Congested Waiting Lists and Organ Allocation

Itai Ashlagi* Ravi Jagadeesan† Pengyu Qian‡

August 22, 2025

Abstract

More than 25% of the kidneys that are recovered from deceased donors in the U.S. and are offered to patients on national waiting list are not utilized. This paper shows that waiting list designs can suffer from a form of congestion that can lead to discarding valuable organs. Congestion arises from the interaction of a physical limit on the number of patients that can sequentially consider an organ before it accrues excess cold ischemic time (i.e., *expires*) with incentives of patients to try to obtain organs of the highest quality. We develop a parsimonious framework to study equilibria of waiting lists with congestion, and show that congestion provides a strong enough externality to substantively affect welfare and wastage. We further show how delegating decisions to doctors can increase congestion. We also discuss how recent changes to waiting list designs affect congestion, and policies and market designs that can mitigate congestion.

*Department of Management Science and Engineering, Stanford University. Email address: iashlagi@stanford.edu.

†Department of Economics, Stanford University. Email address: ravi.jagadeesan@gmail.com.

‡Questrom School of Business, Boston University. Email address: pqian@bu.edu.

1 Introduction

In 2024, 3,571 patients died while waiting for a kidney transplant from a deceased donor in the United States.[1][2] Patients transplanted during 2024 waited on average around 4 years. At the same year more than 28% of recovered kidneys were not used, a rate which has been increasing over recent years (Wood et al., 2025; Mohan et al., 2023; Puttarajappa et al., 2023). This paper seeks to understand how practical features of the organ allocation system can lead to excess non-utilization of grafts. One central feature in this study is that organs have a limited amount of cold ischemic time before they expire. We make three contributions. First, we show how congestion can arise and its externalities on efficiency and wastage. Second, we explore how (the common practice of) delegating decisions to doctors can increase congestion. Third, we discuss how recent changes to waiting list designs affect congestion and policies that mitigate congestion and reduce inefficiencies.

Organs from deceased donors are allocated to patients through waiting lists. When an organ becomes available it is typically offered sequentially to patients in priority order. Once an offer is made, the patient can decline the offer without penalty. Patients at the "top" of the list may be inclined to decline to accept lower quality organs even if they are useful, as such patients stand a chance of being offered a high-quality organ in the future. By contrast, patients further from the top of the waiting list might very well accept lower quality organs, as they have less of a chance of surviving (and remaining healthy enough to undergo transplantation) until being offered a high-quality organ. A large number of organ refusals comes at a cost as organs have limited cold ischemic time. So when there are many patients on the waiting list, lower quality organs are more likely to expire before being offered to a patient who would be willing to accept them—despite there being such patients on the waiting list (Mehrotra et al., 2022). We refer to this feature of equilibrium allocations of waiting lists as *congestion*.

This paper investigates inefficiency of waiting list for organs in the presence of congestion. It turns out that congestion creates a strong enough externality to substantively affect the welfare and wastage induced by waiting lists. When congestion exists, it necessarily hurts both patients who are interested in lower quality organs and ones who are not by raising equilibrium waiting times for all organs.

However, we show that congestion can be alleviated by recently used practices of expediting organs and open offers to reduce wastage of lower quality organs. These policies help not only patients who are interested in lower quality organs, but also patients who are interested

[1] `https://optn.transplant.hrsa.gov/data/`.

[2] During the same year 4,758 patients were removed from the waiting list due to being too sick to transplant.

only in higher quality organs. Intuitively, improving the allocation of lower quality organs entices more patients to accept such organs, and therefore lowers the equilibrium waiting times for higher quality organs as well.

To obtain these conclusions, we develop a parsimonious fluid-limit framework to study waiting lists that may suffer from congestion. In our model, patients and organs both arrive as flows of constant rates. We assume that patients depart from the waitlist stochastically (e.g., due to death), but waiting is otherwise costless. To capture a key dimension of heterogeneity in the organ allocation context, we suppose that organs can be high quality (safe) or low quality (risky).[3] Patients are heterogeneous in how they value risky organs relative to safe organs. And patients' departure rates can depend on their type (e.g., sicker or older patients may depart the waiting list unmatched at a faster rate).

Organs are offered sequentially to patients in the order of arrival and patients can decline offers. To capture the perishability of organs with excess cold ischemic time, we assume that organs expire after being offered to ε mass of patients.[4] A consequence of this perishability and the sequentiality of offers is that organs—including risky organs—are never offered to patients who are not among the top ε on the waiting list. Hence, if many patients strongly prefer safe organs, congestion can arise and risky organs may be wasted.

We first establish the existence of an equilibrium, captured by the steady state of the system. While there maybe multiple equilibria, there is at most one equilibrium that is not congested. The framework allows us to conduct comparative statics and investigate the welfare properties of the steady state.

Congestion is relevant not only to patients who maybe willing to accept risky organs. Indeed, congestion incentivizes such patients to wait for safe organs, thereby raising the waiting time for safe organs and hurting even patients who strongly prefer safe organs. When congestion occurs, all patients would strictly prefer the steady state of a counterfactual with non-perishable organs over the congested steady state.

To demonstrate the severity of the welfare loss caused by congestion. we consider a hypothetical policy of destroying some safe organs. Without congestion, this policy would obviously weakly hurt all patients and strictly hurt some. By contrast, we find that in the presence congestion, it is possible that such a policy strictly benefits all patients. Intuitively, destroying some safe organs should raise the waiting time for such organs. However, this effect makes patients more willing to accept risky organs. Without congestion, the waiting time for risky organs would weakly increase a result. However, this also makes the waiting

[3]For interpretation, we focus on patients of a single blood type and abstract from other forms of incompatibility (e.g., tissue type incompatibility).

[4]We abstract from other forms of decay in the quality of organs with the elapse of cold ischemic time.

list less congested, reducing the number of discards of risky organs. Hence, the waiting time for risky organs could in fact decrease. This in turn makes fewer patients wait for safe organs, so the waiting time for safe organs could actually decrease. The fact that destroying organs can lead to a strict Pareto improvement by reducing wastage shows that congestion causes a strong form of inefficiency on waiting lists.

