. Since we already established uniqueness in Proposition C.1, it suffices to establish that after the algorithm runs, M_i equals the set of rank-i matches that are found in the unique rank stable merge M. We will show this by induction on rank.

Base case: We must prove that after the algorithm runs, M_1 equals the set of rank-one matches that are found in M. There are two ways to contradict this statement: either there is a rank-one match in M that the algorithm doesn't place in M_1, or there is a match in M_1 that isn't one of the rank-one matches in M.

Let's first consider the former. By way of contradiction assume $m \in M \setminus M_1$ and $\rho(m) = 1$. Clearly, the algorithm must place m in Y_1. For it to also not place m into M_1, there must be some other match $m' \in Y_1$ with which m collides. But then m' would weakly supercede m. And since m' is a rank one match, it can't be strictly superceded. Hence $m \notin M$, a contradiction.

Now, let's consider the latter. By way of contradiction, assume some $m \in M_1 \setminus M$. Looking at the algorithm, it is clear that $\rho(m) = 1$. Then, for m to not be in M, it must be that m is weakly superceded by some m'. But since m' and m must collide, looking to the algorithm, it is clear that m can't be in M_1, a contradiction.

Induction step: We must prove that after the algorithm runs, if M_j equals the set of rank-j matches that are found in M, for all $j \in \{1, \ldots, k\}$, then M_{k+1} equals the set of rank-$(k+1)$ matches that are found in M. As before, there are two ways to contradict this statement: either there is a rank-$(k+1)$ match in the unique merge M that the algorithm doesn't place in M_{k+1}, or there is a match in M_{k+1} that isn't one of the rank-$(k+1)$ matches in M.

Let's first consider the former. By way of contradiction assume $m \in M \setminus M_{k+1}$ and $\rho(m) = k+1$. If m collides with any match in M of strictly higher rank, by the induction hypothesis, this match must strictly supercede m, which contradicts that $m \in M$. Hence, the algorithm must place m in Y_{k+1}. For it to also not place m into M_{k+1}, there must be some other match $m' \in Y_{k+1}$ with which m collides. But then m' would weakly supercede m, which means that for m to be in M, there must be some other match $m'' \in M$ that strictly supercedes m'. However, by the induction hypothesis, m'' would have been added to the match in a previous round of the algorithm, which means that X_{k+1} can't contain any match that collides with m''. Hence, Y_{k+1} cannot contain m', a contradiction.

Now, let's consider the latter. By way of contradiction, assume some $m \in M_{k+1} \setminus M$. Now, if $m \notin M$, there must exist some m' that weakly supercedes m without itself being strictly superceded by any match in M. But, for $m \in M_{k+1}$ to hold, any weakly superceding m' must not be in Y_{k+1}. One way for this to happen is for $m' \in M_\ell$ for some $\ell \leq k$. But then, by the induction hypothesis, $\rho(m') = \ell$, which means that $m \notin X_j$ for any $j \geq \ell$. This contradicts the assumption that $m \in M_{k+1}$.

The other way to keep m' from being in Y_{k+1} is that in some round $\ell \leq k$, m' is in $X_{\ell-1}$ and collides with some $m'' \in M_\ell$. By the induction hypothesis, $\rho(m'') = \ell$. Looking to the algorithm, for m'' to be added to the merge while m' is removed, we need $\rho(m') > \ell$. But then, $m'' \in M$ strictly supercedes m', a contradiction.

□

D Simulation Details

We now provide details on the procedure used in Section 5.1.

D.1 Matching Offers

D.1.1 The matching algorithm and examples

We describe solve the linear programming problem subject to cycle length and chain constraints. Because it is computationally burdensome to compute all cycles in a pool with many patients and donors, we first solve the relaxed problem by ignoring the constraint that cycles cannot involve more than three transplants. Specifically, we solve:

$$\begin{array}{lll}
\max\limits_{x_{jk}\in\{0,1\}} & \sum\limits_{j\in\mathcal{P}\cup\mathcal{U}}\sum\limits_{k\in\mathcal{P}\cup\mathcal{A}} c_{jk}\, w_{jk}\, x_{jk} & \\
\text{s.t.} & x_{jk} - \sum\limits_{\ell} x_{k\ell} = 0 & \text{for all } k\in\mathcal{P} \\
 & \left.\begin{array}{l} \sum\limits_{j} x_{jk} \le 1 \\ \sum\limits_{k} x_{jk} \le 1 \end{array}\right\} & \text{for all } k
\end{array}$$

where $\mathcal{A}$ is the set of altruistic donors, $\mathcal{P}$ is the set of pairs, $\mathcal{U}$ is the set of unpaired recipients, $x_{jk} = 1$ denotes a proposed transplant from the donor in $k \in \mathcal{P}\cup\mathcal{A}$ to the patient in $j \in \mathcal{P}\cup\mathcal{U}$, w_{jk} is the weight described in Section D.1.2 below, and $c_{jk} = 1$ if a transplant from k to j is allowed and 0 otherwise. The first constraint ensures that a donor who is part of a pair is only asked to donate an organ if the intended recipient has been proposed a transplant. The second and third constraints ensure that no donor or recipient is involved in more than one transplant.

If the solution to the problem does not involve any long cycles, i.e. there do not exist $j_1, \ldots, j_n \in \mathcal{P}$ such that $x^*_{j_{k+1}j_k} = 1$ for $k \in \{1, \ldots, K-1\}$, $x^*_{j_1 j_K} = 1$, and $K \geq 4$, then it must be that x^* is optimal, given the no long-cycle constraints, and is our desired solution. In 87.3% of simulation days the solution to this relaxed problem yields a feasible match without any further cycle restrictions.

If the solution to this problem contains at least one long cycle, then we proceed as follows. We begin by following the algorithm in **?**. The algorithm includes a constraint that explicitly prohibits all long cycles in x^*, i.e. for each sequence $j_1, \ldots, j_K \in \mathcal{P}$ such that for $k \in \{1, \ldots, K-1\}$, $x^*_{j_{k+1}j_k} = 1$, $x^*_{j_1 j_K} = 1$, and $K \geq 4$, we include a constraint in the problem above to ensure that $x^*_{j_1 j_K} \prod_{k=1}^{K-1} x^*_{j_{k+1}j_k} = 0$. If the solution to the modified problem also contains long cycles, we modify the problem again to prohibit those cycles. We iterate this procedure up to ten times. This procedure yields a feasible solution in about 50% of the remaining cases (about 6.4% of all cases) with an average of approximately 3.9 iterations.

If the algorithm above does not yield a feasible solution even after 10 repetitions, we proceed to the next phase in which we use **?**'s algorithm to compute cycles and explicitly add constraints that prohibit long cycles. This algorithm searches the compatibility graph induced by c to calculate cycles. We enumerate and add a constraint to our program to prohibit any long cycles we have found. The number of cycles is usually small, but sometimes is very large. Therefore, we search for cycles with a time-out of one second. We find a solution to the problem with these additional constraints and terminate our algorithm if the solution is feasible. This procedure is repeated once more, if necessary. At the end of this phase, we are able to find an optimal solution to the full problem in about 99.9% of the simulation days.

For the remaining 0.1% of days in the simulation, our matching algorithm still ends up with long cycles. Whenever this is the case, we attempt to find a solution with the following, alternative algorithm:

$$\begin{array}{lll}
\max\limits_{y_\ell \in \{0,1\}} & \sum\limits_{\ell \in C} w_\ell \, y_\ell & \\
\text{s.t.} & w_\ell = \sum\limits_{i \in \ell}^{|\ell|-1} w_{i,i+1} + w_{|\ell|,1} & \text{if } \ell \text{ is a cycle,} \\
& w_\ell = \sum\limits_{i \in \ell}^{|\ell|} w_{i,i+1} & \text{if } \ell \text{ is a chain,} \\
& \ell \cap \ell' = \emptyset & \text{if } y_\ell \, y_{\ell'} = 1, \\
& c_{i,i+1} = 1,\ c_{|\ell|,1} = 1, \forall i \in \ell & \text{if } \ell \text{ is a cycle,} \\
& c_{i,i+1} = 1, \forall i \in \ell & \text{if } \ell \text{ is a chain.}
\end{array}$$

where C is the set of feasible chains and cycles, $y_\ell = 1$ denotes implementing chain or cycle ℓ, w_ℓ is the sum of the weights for each transplant in the chain or cycle indexed by ℓ, and $c_{jk} = 1$ if a transplant from k to j is allowed and 0 otherwise. Denote the number of pairs/altruistic donors/unpaired recipients in a chain or cycle ℓ by $|\ell|$. The first set of constraints defines the total weight w_ℓ for each chain or cycle; the second set ensures no donor or recipient is involved in more than one transplant; the third set ensures that all simultaneously proposed chains and cycles do not overlap. Because it is computationally burdensome to compute all chains in a pool with many patients and donors, we solve the problem chains with up to length of 5. We abandon the match on that "simulation day" if we still cannot find a solution, and all patients and donors are returned to the kidney exchange pool to wait for the next day for the transplant offer. Even in these rare cases, we can find an optimal match within 1.0 days on average.