Second, we show that in the presence of congestion, the delegation of organ acceptance decisions to risk-averse doctors always harms all patients. In practice, doctors and hospitals make most decisions on behalf of their patients (especially for declining offers). There is also evidence that transplant centers vary substantially in the type of organs they are willing to accept (King et al., 2020). Transplant centers are penalized for adverse post-transplant outcomes (Chan and Roth, 2024; Mildebrath et al., 2024). As a result, hospitals may be less inclined to accept a lower-quality organ than a patient even when a patient may benefit from a lower-quality organ.

This form of moral hazard inflates the relative demand for safe organs. Without congestion, the change in demand increases the waiting time for safe organs but lowers demand for risky organs. However, the increase in relative demand for safe organs also worsens congestion on the waiting list, which we show increases wastage enough to raise the waiting time for risky organs as well. Intuitively, the congestion externality means that increases in the demand for safe organs hurts not only patients who wait for safe organs, but also ones who are willing to accept risky organs. Moreover, while some hospitals may benefit from not using risky organs, hospitals are made worse off in the aggregate by making decisions on patients' behalves due to impact on congestion. Thus, the congestion externality makes the misalignment of incentives between patients and doctors particularly harmful.

Third, we show that efforts to improve market thickness can worsen congestion and harm patients. This insight may shed light on some of the increase in organ discards. On March 2021, the Organ Procurement and Transplantation Network (OPTN) has moved from offering a "local" priority to patients based on an Organ Procurement Organization (OPO) service area[5] to offer priority to patients within 250NM from the donor hospital. This change increased the number of patients that are given local priority.[6] At the same time the discard rate of kidneys from deceased donors has increased from 21.2% in 2020 to 24.5% to 2021.[7]

[5]Referred to as a donor service area.

[6]From 2020 to 2021, the median distance between the donor's hospital to the patient hospital from around 65 to 120 miles.

[7]There has been further additional increase to 26.5% in 2022 and 27% in 2023. Other factors may of course play a role in this; e.g. modifications to the CMS Final Rule for measuring the performance of OPOs have took place in the beginning of 2022 (which have potentially led to an increase in the procurement of worse organs).

1.1 Related Literature

Methodologically, we show how to use monotone comparative statics methods (Milgrom and Roberts, 1994) to analyze queuing systems whose steady states are not tractable in closed form, even when there are a multiplicity of steady states. This is important to our analysis of congested waiting lists with heterogeneous patient departure rates.

This paper contributes to a growing literature on waitlist systems with private information. Several papers study the design of the waiting list to maximize efficiency or other objectives (Su and Zenios, 2004; Bloch and Cantala, 2017; Leshno, 2022; Arnosti and Shi, 2020; Waldinger, 2021; Nikzad and Strack, 2023; Ashlagi et al., 2025). These papers assume that agents' utilities are quasi-linear with waiting times. In our model agents depart exogenously and have expected utility with respect to the risk of not getting an object. A key difference is that these papers assume that the object is always allocated if the waiting list is not empty. Notably, in Su and Zenios (2004) and Bloch and Cantala (2017) objects may be discarded if they are declined by all agents in the queue; they find that Last-Come-First-Serve and Random order increase utilization by reducing selectivity at the expense or welfare. In our model, objects are not utilized due to congestion.

Several papers study the (re)design of waitlist rules in organ allocation policies with rational agents. Su and Zenios (2004, 2006); Agarwal et al. (2021); Shi and Yin (2022) and Sweat (2023) identify policies that can improve efficiency and reduce utilization by limiting choice, changing priorities, or allocating some organs to agents down the list. These papers assume that organs do not perish before offered to all agents on the list. Zhang (2010); De Mel et al. (2021) and Doval et al. (2024) identify and address inefficiencies that arise due to herding.

Our paper builds on the model of Shi and Yin (2022), which assumes that organs do not expire ($\varepsilon = \infty$) and thus congestion cannot occur. The key finding is that waiting times do not clear the market; agents who decide to wait for a safe organ and die before getting such an organ, would have been better off accepting a risky organ. Instead we focus on the negative externalities generated from congestion.

A notable paper that allows objects to perish is Castro et al. (2021). Their model assumes agents have identical preferences and no departures but their main finding for improving efficiency would be effective in our model as well: skipping agents on the list to offer less desired objects to agents who would accept them (rather than waiting for better objects).[8]

This paper contributes more broadly to the study of the frictions and inefficiencies that arise in organ allocation and congested systems.

[8]The medical literature also discusses such remedies (Gentry and Segev, 2013; Mohan and Schold, 2022; Guan et al., 2025).

2 Background

2.1 Generic data

In the United States there are over 500,000 people with end-stage renal disease. Medicare provides near-universal coverage for this disease and has spent over $35 billion treating kidney failure in 2016 (USRDS, 2018). Transplantation is the preferred treatment and it does not only improve health outcomes, it further saves Medicare more than $200,000 over dialysis costs over the life a transplanted patient (Wolfe et al., 1999; Held et al., 2016).

The allocation of deceased donors is done through waiting list but there is a severe shortage of organs for transplantation. In 2024 47,159 patients joined the waiting list, more than 66% of them listed while on dialysis. At the same year, there were 21,341 kidney transplants from deceased donors in the US 3600 died while for a kidney transplant and 4,758 patients were removed from the list due to being too sick to be transplanted. Figure 1 shows the annual trends of the deceased donor waiting list. [9]