Figure D3 illustrates a few kidney exchange pools. The left panel shows compatibility as captured by c, feasible transplants, and the optimal match. Blue dots denote patient-donor pairs, and magenta dots denote altruistic donors. We ignore unpaired patients in this illustration for simplicity. A blue arrow depicts a feasible transplant with the origin of the arrow denoting the donor. Red and green arrows depict cycles and chains, respectively, in the optimal match. Given feasible transplants on left, our match algorithm offers the one on the right. The figure shows there may be several feasible transplants, and in these cases, the optimal match may be relatively easy to determine. Figure D4, on the other hand, illustrates a relatively hard-to-match problem where the optimal match is relatively more difficult to determine.

D.1.2 Weights

We attempt to closely match the weights, w_{ij}, on a NKR transplant between patient i and donor j. These weights are designed to favor patients who are highly sensitized, in other

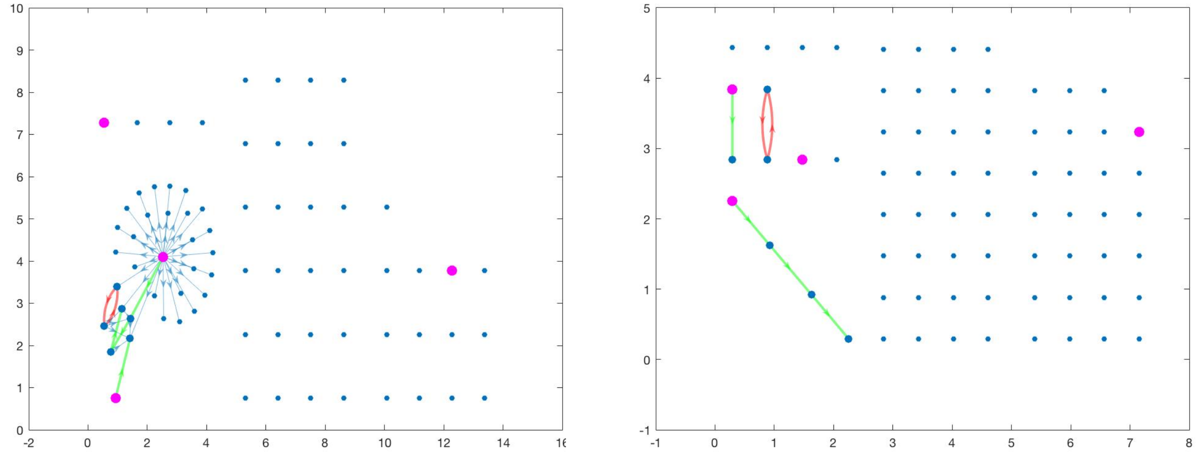


Figure D3: Optimal matches for two simple pools

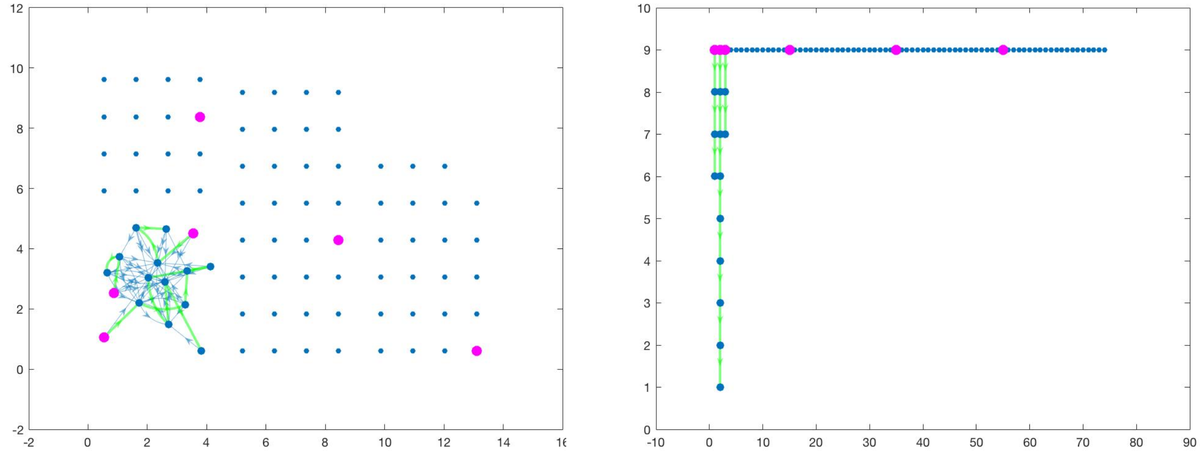


Figure D4: Optimal matches for two more complicated pools

words, who are harder to transplant. To define weights, w, NKR first defines a matching power for each submission. Each patient has a Patient Match Power (PMP), a number between 0 and 1, that is a fraction of compatible donors in the NKR pool for that patient. A low PMP for patient i implies that few donors are compatible with patient i. Similarly, the Donor Match Power (DMP) is defined as the fraction of patients in the NKR pool with whom that donor is compatible. Because these quantities and the pool used by the NKR to compute these match powers are not directly observed in our dataset, we calculate them using our sample.

Given these characteristics, NKR calculates a scaled measure of how likely a feasible transplant can occur between i and j, $WNKR_{ij}$. Specifically,

$$WNKR_{ij} = PMP \times DMP \times 10,200.$$

A low $WNKR_{ij}$ correlates with a transplant between i and j being unlikely. It is important to note that the magnitude of $WNKR_{ij}$ is not related to the success of a transplant if it

$WNKR_{ij}$ interval	w_{ij}
$WNKR_{ij} > 70$	1
$25 < WNKR_{ij} \leq 70$	1.01
$5 < WNKR_{ij} \leq 25$	1.2
$WNKR_{ij} \leq 5$	1.5

Table D2: Weights used by the NKR

turns out to be feasible. These weights therefore accord higher priority to hard-to-match patients and donors. Using $WNKR_{ij}$, NKR assigns the weights w_{ij} as follows:

Because these weights are less than 2, they typically maximize the total number of transplants. However, these weights may sometimes result in two transplants, each with weight 1.5, instead of three transplants with weight 1 each.

D.2 Arrival and Departure

D.2.1 Arrival process

We assume the daily number of submissions in the NKR is given by a Poisson distribution with parameter λ, where λ represents the mean daily arrival rate for NKR. We estimate that parameter to be $\lambda = 1.975$. In each period, our simulations draw a number, say n_t, from this distribution. Then we draw n_t submissions with replacement from the entire pool that ever registered in the NKR during the April 2012 to June 2014 sample period.

Figure D5 shows the fit of the arrival per day distributions of NKR and Poisson. Notice that NKR's distribution has more 0 arrivals per days than the poisson distribution. This mass point is explained by weekends, which appear to have a much lower arrival rate.[5] Figure D6 shows the arrival per day distribution of NKR for weekdays and our estimated poisson distribution, which shows a better fit.

D.2.2 Departure process

To model departures, we estimate an interval censored hazard model to calculate the rate at which patients and/or donors depart the NKR without a transplant. Specifically, let t_i^a, t_i^*, and t_i^d be the (latent) arrival, transplant, and departure dates for an unpaired patient, donor, or patient-donor pair i. Our dataset records t_{0i} and t_i^* if i was transplanted. Further, if i is transplanted, then we know that $t_i^d > t_i^*$. If i is not transplanted, then in most cases we observe t_i^a, but in some cases, we only know that t_i^a belongs to an interval $\left[t_i^{a-}, t_i^{a+}\right]$ (typically

[5] Only 40 arrivals in weekends over the course of 140 weeks.

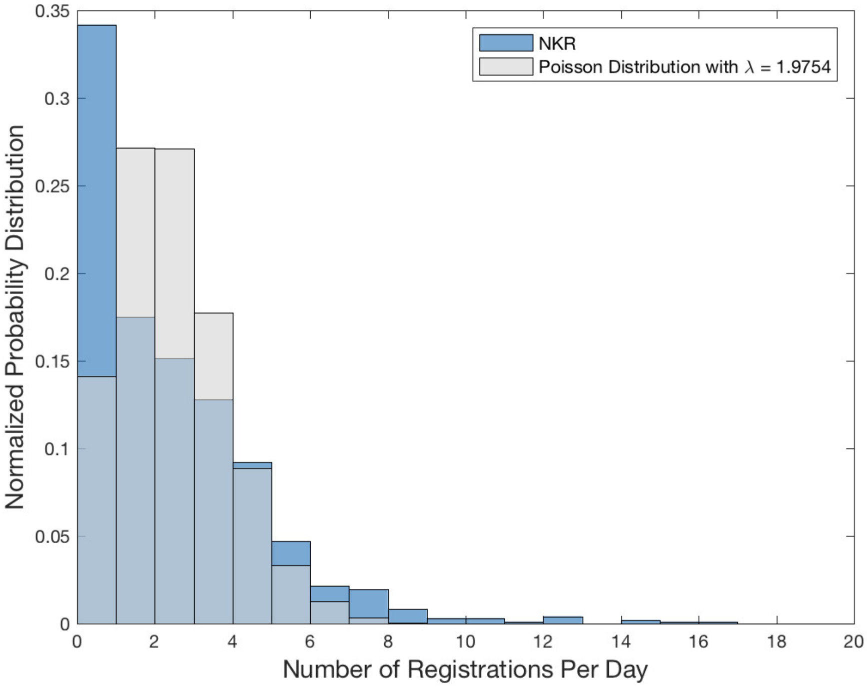


Figure D5: Distribution of NKR and Poisson Number of Submissions Per Day

within a week). If i departed without a transplant, we observe t_i^deither exactly or up to a small interval. If i remains in the NKR at the end of our sample, then we know that $t_i^d > T$. Using these observations, we can construct bounds on the duration τ_i that each unit i remains in the NKR without a transplant.

With these observations, we estimate the exponential hazards model. The model is characterized by a survival function

$$S_i(\tau) = \exp(-\lambda_i t),$$

where we use the parametric form

$$\lambda_i = \alpha + z_i \beta.$$

The likelihood of the model for the interval censored survival data is straightforward to derive, and estimation via `intcens` in STATA is straightforward.

Table D3 presents the estimates. Note that the hypothesis tests in a hazard model are reported relative to 1, which implies no effect. As can be seen, patients with blood types that are easier to match or who are paired with easier-to-match donors have a higher departure rate. This finding is consistent with patients and patient-donor pairs departing in response to transplantation opportunities elsewhere, either through direct donation, deceased donor transplants, or live-donor exchanges outside the NKR.

Table D3: Departure Hazard Rate Estimates

	(1) Patient-Donor Pairs	(2) Unpaired Patients	(3) Altruistic Donors
Patient Matching Power	1.824***	16.35***	
	(0.244)	(2.446)	
Donor Matching Power	0.0699		0.000167
	(0.137)		(0.00126)
Patient Age	0.994*	1.002	
	(0.00349)	(0.00383)	
Donor Age	1.008*		1.011
	(0.00442)		(0.0140)
AB Blood-type Patient	2.465***	1.557*	
	(0.698)	(0.390)	
A Blood-type Patient	1.184	1.294	
	(0.160)	(0.265)	
B Blood-type Patient	1.077	0.635*	
	(0.172)	(0.158)	
AB Blood-type Donor	0.584**		1.249
	(0.150)		(1.357)
A Blood-type Donor	0.667***		1.562
	(0.0832)		(0.654)
B Blood-type Donor	0.608***		0.764
	(0.0957)		(0.520)
Constant	0.00578***	0.000892***	0.0656
	(0.00565)	(0.000203)	(0.234)
Observations	1,264	498	164

Note: Interval censored exponential hazard model. Patient (Donor) Match Power is the fraction of donors (patient) in the NKR pool over the course of a sample a given patient (donor) is compatible with. Sample restricted to patients and donors that registered after April 2012.

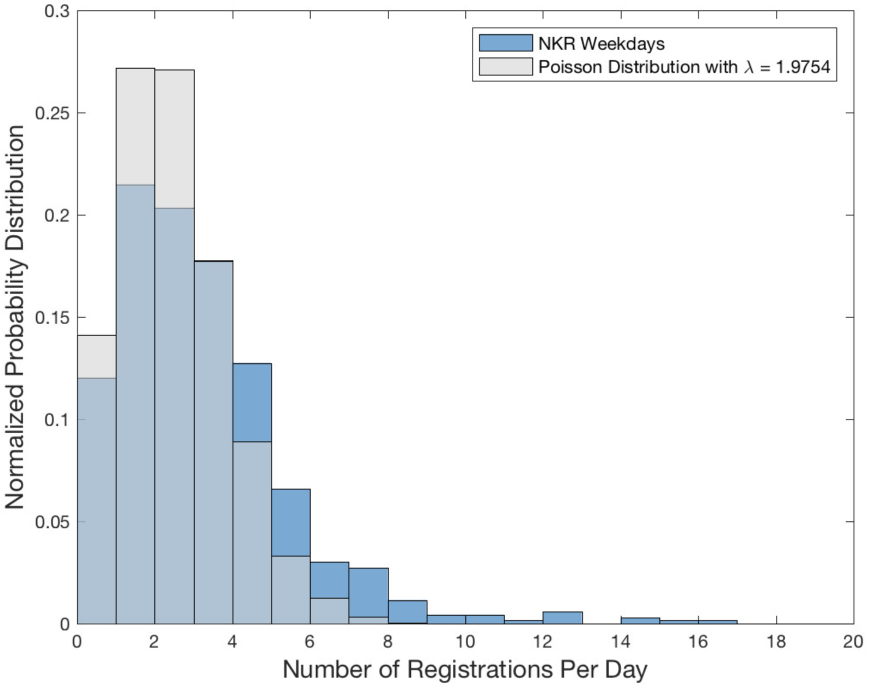


Figure D6: Distribution of NKR and Poisson Number of Submissions Per Day for Weekdays

D.3 Compatibility and acceptance

To calculate whether donor j is compatible with patient i, we use the blood types of the patient and donor, the tissue type of the donor, and the list of unacceptable antigens listed by the patient. There are three additional ways in which a transplant between a patient and a donor can be prohibited.

First, upon registration, each patient can declare criteria for excluding donors based on a variety of characteristics. These include thresholds for the maximum donor age and minimum donor weight that are acceptable. These criteria are recorded in our dataset.

Second, upon arrival a patient can list as unacceptable any number of specific donors who were in the NKR pool at the time. This rejection can be done for any reason, including known pathologies. Patients can also exclude donors later, but according to our understanding, the practice is most common during registration. Our dataset includes the anonymized identifiers for these excluded donors.

Third, when a transplant is proposed, a patient may refuse the specific donor. A first phase of refusals is at the patient's discretion (with advise from his/her surgeon). If a patient chooses to proceed after the first phase, a final tissue-type compatibility test is conducted. We refer to this as the second phase.

In our simulations, we initialize $c_{ij} = 1$ if j is compatible with i and if j is not excluded by i. Otherwise, we set $c_{ij} = 0$. If j is offered to i during the simulation and a transplant is ruled out during either the first or second phases in the third type of exclusion, then we set $c_{ij} = 0$ for future "simulation days."

D.4 Burn-in and calculating standard errors

We may start our simulations from any initial state because the effect of the chosen initial state fades over time. A convenient choice is to pick an initial pool with no unpaired patients, altruistic donors, or patient-donor pairs. Although the initial pool does not affect long-run averages with enough simulations, it is advisable to discard or burn-in a portion of the initial chain in order to improve the estimates' precision. We simulate a chain with 100,000 to 400,000 days after a burn-in of 2,000 days. We used longer chains for larger scales as convergence is faster for small market sizes. We checked to ensure that the potential scale reduction factor for the number of transplants per day that is close to 1. Numbers close to shis value suggest that the chain has converged. Marginal products are computed using finite differences in arrival rates.

Our simulations produce a series $y_1, \ldots, y_T$ of the transplants that occur on each day after an initial burn period. We estimate $f(\boldsymbol{q})$ as the sample mean of the y_t and calculate the standard errors of this estimate using the non-overlapping batch means estimator by following Chapter 12 in **?**. The method divides the time series of y_t into batches, calculates the sample mean in each of those batches, and uses the variability in sample means to estimate the standard error of $f(\boldsymbol{q})$. We use the commonly recommended batch size of approximately $\sqrt{T}$. The procedure is a simple and popular method that accounts for autocorrelation of the y_t.

D.5 Calibration

Our simulation procedure is tailored to match the procedures and practices used by the NKR. In most cases, the data or institutional knowledge directly tell us the parameters; e.g., the weights w_{ij} are chosen to match NKR's practices. However, there are a few aspects of the real-world procedures and outcomes in the NKR on which we don't directly have information. We model and parametrize these aspects in our simulation model and calibrate them to match the realized number of matches in the NKR.