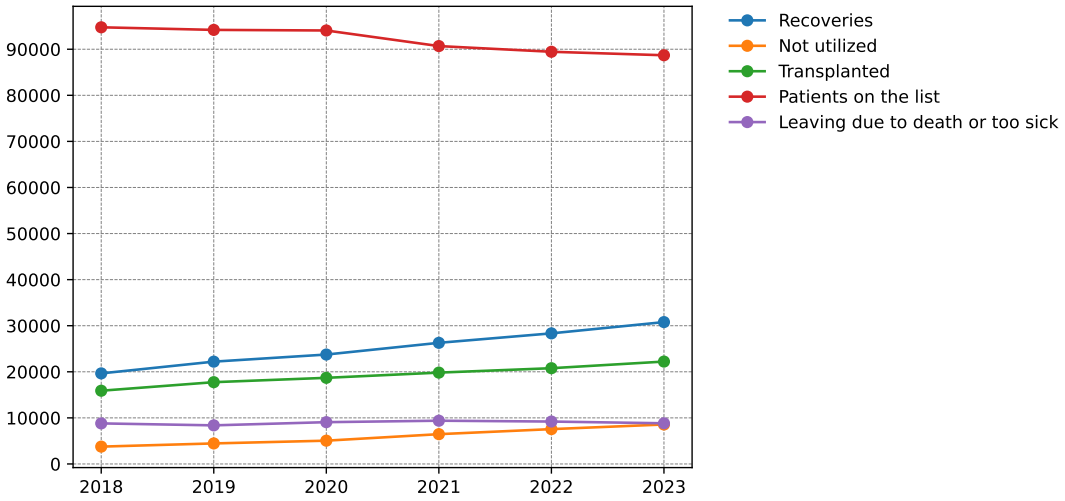


Figure 1: Annual trends in the deceased donor waiting list for cadaver kidneys.

2.2 Institutional features and data

There are a few features that are central in the economics of a waiting list for organ allocation. First, kidneys for transplantation come with heterogeneous qualities or risk. How risky a donors' kidney depends on her kidney function, age, BMI, cause of death, etc. A common measure that captures the risk of an donor is the Kidney Donor Risk Index (KDRI). A

[9]Data presented in this paper is from the United Network for Organ Sharing (UNOS), which includes the Standard Transplant Analysis and Research (STAR) files for de-identified information on all patients, as well as the Potential Transplant Recipient (PTR) files de-identified organ offers.

percentile conversion of this measure is called the Kidney Donor Percentile Index (KDPI) and lower KDPI values are associated with higher donor quality and longer expected longevity.[10] While only around 40% of patients survive on dialysis for more than 5 years (USRDS, 2013), even the lowest 15% of organs based on KDPI organs have more than 60% five year graft survival. A large fraction of non-utilized kidneys have qualities similar to organs that have been transplanted (Mohan et al., 2018; Aubert et al., 2019). For example, in 2023 there were 1547 deceased donors whose both kidneys recovered but only one of them transplanted (Figure 2). It is therefore plausible that many kidneys that are currently not utilized could in fact be valuable as grafts, leading to numerous excess deaths of patients on the waiting list.

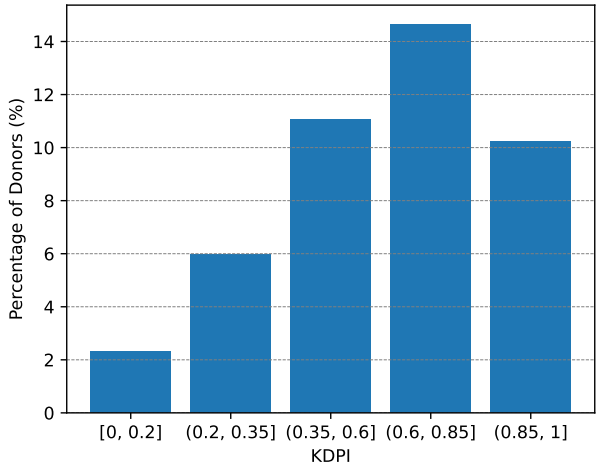


Figure 2: Percent of donors in each KDPI group for which two kidneys were recovered but only one was transplanted.

Second, organs from deceased donors are organized through waiting lists and offers can be declined by patients and doctors without penalty. The waiting lists for deceased donors are ordered according to priorities that include medical characteristics, demographics as well as waiting time. When an organ becomes available, it is offered to patients according to priority order.[11] Once an offer is made, the patient or her doctor can decline the offer without affecting the patient's priority. [12]

Since patients at the "top" of the list may be inclined to decline offers of organs with low

[10]KDRI is based on basic lab test prior to recovery. In practice, however, further information regarding organ quality is revealed after recovery and during the allocation process (biopsy results, pump values, etc.)

[11]For details of the current organ allocation policies in the US see `https://optn.transplant.hrsa.gov/media/eavh5bf3/optn_policies.pdf`. For example, healthier patients are prioritized for highest quality organs. increased priority is also given to patients with a perfect match or patients who are highly sensitized.

[12]In practice, they may need to do so because particular organs are not a good enough match for them (e.g., since the patient may reject the organ due to existing antibodies, which is discovered through a crossmatch test.)

quality, such organs may accumulate many refusals. Figure 3 provides the number of offers that organs that are eventually transplanted accumulate by their quality. So the process of

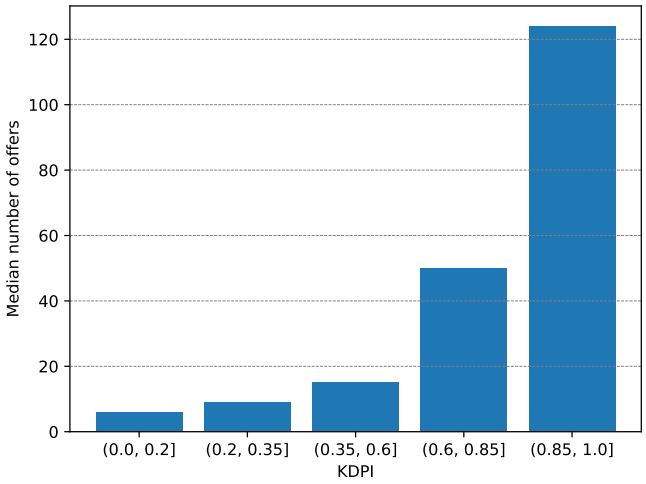


Figure 3: Median number of offers for transplanted organs by Kidney Donor Percentile Index (KDPI).

repeatedly declining organs affects which patients receive which organs, but not necessarily which organs are utilized or the overall efficiency of the allocation system.

Third, recovered organs have a limited time to be placed. A long cold-ischemic time (the time on ice since recovery to being used) can have a substantial impact on the graft survival (Quiroga et al., 2006; Peters-Sengers et al., 2019; Debout et al., 2015). Recovered kidneys are medically viable for only around 24-36 hours after procurement and declines in quality as CIT increases. Refusal reasons due to long-cold-ischemic time become very common already 6 hours after recovery (Figure 4).[13] So lower quality organs, which require many offers to patients on the list, may therefore accrue excess cold ischemic time and *expire* (Barah et al., 2022; Guan et al., 2025) before reaching a patient who would be willing to accept the offer –despite there being such patients on the waiting list (Mehrotra et al., 2022).