There are two main sets of parameters we need to calibrate. First, we do not have direct data on the frictions of translating proposed transplants into surgeries. As mentioned in Section D.3, the various acceptance phases may result in some transplants not being consummated. Each phase introduces a time-lag between transplants being proposed and finalized as well as the chance of a match being aborted. Roughly speaking, these frictions reduce the number of transplants facilitated by the NKR.

We parametrize these phases and calibrate the parameters to best fit the observed number of transplants by patient type. To do so, we simulate outcomes predicted by our model for various lengths (number of days) and various failure probabilities for both phases. The first phase parameters can be interpreted as controlling the frictions in the system because proposed matches are refused, whereas the second phase parameters govern the frictions due to biological compatibility tests.

Second, as mentioned earlier, when chains are aborted because of a refusal, NKR usually tries to use the donor, called the bridge donor, of the last transplanted patient for a new chain. However, the exchange prefers not to wait too long to start a chain with this donor. If a new chain cannot be found, the donor is offered to a patient without a related donor. Unfortunately, we do not know of a consistent policy rule followed by the NKR. We therefore also experimented with the number of days the NKR tries to match a bridge donor.

In summary, we calibrated five parameters: (i) the number of days a bridge donor can initiate a new chain, (ii) the number of days for consent and the probability of consent for the two phases of match acceptance, and (iii) the number of days taken for testing and the probability of failing the tissue type biological test.

These parameters parsimoniously set the nature of frictions and policies in the NKR. However, the limited heterogeneity, especially over time, is clearly a simplification. Most likely, NKR policies and their ability to translate proposed matches into transplants evolves over time and includes some ad hoc modifications to their basic procedure.

For our calibrations, we conduct our simulations by setting an initial market with the patients and donors who were present in the NKR on 1 April, 2012. This is the date from which we have clear registration data. Then, for each parameter set, we run 100 simulations until December 2014, the last date of the available data.

We calibrated our parameters to match two sets of moments. The first set is the transplantation probabilities, and days in the exchange for transplanted patients and donors. Each of these three moments are split by whether we are considering an unpaired patient, a patient-donor pair, or an altruistic donor. Therefore, there are are six such moments in total. The observed quantities and the simulated quantities at the calibrated parameters are summarized in Table D4, with even finer partitions. The second set are annual statistics on the number of patients and donors waiting for a transplant in the NKR. Each of these three statistics is calculated by type (patient-donor pair, altruistic donor or unpaired patient) and by year (2012, 2013, 2014). There are a total of nine such moments. The trends in the data and the simulations from our preferred parameters are detailed in Figures D8 - D10, with the overall trends depicted in Figure D7. The dashed lines depict data, darker lines are the mean of 100 simulations, and dashed lines are 95% confidence interval from the simulations. Although the averages are not as well matched in Figures D10 and D9, the observed quantities are within the 95% confidence intervals for the model. The final set of moments are the fraction of transplants in two-way cycles, three-way cycles and the average chain length. These statistics are summarized in Table D6.

These three sets of moments capture the transplants produced by the NKR by type, the types of exchanges arranged and the probability that agents are matched. These are key quantities used in our theoretical and empirical exercises as they correspond to average products and current incentives. For example, transplant probabilities directly translate to average products. Other moments combine to yield information on marginal productions. For example, chain length captures the average number of transplants produced by each altruistic donor and fraction of transplants in two-way and three-way cycles indicate the alternative avenues through which patients and donors in the chains could be transplanted. Chain length also provides important information on frictions and bridge donor wait time policy. Finally, wait time statistics are useful for appropriately replicating the flow rates in our simulations, which aligns well with the preferred interpretation of the quantities in our theoretical model. As shown in the tables and figures referenced above, our highly parsimonious model does well in matching a large number of moments.

To arrive at these calibrated parameters and assess their fit relative to others, we worked with a statistical loss function. At each test parameter, the difference between the moments observed in the data and calculated from the simulation were stacked. Formally, we calculated the vector $\hat{m} - m_S(\theta)$, where $\hat{m}$ denotes the statistics observed in the dataset and $m_S(\theta)$ denotes the simulated quantities. We then constructed a loss function $[\hat{m} - m_S(\theta)]' W [\hat{m} - m_S(\theta)]$ using a weight matrix W to summarize a measure of fit. The weight matrix was set to the inverse covariance of the moments estimated using 10,000 bootstrap samples from the NKR dataset. To construct this matrix, we sampled unpaired patients, altruistic donors and patient-donor pairs from the NKR dataset. Transplant probabilities and days in the exchange by transplanted status were computed for each sample using the observed outcome for each draw. The observed dates of arrival and departure for the bootstrap sample were used to compute the second set of moments on stocks and transplants by year. Similarly, fraction of transplants in two-way and three-way cycles were computed based on whether or not the patient is transplanted and the observed mode of transplant in the realized data. Average chain lengths for the bootstrap sample were constructed using the realized chain length in the data and weighting each observation by $1/l_i$ where l_i is the observed chain length for agent i in the data. This ensures that all chains are equally weighted, as done in the original computation.[6]

While this loss function now makes it possible in principle to find the value of the parameters θ that minimizes it, yielding a simulated minimum distance estimator, it is computationally infeasible to do so because it takes more than a week on our cluster to compute $m_S(\theta)$ for each θ. Instead, we first experimented with short simulations with few iterations to zoom in on a set of parameters. We then conducted complete simulations for five variants. The parameters with the best fit were 14 days of waiting and 75% success rate for each of the two phases, and 30 days of wait before ending the chain with a potential bridge donor.

In addition to the preferred parameters, we computed complete simulations for six other parameters that vary the frictions in the market:

1. Lower Friction: Increases the success rate from the baseline of 75% to 80% in each of the two phases.

[6] Strictly speaking, this bootstrap procedure imposes independence assumptions that may not be feasible in the true data generating process. For example, the bootstrap sample may contain a set of transplants that may not be possible to arrange based on the draws of all patients and donors. A parametric bootstrap that uses a weight matrix as a function of the parameter θ also finds that the objective function is minimized at our preferred parameters. We view our bootstrap calculation primarily as an approximation to account for moments that may be highly correlated and adjust for the scale of various moments.

2. Higher Friction: Reduces the success rate from 75% to 70%.

3. Shorter Wait Time: Uses the baseline success rates, but sets the delay in each phase to 7 days.

4. Longer Wait Time: Uses the baseline success rates, but sets the delay in each phase to 21 days.

5. Shorter Bridge Donor Wait Time: Uses the baseline friction parameters, but increases the bridge donor wait time to 7 days.

6. Longer Bridge Donor Wait Time: Uses the baseline friction parameters, but reduces the bridge donor wait time to 60 days.

Tables D5 and D6 show the first and third set of moments produced by these alternative models, with patient-donor pair classifications aggregated in the interest of space. The fit of the alternative models is noticably worse on match probabilities and wait times for most of the types, indicating that our preferred parameters do significantly better than the others. Figure D13 shows the second set of moments. Our preferred parameters are slightly worse than some others on this dimension, but does better than others. Table D7 aggregates all three sets of moments into the loss function described above. It shows that the value of the objective function is the lowest at our preferred parameters. A change in the bridge donor waiting policy does not substantially affect the overall loss function. A shorter bridge donor waiting time results in an average chain length that is too low. A longer waiting time does better on chain length, but does poorly relative to our preferred parameters on the overall transplant probabilities in Table D5.

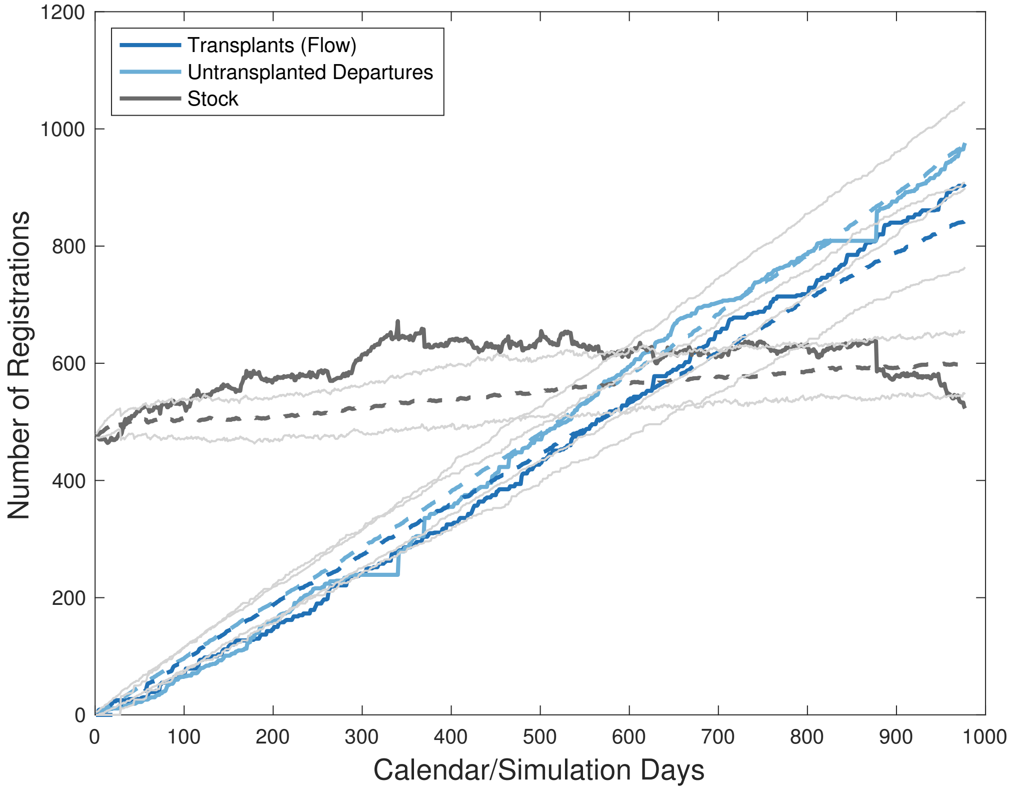


Figure D7: Calibration for All Submissions

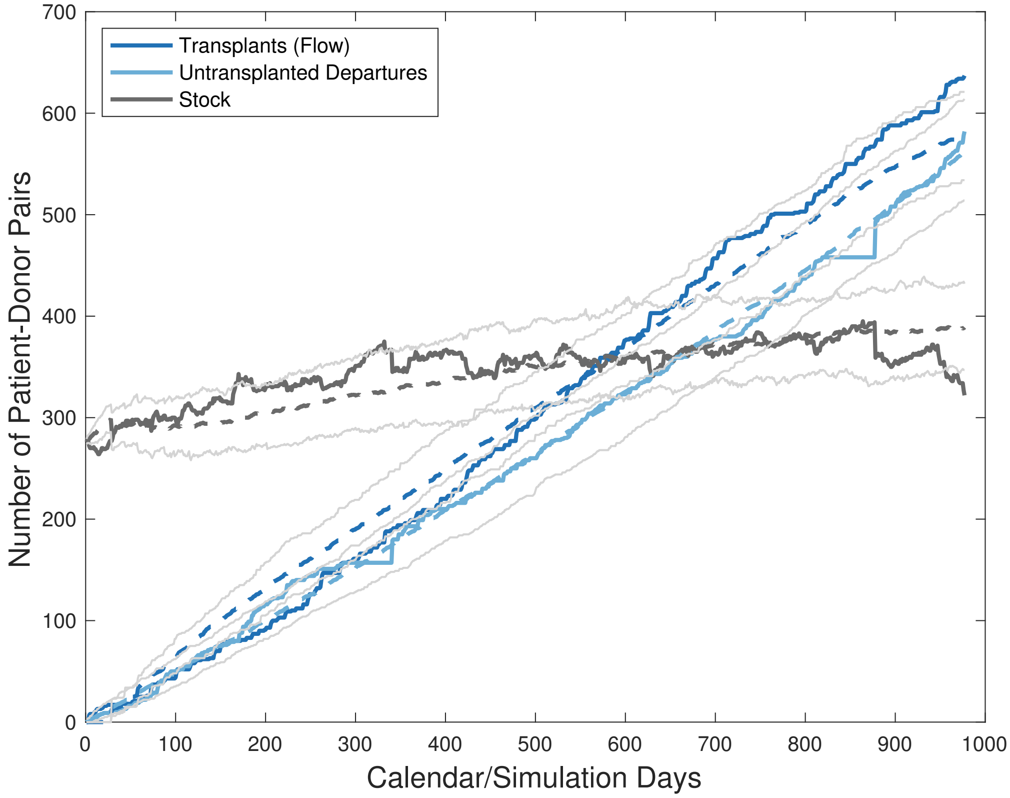


Figure D8: Calibration for Pairs

Note: Observed Quantities and Simulations are shown with solid and dashed lines respectively. Solid grey lines represent 95% confidence intervals.

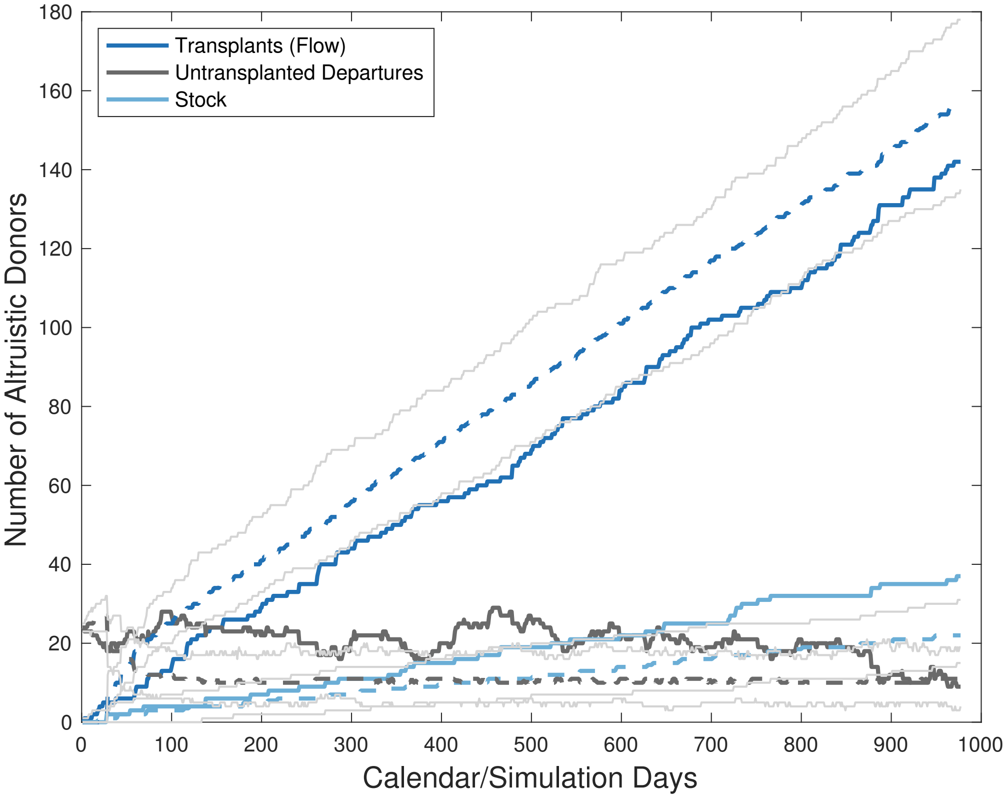


Figure D9: Calibration for Altruistic Donors

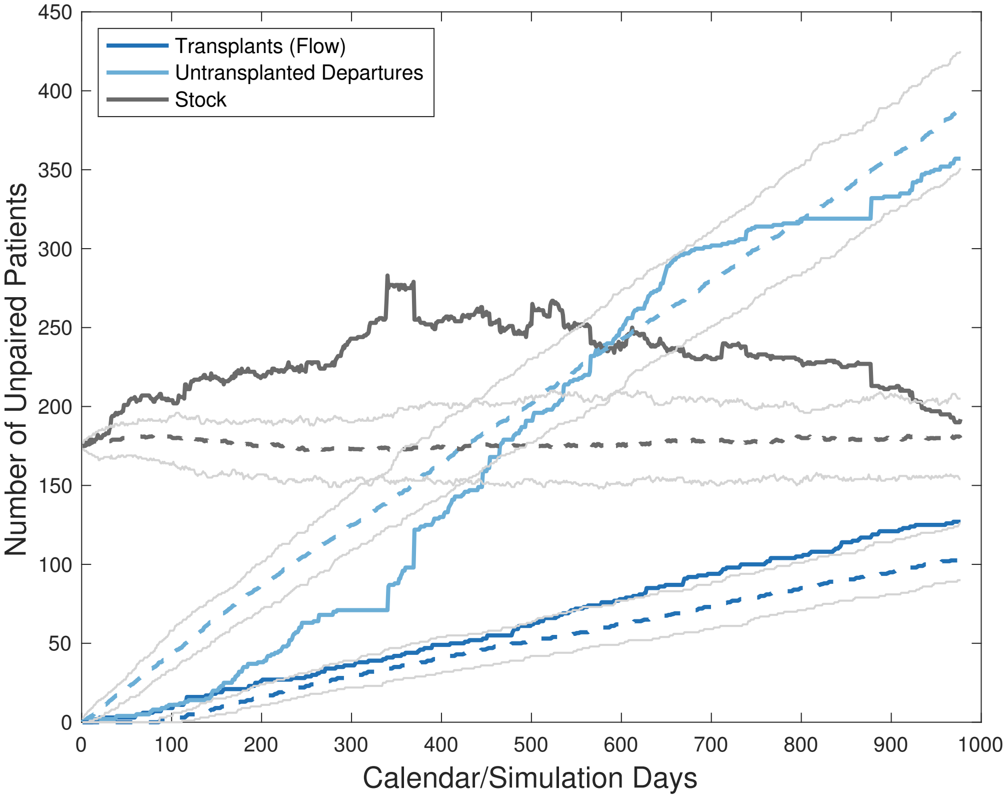


Figure D10: Calibration for Unpaired Patients

Note: Observed Quantities and Simulations are shown with solid and dashed lines respectively. Solid grey lines represent 95% confidence intervals.