Finally, in practice doctors often make decisions to decline an offer on behalf of their patients. Guan et al. (2025) distinguish between a transplant center refusing an organ for a single patient or simultaneously refusing it for multiple patients and they document that the majority of the refusals are of the latter category.[14] At the same time survey data suggests that patients are more inclined to accept low quality organs than their doctors.(Mehrotra et al., 2022; Sepulveda et al., 2023). Our analysis focuses on patients' decisions, but we return in Section 7 to discuss how do equilibrium outcomes change under moral hazard.

[13]See also Guan et al. (2025) for a detailed analysis of offer refusals.

[14]In 2022-2023 more than 69% of organs belong are multi-patient refusals.

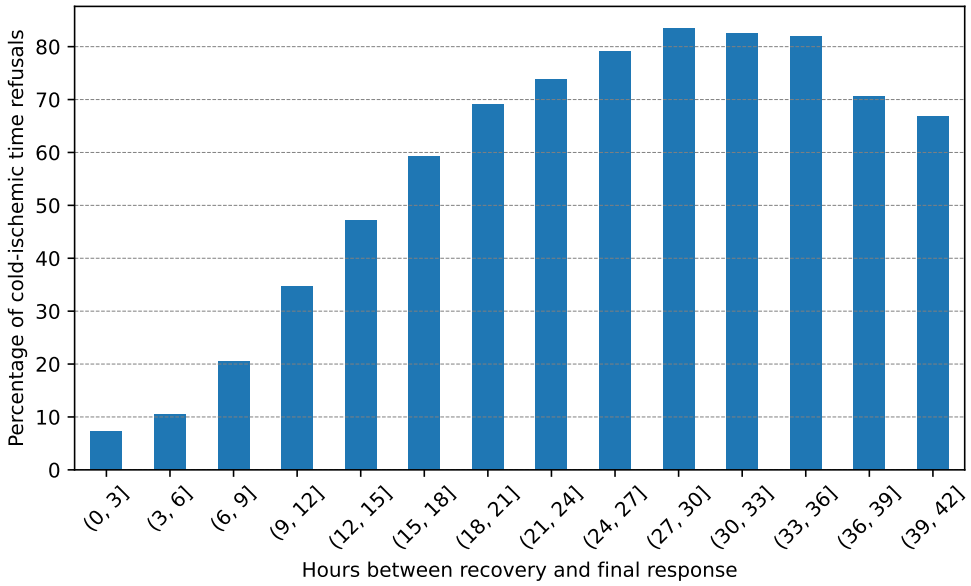


Figure 4: Percent of offer refusals due to long cold-ischemic time arriving x hours since organ recovery.

3 Model

This section outlines a continuous-time dynamic fluid model for matching organs and patients. Thus, we suppose a continuum of patients arrive on the waitlist at a flow of a constant rate p. The model makes the equilibrium analysis and comparative statics tractable by avoiding fluctuations due to stochastic arrivals and departures.

3.1 Organs

There are two types of organs, safe grafts (denoted by s) and risky grafts (denoted by r). A continuum of safe organs (resp. risky organs) arrives at a flow of constant rate sp (resp. rp). We assume that $s < 1$, capturing the shortage of safe organs.

Each arriving organ is offered sequentially to patients in descending order of waiting time. However, an organ expires after $\varepsilon > 0$ mass of patients decline the organ. Here, ε can be interpreted as a measure of the lifetime of graft (called "cold ischemic time" in the medical literature). Thus, each organ goes to a highest-ranking patient on the waitlist (at the time of the organ's arrival) who is willing to accept the organ, except that patients who have more than ε mass of patients ahead of them on the waitlist do not receive organs.

3.2 Patients' Preferences and Decisions

Patients are heterogeneous in their individual characteristics, as well as in the characteristics of transplant centers that manage their care. We summarize these characteristics in a type θ, which we assume is drawn independently across patients from an atomless probability distribution $f(\theta)$ on a measurable space Θ.

An unmatched patient of type θ remains on the waiting list for an exponentially distributed time with rate $\delta(\theta) > 0$, after which they depart if still unmatched. We assume that $\underline{\delta} \triangleq \inf_{\theta \in \Theta} \delta(\theta) > 0$. Note that agents' waiting time is upper bounded by $\bar{t} \triangleq \underline{\delta}^{-1} \cdot \log(s^{-1})$.[15] We denote the set of possible agent waiting times by $T \triangleq [0, \bar{t}]$, and refer to an agent who has waited t as an agent at position t. Upon receiving an offer, the agent decides whether to accept it. We restrict attention to steady-state equilibria where the state of the waitlist does not change over time, and the agents make decisions solely based on their type and position. An agent of type θ has a decision rule $D_\theta : \mathbb{R}^+ \to \{s, r\}$ such that $D_\theta(t) = s$ means agent only accepts safe organ when at position t, and $D_\theta(t) = r$ means agent accepts both types of organ when at position t. Every agent chooses her decision rule in order to maximize her (Bernoulli) utility

$$u_{\text{patient}} = \begin{cases} 1 & \text{if get safe organ} \\ \nu(\theta) & \text{if get risky organ} \\ 0 & \text{if do not get an organ} \end{cases} .$$

where $\nu(\theta) \in [0, 1]$ is the true willingness to accept a risky organ.

In Section 7, we introduce an extension in which organ acceptance/rejection decisions may be delegated to hospitals.

3.3 Discussion of Assumptions

To isolate the impact of organ expiry due to excess cold ischemic time, our model abstracts away features of the national waiting list for deceased donor kidneys.

In practice, the waiting list is based on priorities, which depend on medical and geographic characteristics of the organ and the patient including waiting times.[16] However, a majority of organs from deceased donors are effectively allocated based on waiting times within a local region.[17] Our analysis can be interpreted as applying within any of the majority of priority

[15] This bound corresponds to the case where all agents wait for a safe organ, and the departure rate for every agent is assumed to be at its lower bound.