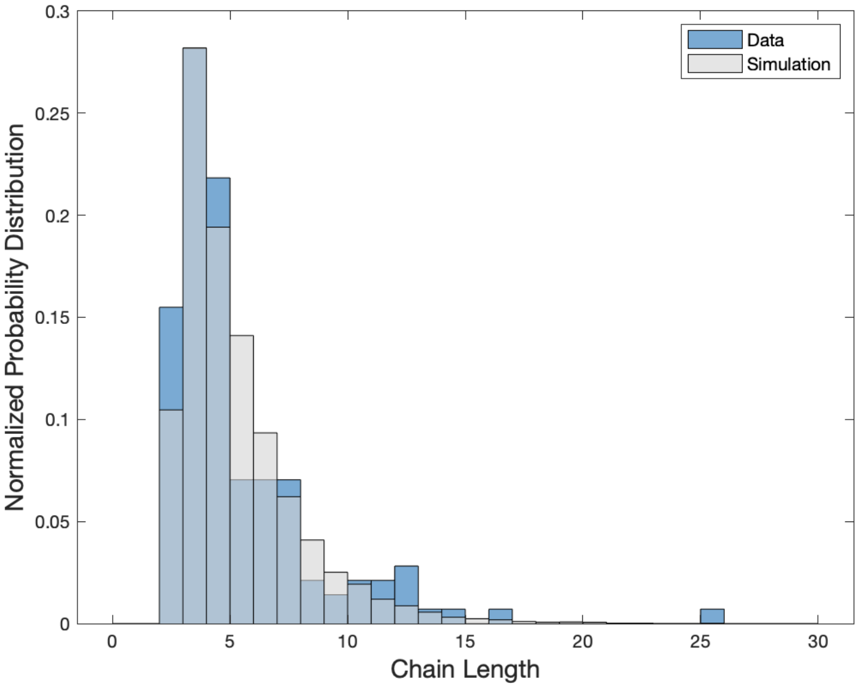


Figure D11: Chain Length Distribution

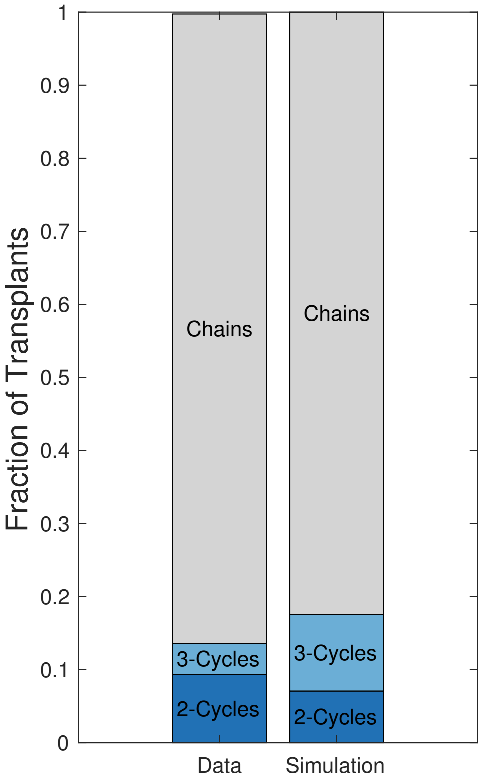


Figure D12: Cycle - Chain Ratio

Table D4: Model Fit

	Patient Match Probability	Donor Match Probability	Mean Days in the NKR			
			Transplanted Patients	Untransplanted Patients	Transplanted Donors	Untransplanted Donors
Patient-Donor Pairs	0.56	0.55	146.67	197.58	140.53	189.92
Unpaired Patients	0.30	-	63.85	135.40	-	-
Altruistic Donors	-	0.79	-	-	105.98	74.91
Under-Demanded	0.41	0.40	195.19	223.27	187.97	215.60
Over-Demanded	0.65	0.66	117.55	129.45	115.37	128.52
Self-Demanded	0.64	0.62	133.09	188.89	126.22	178.06
			Panel B: Simulated Quantities			
Patient-Donor Pairs	0.57	0.54	105.43	183.30	131.69	179.17
Unpaired Patients	0.25	-	90.77	103.71	-	-
Altruistic Donors	-	0.88	-	-	54.65	55.56
Under-Demanded	0.36	0.34	163.68	210.56	196.51	208.84
Over-Demanded	0.68	0.65	76.62	137.06	90.24	133.83
Self-Demanded	0.70	0.66	92.27	156.00	120.37	151.15

Table D5: Model Fit for Alternative Parameters

	Patient Match Probability	Donor Match Probability	Mean Days in the NKR			
			Transplanted Patients	Untransplanted Patients	Transplanted Donors	Untransplanted Donors
			Panel A: Patient-Donor Pairs			
Observed Quantities	0.56	0.55	146.67	197.58	140.53	189.92
Preferred Parameters	0.57	0.54	105.43	183.30	131.69	179.17
Lower Friction	0.60	0.57	98.00	184.64	123.01	179.81
Higher Friction	0.54	0.51	116.42	188.93	144.27	186.02
Shorter Wait Time	0.64	0.62	78.95	189.51	97.65	182.74
Longer Wait Time	0.53	0.49	130.02	184.76	164.38	183.49
Shorter Bridge Donor Wait Time	0.57	0.55	111.97	183.51	132.08	180.95
Longer Bridge Donor Wait Time	0.59	0.55	100.81	184.83	134.98	179.46
			Panel B: Unpaired Patients			
Observed Quantities	0.30	-	63.85	135.40	-	-
Preferred Parameters	0.25	-	90.77	103.71	-	-
Lower Friction	0.25	-	87.90	103.29	-	-
Higher Friction	0.24	-	93.23	103.57	-	-
Shorter Wait Time	0.29	-	63.07	111.48	-	-
Longer Wait Time	0.19	-	113.68	102.91	-	-
Shorter Bridge Donor Wait Time	0.27	-	86.71	105.96	-	-
Longer Bridge Donor Wait Time	0.21	-	89.76	106.33	-	-
			Panel C: Altruistic Donors			
Observed Quantities	-	0.79	-	-	105.98	74.91
Preferred Parameters	-	0.88	-	-	54.65	55.56
Lower Friction	-	0.88	-	-	53.42	53.14
Higher Friction	-	0.87	-	-	56.76	58.20
Shorter Wait Time	-	0.91	-	-	38.26	56.29
Longer Wait Time	-	0.84	-	-	74.45	60.13
Shorter Bridge Donor Wait Time	-	0.89	-	-	50.34	48.34
Longer Bridge Donor Wait Time	-	0.86	-	-	61.10	69.16

Table D6: Statistics on Chains and Cycles for Alternative Parameters

	Share of Transplants			Mean Chain Length
	2-way cycles	3-way cycles	Chains	
Observed Quantities	0.09	0.04	0.86	4.85
Preferred Parameters	0.07	0.10	0.82	4.77
Lower Friction	0.07	0.10	0.83	4.99
Higher Friction	0.07	0.11	0.82	4.52
Shorter Wait Time	0.06	0.08	0.86	5.46
Longer Wait Time	0.07	0.10	0.82	4.57
Shorter Bridge Donor Wait Time	0.08	0.14	0.78	4.48
Longer Bridge Donor Wait Time	0.06	0.07	0.86	5.18