[16] For example, in kidney allocation in the US, the healthy and young patients are prioritized for organs with the top 20% quality (KDPI).

[17] For each organ typically few patients are assigned a high priority (as few patients have a perfect match, or are very sensitized yet still compatible with the organ). For other patients compatible with the organ,

classes for which organs are effectively allocated by waiting times.

In practice, there are numerous different qualities of organs. For tractability, we consider two types of organs. In practice, organ quality can degrade continuously over time (Sampaio et al., 2018). For tractability, we assume that the quality of the organ remains fixed for ε offers after which it immediately expires.

Last, we assume that agents depart according to an exponential distribution. This assumption is also close to empirical observations and has been widely applied in the literature via COX regression models (Medin et al., 2000; Schaubel et al., 2009; Agarwal et al., 2021).

4 Characterization of Equilibria

4.1 Equilibrium definition

We first formally define the dynamic game among agents under an FCFS mechanism with deferrals, and its steady-state equilibrium. Some of the definitions are adapted from Ashlagi et al. (2025).

Definition 1. A *steady-state equilibrium* of the FCFS policy is defined by a tuple $(A, b, m, \{D_\theta\}_{\theta\in\Theta})$ that satisfies the following properties:

1. **Well-definedness** of agent densities, organ offers, allocation rates, and agent decisions.

    - *The densities of agents at all positions are well-defined:* For each $t \in T$, $A(t)$ is a Lebesgue measure defined on Θ and has a density function $a(t,\cdot) : \Theta \to \mathbb{R}_+$. The function $a(t,\theta)$ gives the density of agents of type θ at position t. For each $\theta \in \Theta$, $a(\cdot,\theta) : T \to \mathbb{R}_+$ is left-continuous and weakly decreasing.
    - *The densities of organs offered to agents at every position are well-defined:* Define $b : T \times \{s, r\} \to \mathbb{R}_+$, and for $x \in \{s, r\}$, $b(\cdot, x)$ is right-continuous and weakly increasing. We denote by $b(t, x)$ the measure of organs with type x that are offered to agents with waiting time t.
    - *The allocation rates of organs to agents at every position are well-defined and feasible:* Let $m(\cdot,\cdot, s)$ and $m(\cdot,\cdot, r)$ be two Lebesgue measures defined on $T \times \Theta$. The restriction of m to fixed $t \in T$ and $x \in \{s, r\}$ has a density, which we denote by $m(t,\cdot, x)$. For all $t \in T$, $x \in \{s, r\}$, $\int_{\theta\in\Theta} m(t,\theta, x)\, d\theta \le b(t, x)$, i.e., the total

priority is largely affected by dialysis time. See `https://optn.transplant.hrsa.gov/media/eavh5bf3/optn_policies.pdf` for organ allocation policies in the US.

rate type-x organs are allocated to agents with waiting time t is no more than the rate it is offered; also, $\sum_{x\in\{s,r\}} m(t,\theta,x) \leq a(t,\theta)$, i.e., the total rate organs are allocated to agents of type θ at position t is no more than the density of these agents.

- *The agents' decisions are respected:* $D_\theta : T \to \{s,r\}$ is left-continuous for all θ, and $m(t,\theta,x) > 0$ only if: (a) $x = s$ (when offered a safe organ, an agent who remains on the waitlist always accepts it), or (b) $x = r$ and $D_\theta(t) = r$ (an agent is allocated a risky organ only if she chooses to accept it).

2. **Conservation laws**.

- *Conservation law for agents:* Recall that agent densities $a(t,\theta)$ are left-continuous and non-increasing in their first argument. By Froda's theorem, $a(\cdot,\theta)$ has countably many jumps on T. For any $t \in T$ where $a(\cdot,\theta)$ jumps,

$$a(t,\theta) - \sum_{x\in\{s,r\}} m(t,\theta,x) = \lim_{t'\to t^+} a(t',\theta)\,,$$

i.e., agents at position t depart due to being matched. For any $t \in T$ where $a(\cdot,\theta)$ is continuous, there exists $\xi > 0$ such that $a(\cdot,\theta)$ is continuous on $(t-\xi, t+\xi)$:

$$a(t_1,\theta) = a(t_2,\theta)\cdot e^{\delta(\theta)(t_2-t_1)}\,, \quad \forall t_1,t_2 \in (t-\xi,t+\xi)\,,$$

i.e., agents at position t depart due to exogenous departure process.

- *Conservation law for organs:* Organs remain available until they are allocated. For any $t \in T$ and $x \in \{s,r\}$,

$$b(t,x) - \int_{\theta\in\Theta} m(t,\theta,x)\,d\theta = \lim_{t'\to t^-} b(t',x)\,.$$

Moreover, the initial conditions are $b(\bar{t},s) = sp$, $b(\bar{t},r) = rp$, i.e., all arriving organs are first offered to agents who wait the longest.

3. **Market clearing conditions.**

- *Well-definedness of agents' demand.* For any t and x, the set of θ demanding x at t is Lebesgue measurable.
- *Organs are not intentionally withheld at any position.* If $\int_\Theta m(t,\theta,x)\,d\theta < b(t,x)$, then for every agent type $\theta \in \Theta$ demanding x at position t, $m(t,\theta,x) = a(t,\theta)$.

- *The highest demanded organ type at each position is fully allocated.* If x is the highest demanded type at t, then $\int_\Theta m(t,\theta,x)\,d\theta = b(t,x)$.

4. **Optimality of an agent's decision.** Each agent's decision rule D_θ maximizes her utility defined in Section 3.2.

5. **Cold-ischemic time limit.** An organ expires after being declined by ε measure of agents: for $x \in \{s,r\}$, we have $\int_{t\in T} b(t,x)\,dt \le \varepsilon$.

4.2 Equilibrium Properties and Congestion

In the following, we show that the steady-state equilibrium defined above exhibits a remarkably simple structure.

Proposition 1 (Waiting time cutoffs)**.** *Let $(A, b, m, \{D_\theta\}_{\theta\in\Theta})$ be a steady-state equilibrium of the FCFS policy exists, as defined in Definition 1. Then there exist $0 \le T_{\text{risky}} < T_{\text{safe}} \le \bar{t}$ such that*

- No agents wait beyond T_{safe}: *For all $\theta \in \Theta$ and $t > T_{\text{safe}}$, $a(t,\theta) = 0$.*
- All safe organs are allocated at T_{safe}: *$\int_{\theta\in\Theta} m(T_{\text{safe}},\theta,s)\,d\theta = sp$, and $b(t,s) = 0$ for all $t < T_{\text{safe}}$.*
- Risky organs are allocated only at time T_{risky}: *$\int_{\theta\in\Theta} m(t',\theta,r)\,d\theta = 0$ for every $t' \neq T_{\text{risky}}$.*

The proof of Proposition 1 is in Appendix A.

Proposition 1 shows that under the FCFS policy, waiting times function as prices that clear the market, with higher quality organs commanding longer waits.

Suppose that a patient of type θ is offered a risky organ at time t but knows that they will receive an offer of a safe organ at time $t' \ge t$. Maximizing expected utility taking into account the exogenous rate of unmatched departure, they will accept the risky organ if $t' - t > W_\nu(\theta)$, where

$$W_\nu(\theta) = \frac{-\log \nu(\theta)}{\delta(\theta)},$$

and accept the safe organ if $t' - t < W_\nu(\theta)$. (If $t' - t = W_\nu(\theta)$, then the patient has multiple optimal strategies.)

We call the quantity $W_\nu(\theta)$ the *willingness to wait for quality* for a patient with type θ. Throughout, we assume that W_ν has a distribution that has strictly positive density on $(0,\infty)$, though possibly with point masses at ∞. Hence the organ assignment in steady-state

equilibrium is determined by an agent's willingness to wait: A type θ agent rejects a risky organ and waits for a safe one if

$$W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}\,,$$

and accepts a risky organ if

$$W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}\,.$$

Equilibrium waiting times can be characterized in terms of the distribution of willingness to wait.

Proposition 2 (Steady-State Waiting Times)**.** *A steady-state equilibrium under the FCFS policy exists if and only if there is a pair of waiting times* $0 \leq T_{\text{risky}} < T_{\text{safe}} \leq \bar{t}$ *that satisfies the following conditions:*

1. *All safe organs are allocated (recall that there is shortage of safe organs):*

$$\int_{\theta: W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta) T_{\text{safe}}} f(\theta) d\theta = \ s\,. \tag{1}$$

2. *Demand of risky organs cannot exceed their supply:*

$$\int_{\theta: W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta) T_{\text{risky}}} f(\theta) d\theta \leq \ r\,. \tag{2}$$

3. *An organ expires after being declined by* ε *measure of agents:*

$$p \int_{\theta: W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}} \int_{T_{\text{risky}}}^{T_{\text{safe}}} e^{-\delta(\theta) t} dt f(\theta) d\theta \leq \ \varepsilon\,. \tag{3}$$

4. *The organs are not intentionally withheld:*

$$\text{At least one of the inequalities (2),(3), and } T_{\text{risky}} \geq 0 \text{ holds with equality.} \tag{4}$$

The proof of Proposition 2 is in Appendix A.

The proposition establishes that the key statistics of the equilibrium of the waiting list are the waiting times for safe and risky organs, which will allow us to conduct comparative statics. Next we establish the existence of a steady-state equilibrium.

Proposition 3 (Existence of Steady-State Equilibrium)**.** *Given a market* $(s, r, p, f(\cdot), \nu(\cdot), \delta(\cdot))$*, there exists a steady-state equilibrium.*

The proof of Proposition 3 is in Appendix A. We use Kakutani's fixed point theorem to show the existence.

In standard queuing systems, a waiting time for a good is positive only if the good is in excess demand—in particular, only if all units of the good are allocated. By contrast, due to the perishability of organs, the example illustrates that the system can enter a counterintuitive state in which the waiting time for risky organs is positive but nevertheless a risky organs are wasted (as will also be illustrated in Example 1). The following is a formal definition of congestion.

Definition 2. The steady state is *congested* if $T_{\text{risky}} > 0$ and a positive mass of risky organs are discarded.

4.3 Multiplicity and Equilibrium Selection

The following example illustrates that the steady-state equilibrium is not necessarily unique and there can be an congested equilibrium:

Example 1 (Multiple Steady-State Equilibria)**.** There are two patient types, *healthy* (θ_h) and *unhealthy* (θ_u), each arriving at rate 0.5. Both safe and risky organs arrive at rate 0.4. The departure rates are $\delta(\theta_h) = 0.15$, $\delta(\theta_u) = 0.5$, reflecting that healthy patients have a longer expected lifetime.

Patient utilities from receiving risky organs are $\nu(\theta_h) = 0.4$, $\nu(\theta_u) = 0.6$, indicating that healthy patients have a stronger preference for safe organs. Organs expire after rejection by a measure $\varepsilon = 0.46$ of patients.

Patients' willingness to wait for quality are $W_\nu(\theta_h) \approx 6.11$ and $W_\nu(\theta_u) \approx 1.02$. These values represent decision thresholds determining whether a patient type θ accepts a risky organ or waits for a safe one: patients accept the risky organ if the required additional waiting time exceeds this threshold; otherwise, they reject and wait for a safe organ.

We identify two equilibria:

- *Equilibrium 1: Uncongested Equilibrium.* Here, $T_{\text{risky}} \approx 0.45$, $T_{\text{safe}} \approx 1.49$. Since $W_\nu(\theta_u) < T_{\text{safe}} - T_{\text{risky}} < W_\nu(\theta_h)$, healthy patients reject risky organs, while unhealthy patients accept risky organs. Equation (1) is satisfied, and inequality (2) holds with equality. Moreover, the left-hand-side of (3) is approximately 0.45, smaller than ϵ; thus (3) holds, implying no organ wastage. By Proposition 2, this pair of waiting times constitutes an equilibrium.