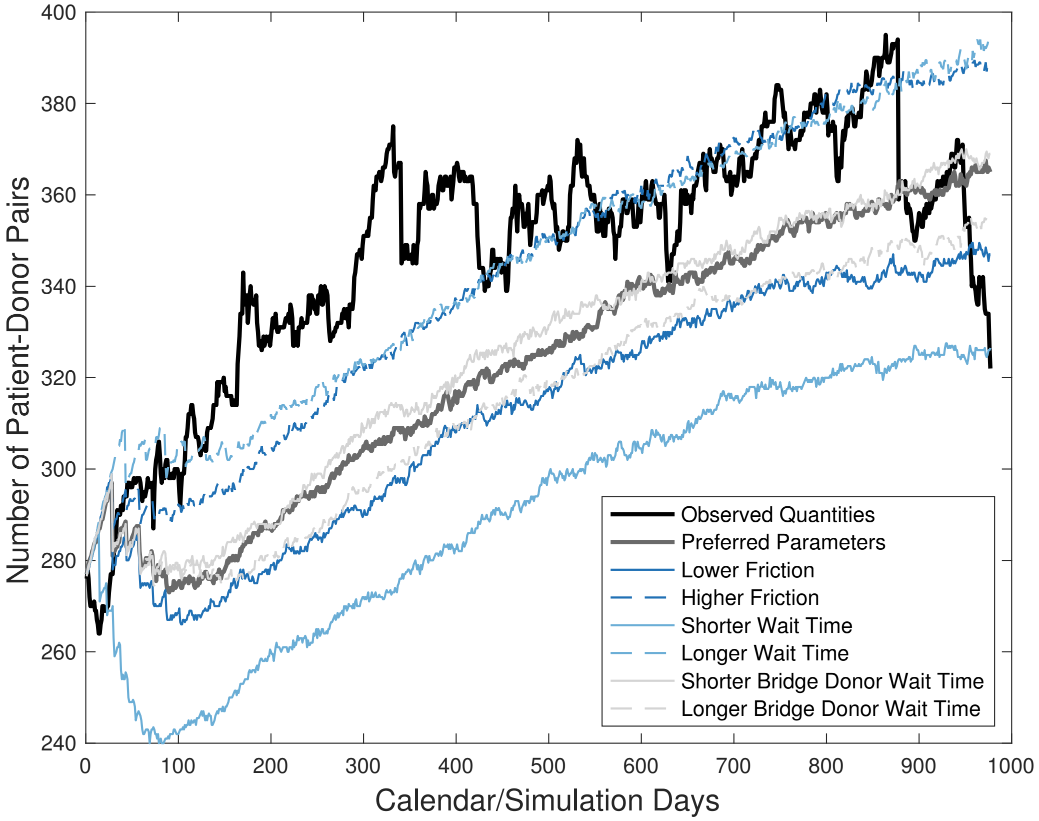


Figure D13: Platform Size for Alternative Parameters

Table D7: Objective Function for Alternative Parameters

	Objective Function
Preferred Parameters	372.5
Lower Friction	530.6
Higher Friction	450.2
Shorter Wait Time	773.5
Longer Wait Time	436.0
Shorter Bridge Donor Wait Time	522.2
Longer Bridge Donor Wait Time	456.4

Table D8: Regression Tree Summary Statistics

	N	Match Probability			Marginal Product			Points per Transplantation
		Mean	S.E.	Within Category Standard Deviation	Mean	S.E.	Within Category Standard Deviation	
Panel A: Altruistic Donors								
Non-O Donor	102	0.84	(0.01)	0.04	0.86	(0.04)	0.20	1.02
O Donor	62	0.93	(0.01)	0.00	1.74	(0.05)	0.11	1.88
Panel B: Patient-Donor Pairs								
O Patient, Non-O Donor	502	0.27	(0.01)	0.16	0.02	(0.02)	0.18	-0.91
O Patient, O Donor, PRA >= 89%	104	0.30	(0.01)	0.22	0.14	(0.04)	0.21	-0.51
O Patient, O Donor, PRA < 89%	135	0.79	(0.01)	0.01	0.67	(0.04)	0.20	-0.15
Non-O Patient, O Donor, PRA >= 95%	138	0.31	(0.01)	0.23	0.09	(0.04)	0.21	-0.71
Non-O Patient, Non-O Donor, PRA < 95%	284	0.85	(0.01)	0.03	0.67	(0.02)	0.24	-0.21
Non-O Patient, O Donor, PRA < 95%	102	0.82	(0.01)	0.01	1.24	(0.04)	0.29	0.52
Panel C: Unpaired Patients								
Unpaired Patients	501	0.22	(0.01)	0.11	-0.03	(0.02)	0.16	-1.13

Notes: Categories are determined by regression tree analysis to predict marginal products as a function of whether a submission is a pair or altruist, blood types, and the patient's PRA. Our procedure followed standard recommendations in **?**. Specifically, we used 10-fold cross-validation to pick the penalty parameter on the number of nodes, required each leaf to have at least 20 observations and pruned a leaf if it does not increase the overall fit by at least 2%. The resulting tree is depicted in Figure 7. Standard errors for the simulations are calculated by following Chapter 12 of **?**. The within category standard deviation is estimated using shrinkage methods recommended in **?**.

E Robustness analyses

This section assesses the robustness of our results to calibrated parameters and the weights used by the NKR (Table D2). The calibrated parameters specifically refer to frictions in consummating proposed transplants due to longer waiting times but higher approval rates produce similar moments as lower waiting times and lower acceptance rates. Chain lengths, however, are increasing in acceptance rates and are best matched by our baseline parameters. We compare our baseline results with two substantially different parameters and a specification with equal weights. The first, labelled "Higher Wait-time and Lower Frictions," has two weeks and three weeks for each of the two approval periods (approval and biological testing) but increases the acceptance rates in each phase from 75% to 80%. The second, labelled "Lower Wait-time and Higher Friction," uses three days and three weeks for each phase, respectively, but decreases the acceptance rates in each phase from 75% to 70%. Finally, the estimates labelled "Equal Weights" use the baseline parameters, but set all the weights in Table D2 to one.

The qualitative and quantitative findings are robust to these alternative parameters. Figure E14 plots average products, as in Figure 5. These alternative parameters yield average product functions that closely follow the baseline. Table E9 shows the inefficiency estimates as in Table 3. The estimated inefficiency is within 5-10% of the baseline. Figure E15 shows marginal product versus matching probability of registrations aggregated by category, as in 6. These results are also qualitatively similar. Table E10 shows marginal product, matching probability, and point system summary statistics, as in Table D8. Again, the points system under the alternative parameters are similar in magnitude. Similarly, the estimates are not sensitive to the weights as they are used only to break ties.

Table E9: Robustness: Total Efficiency Loss

		Efficiency Loss		
	Number of Hospitals	Additional Kidney Exchange Transplants		
		Base	Higher Waittime Lower Friction	Lower Waittime Higher Friction
Panel A: All Hospitals				
All Hospitals	164	505.1	521.6	471.3
Panel B: By hospital size (number of PKEs per year)				
Top Quartile	42	254.2	263.8	237.1
2nd Quartile	48	142.2	149.0	133.9
3rd Quartile	40	81.9	81.2	75.2
Bottom Quartile	34	26.8	27.6	25.1
Panel C: By Platform Membership				
NKR	68	256.4	265.7	240.3
Only UNOS and APD	45	119.8	123.5	110.7
None	51	128.8	132.5	120.3
Panel D: By NKR Participation Rate (Fraction of PKEs facilitated through the NKR)				
Top Quartile	17	17.7	17.8	16.4
2nd Quartile	17	49.2	51.0	46.4
3rd Quartile	17	89.0	92.3	83.2
Bottom Quartile	17	100.4	104.5	94.3

Notes: Constructed as in Table 3.

Table E10: Robustness: Points System

	Match Probability			Marginal Product			Points per Transplantation		
	Baseline	Higher Wait Time Lower Friction	Lower Wait Time Higher Friction	Baseline	Higher Wait Time Lower Friction	Lower Wait Time Higher Friction	Baseline	Higher Wait Time Lower Friction	Lower Wait Time Higher Friction
				Panel A: Altruistic Donors					
Non-O Donor	0.84	0.83	0.85	0.86	0.64	0.84	1.02	0.78	0.99
O Donor	0.93	0.92	0.94	1.74	1.74	1.85	1.88	1.88	1.96
				Panel B: Patient-Donor Pairs					
O Patient, Non-O Donor	0.27	0.28	0.28	0.02	0.02	0.08	-0.91	-0.91	-0.73
O Patient, O Donor, PRA >= 89%	0.30	0.31	0.29	0.14	0.02	-0.01	-0.51	-0.95	-1.04
O Patient, O Donor, PRA < 89%	0.79	0.80	0.81	0.67	0.63	0.78	-0.15	-0.21	-0.04
Non-O Patient, O Donor, PRA >= 95%	0.31	0.33	0.31	0.09	-0.01	0.22	-0.71	-1.02	-0.30
Non-O Patient, Non-O Donor, PRA < 95%	0.85	0.85	0.87	0.67	0.61	0.72	-0.21	-0.29	-0.17
Non-O Patient, O Donor, PRA < 95%	0.82	0.82	0.84	1.24	1.32	1.72	0.52	0.61	1.04
				Panel C: Unpaired Patients					
Unpaired Patients	0.22	0.21	0.22	-0.03	-0.02	0.07	-1.13	-1.12	-0.70

Notes: Constructed as in Table D8.