- *Equilibrium 2: Congested equilibrium.* Here, $T_{\text{risky}} \approx 2.30$, $T_{\text{safe}} \approx 3.31$. Since $T_{\text{safe}} - T_{\text{risky}} < W_\nu(\theta_h), W_\nu(\theta_u)$, both patient types reject risky organs. Equation (1) is satisfied, and inequality (3) holds with equality, meaning that all risky organs reach their maximum cold-ischemic time. Since no patient accepts risky organs, (2) also holds. Proposition 2 verifies this is an equilibrium.

The above example of multiple equilibria illustrates an interesting *self-fulfilling* congestion phenomenon. Starting from Equilibrium 1, if unhealthy patients switch to choosing safe organs, T_{safe} does not increase significantly because they depart much faster than healthy patients. However, their shift creates congestion and raises T_{risky} enough to make risky organs unattractive to unhealthy patients, thereby sustaining the new equilibrium.[18]

While there could be multiple equilibria in this system, the next result shows that all equilibria must be Pareto ranked.

Proposition 4. *All equilibria can be Pareto ranked. If there exists an uncongested equilibrium, then it Pareto-dominates all other equilibria. Moreover, there is at most one uncongested equilibrium.*

Proof. Suppose $(T_{\text{risky}}, T_{\text{safe}})$ and $(T'_{\text{risky}}, T'_{\text{safe}})$ are both equilibrium waiting time pairs. If $T_{\text{safe}} - T_{\text{risky}} = T'_{\text{safe}} - T'_{\text{risky}}$, the patients' strategy profile stays the same, so $T_{\text{safe}} = T'_{\text{safe}}$, and hence $T_{\text{risky}} = T'_{\text{risky}}$. If $T_{\text{safe}} - T_{\text{risky}} > T'_{\text{safe}} - T'_{\text{risky}}$, then more patients choose to wait for a safe organ in the latter equilibrium. Therefore $T'_{\text{safe}} > T_{\text{safe}}$, implying that $T'_{\text{risky}} > T_{\text{risky}}$, i.e., the former equilibrium Pareto dominates the latter one.

Suppose that $(T_{\text{risky}}, T_{\text{safe}})$ strictly Pareto dominates $(T'_{\text{risky}}, T'_{\text{safe}})$. From earlier, we have $T_{\text{safe}} - T_{\text{risky}} > T'_{\text{safe}} - T'_{\text{risky}}$, so fewer patients wait for a risky organ in the latter equilibrium. Assume, for contradiction, that the latter equilibrium is uncongested. Since no risky organ is discarded in the uncongested equilibrium, we have $T'_{\text{risky}} \leq T_{\text{risky}}$, which is a contradiction. Therefore an uncongested equilibrium, if it exists, must be unique and Pareto optimal. □

Throughout the rest of the paper, we focus on the Pareto-optimal equilibrium to tie our hands and minimize the scope of the inefficiencies that we highlight. The proposition tells us that this selection rule always selects the uncongested equilibrium if one exists.

[18] This example technically violates our assumption that patient types are drawn from an atomless distribution. However, it is straightforward to construct a continuous distribution that closely approximates the two-point distribution used here while exhibiting the same phenomenon. We choose this discrete example for clarity, as it conveys the intuition in the simplest possible way.

4.4 Explicit Equilibrium Characterization with Homogeneous Departure Rates

When the departure rate of patients is homogeneous, i.e., $\delta(\theta) \equiv \delta$ for all $\theta \in \Theta$, we can establish the existence of a unique equilibrium and explicitly characterize it.

Proposition 5 (Existence and Uniqueness when $\delta(\theta) \equiv \delta$)**.** *Given a market* $(s, r, p, f(\cdot), \nu(\cdot), \delta)$*, there exists a unique equilibrium pair of waiting times* $(T_{\text{safe}}, T_{\text{risky}})$*.*

A key quantity that turns out to determine whether congestion arises is the *degree of congestion*.

Definition 3. If $\delta(\theta) \equiv \delta$, the *degree of congestion* is defined as $c = \frac{sp}{sp+\varepsilon\delta}$.

Intuitively, the degree of congestion is the probability that a patient with ε patients ahead in the queue will eventually receive a safe kidney (the patient would be offered risky organs just before they expire).

When the degree of congestion is low, patients who are offered near-expiry risky organs have little incentive to decline and wait for a safe kidney—behavior that would lead to congestion if done by many patients. We provide a simple characterization of when congestion arises and the equilibrium waiting times under congestion in terms of the degree of congestion.

Proposition 6 (Explicit Characterization of Equilibrium when $\delta(\theta) \equiv \delta$)**.** *If* $\delta(\theta) \equiv \delta$*,*

(a) The unique steady-state equilibrium is congested if and only if

$$F_\nu(c) > \max\left\{\frac{s}{c}, \frac{s}{s+rc}\right\},$$

where F_ν *is the cumulative distribution function of* $\nu(\theta)$*.*

(b) If the steady-state equilibrium is congested, the equilibrium waiting times are

$$T_{\text{risky}} = \frac{1}{\delta}\log\left(\frac{cF_\nu(c)}{s}\right) \quad \textit{and} \quad T_{\text{safe}} = \frac{1}{\delta}\log\left(\frac{F_\nu(c)}{s}\right).$$

The characterization above offers useful insights. First, part (a) suggests that, all else equal, a higher $F_\nu(\cdot)$—indicating a shift toward more conservative patient behavior—makes congestion more likely. Second, part (b) implies that an increase in $F_\nu(\cdot)$ leads to longer equilibrium waiting times for both safe and risky organs.

5 Congestion

5.1 Welfare Effects of Congestion

In this section we show that when congestion arises, it hurts all patients. Clearly, by leading to the discard of risky organs, congestion hurts those who would accept such organs. However, it turns out that congestion also affects those who decline risky organs: with a higher waiting time for risky organs, patients who are offered risky organs have a better chance of obtaining a safe organ, making congestion increase the demand for safe organs. The following proposition formalizes this intuition by analyzing the welfare effects of relaxing the cold ischemic time constraint (raising ε).

Proposition 7. *If the Pareto-optimal equilibrium is congested, then raising ε strictly raises the equilibrium utility of all types.*

Proof. Because the Pareto-optimal equilibrium is congested, only constraints (1) and (3) are binding. Define $\Delta T \triangleq T_{\text{safe}} - T_{\text{risky}}$ and rewrite T_{risky} as $T_{\text{safe}} - \Delta T$. Let $(\Delta T^*, T^*_{\text{safe}})$ denote the Pareto-optimal equilibrium prior to the increase in ϵ. The proof proceeds in two steps.