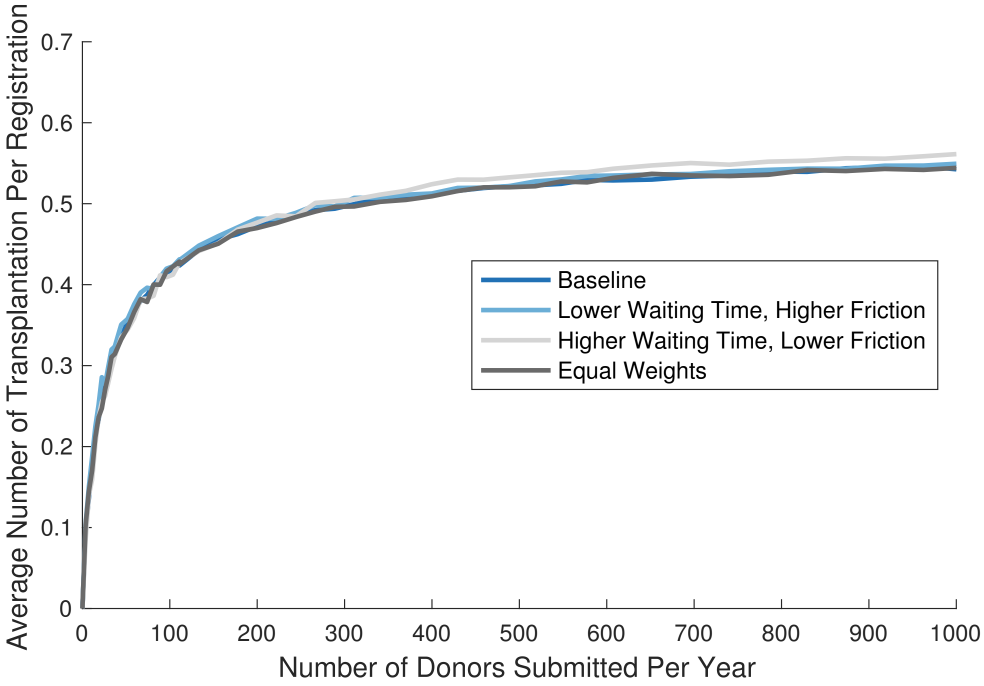
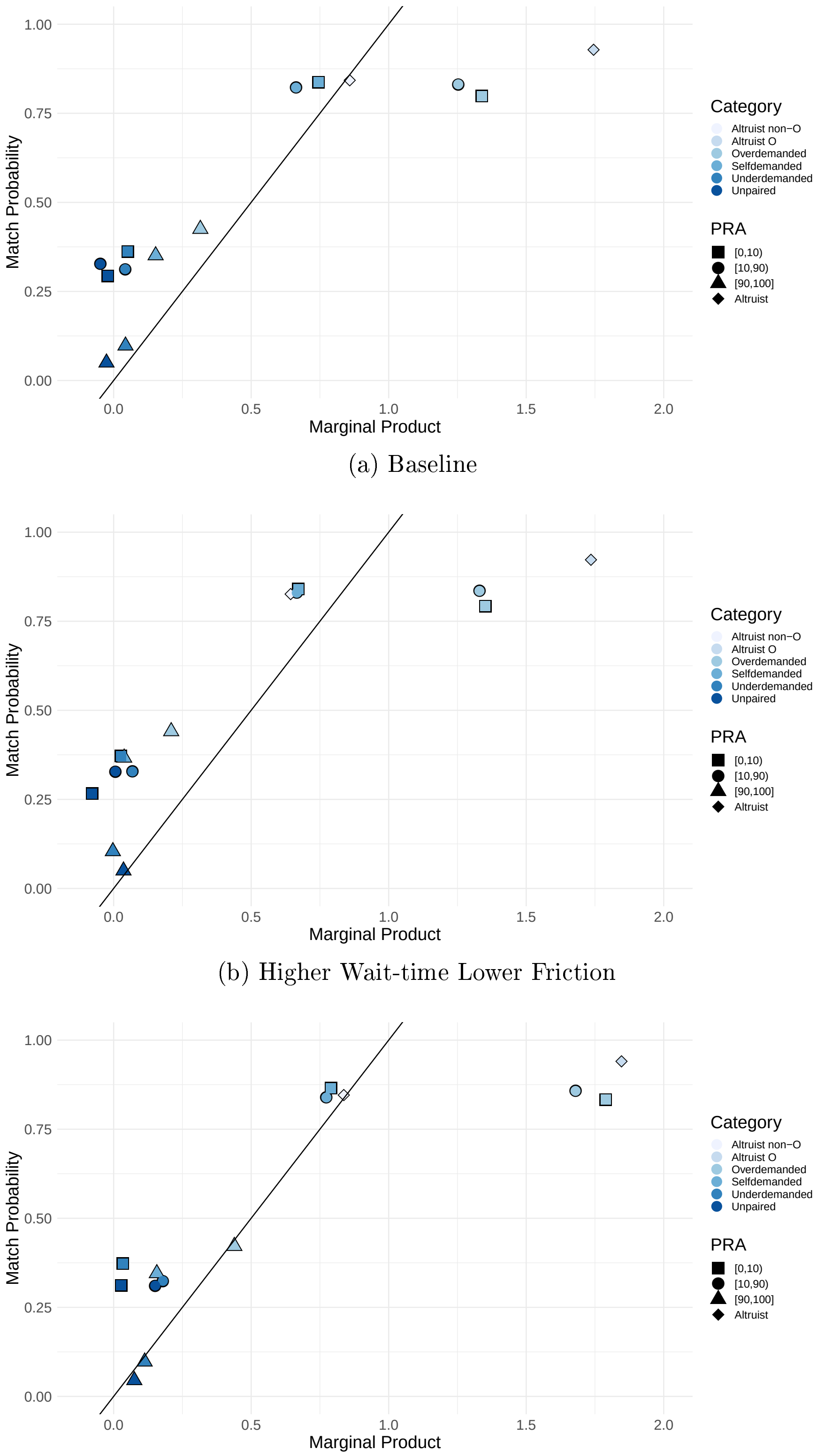


Figure E14: Robustness: Production Function versus Scale

Notes: Constructed as in figure 5.



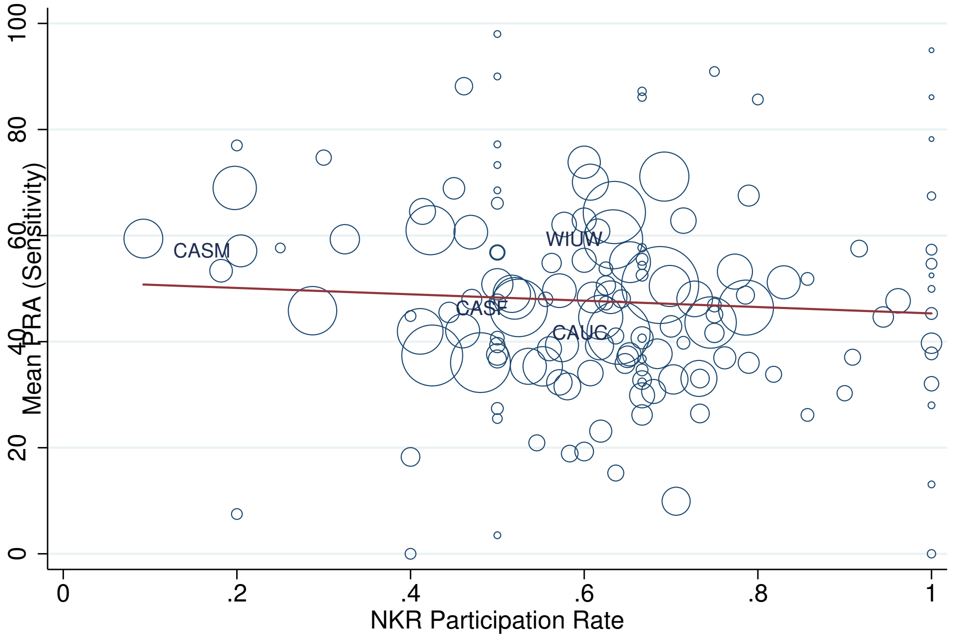
(a) Baseline

(b) Higher Wait-time Lower Friction

(c) Lower Wait-time Higher Friction

Figure E15: Robustness: Private versus Socially Optimal Rewards for Submission Types

Notes: Constructed as in Figure 6.



Note: NKR Participation Rate is defined as # Donors Submitted to the NKR / # Donors Submitted to the N

Figure E16: Selection versus Participation Rate