Step 1. Express T_{safe} as a function of ΔT using (1). Denote the left-hand-side of (1) by:

$$H(T_{\text{safe}}, \Delta T) \triangleq \int_{\theta: W_\nu(\theta) \geq \Delta T} e^{-\delta(\theta) T_{\text{safe}}} f(\theta)\, d\theta\,.$$

For any $\Delta T \geq 0$, define the mapping

$$h(\Delta T) \triangleq \{T_{\text{safe}} \geq 0 : H(T_{\text{safe}}, \Delta T) = s\}\ .$$

We claim that this correspondence is single-valued, continuous and non-increasing on its domain $[0, \overline{\Delta T}]$, where $\overline{\Delta T}$ is uniquely determined by (1) and $T_{\text{risky}} = 0$. The proof of this claim is in Appendix B.

Step 2. Proving the proposition with monotone comparative statics. Substituting $T_{\text{safe}} = h(\Delta T)$ into (3), we can rewrite it as a fixed point equation:

$$\varepsilon - p \int_{\theta: W_\nu(\theta) \geq \Delta T} \frac{1}{\delta(\theta)} \left(e^{\delta(\theta)\Delta T} - 1\right) e^{-\delta(\theta) h(\Delta T)} f(\theta)\, d\theta + \Delta T = \Delta T\,, \tag{5}$$

Denote the left-hand-side by $F(\Delta T, h(\Delta T), \epsilon)$. We restrict attention to the largest fixed point ΔT of (5), equivalently, the smallest T_{safe}, as we focus on the Pareto-optimal equilibrium. Specifically, we consider a γ-neighborhood of $(\Delta T^*, T^*_{\text{safe}})$ for small $\gamma > 0$, where γ is chosen

such that all points in this neighborhood strictly satisfy inequality (2) (by continuity of the left-hand side in both ΔT and T_{safe}) and the constraint $T_{\text{safe}} - \Delta T > 0$.

The following conditions, necessary for the application of Corollary 3 in Milgrom and Roberts (1994), can be easily verified:

1. For each T_{safe}, ϵ, H is continuous in ΔT.
2. For all T_{safe}, ϵ, H is continuous in T_{safe}.
3. For ΔT, T_{safe}, H is strictly increasing in ϵ.

Here $T_{\text{safe}} = h(\Delta T)$ where $h(\cdot)$ is a continuous function. It follows that the largest fixed point is increasing in ϵ. Since T_{safe} is decreasing in ΔT by step 1, it follows that T_{safe} is decreasing in ϵ. Moreover, because $T_{\text{risky}} = T_{\text{safe}} - \Delta T$, it follows that T_{safe} is decreasing in ϵ as well. This concludes the proof. □

6 Constrained Inefficiency of Congested Waiting Lists

Intuitively, increasing the pool of safe organs should improve welfare. If not, then the equilibrium is constrained inefficient in that discarding safe organs might improve welfare even without changing the underlying queuing structure. In this section, we show that this issue can arise if the steady state is congested, but (as is intuitive) not if the steady state is uncongested.

We first show that "artificially discarding" safe kidneys can never help any patient without congestion.

Proposition 8. *If the steady state is not congested, then raising s:*

(a) weakly raises the steady-state expected utility of all types of patients and

(b) strictly raises the steady-state expected utility of types that obtain safe organs with positive probability.

Proof. We can drop constraint (3), and focus on the remaining three constraints. We have

$$\begin{aligned} \int_{\theta: W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta) T_{\text{safe}}} f(\theta) d\theta = \ & s \\ \int_{\theta: W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta) T_{\text{risky}}} f(\theta) d\theta \leq \ & r \\ T_{\text{risky}} \geq 0 & \end{aligned}$$

Note that the set of $(\Delta T, T_{\text{safe}})$ satisfying the first and third constraints is a segment of continuous, non-increasing curve $(\Delta T, T_s(\Delta T))$ on interval $[0, \overline{\Delta T}]$.

Now we determine how an increase in s change this curve. For fixed ΔT, when s increases in the first equation, T_{safe} needs to become smaller to make equation hold. As a result, the curve moves downward. Since $\overline{\Delta T}$ solves $\Delta T = T_s(\Delta T)$, we have that $\overline{\Delta T}$ moves to the left.

Note that the feasible set of the second constraint does not change with s. If $T_{\text{risky}} \geq 0$ is binding, we can plot the intersection between $T_{\text{safe}} = \Delta T$ and $T_{\text{safe}} = T_s(\Delta T)$ the intersection moves to the left and down, meaning T_{safe} becomes smaller, and T_{risky} stays at zero. The proposition is proved.

If $T_{\text{risky}} \geq 0$ is not binding, it must be that the second inequality is binding. We can plot the curve for the second constraint when it holds with equality. When ΔT becomes larger, T_{risky} needs to become larger for the equality to hold, meaning T_{safe} also becomes larger. As a result, the curve corresponding to the equality version of the second constraint is a non-decreasing curve. When s increases, the curve of the first constraint moves downward, hence the intersection of the first two constraints move left and down. This means that T_{safe} decreases and ΔT decreases. Given the second constraint, T_{risky} need to decrease. This proves the proposition. □

The following example shows that when the state is congested, discarding safe organs may result in a Pareto improvement; alternatively, the arrival of more safe organs could harm all patients.

Example 2 (**Pareto improvement when discarding safe organs).** Suppose that $\delta(\theta) \equiv \delta$ for all types θ, and that there exists a constant $\alpha \geq 2$ such that $F_\nu(x) = x^\alpha$ for $0 \leq x \leq 1$. If $\varepsilon < \frac{sp}{\delta} \min\left\{\alpha - 1, s^{\frac{-1}{\alpha+1}} - 1\right\}$ and $r > 1 - \frac{s}{c}$, then increasing the arrival rate s to any $s' \leq \min\{c^{\alpha+1}, (1-r)c\}$ strictly lowers the steady-state expected utility of all types.

Proof. Since departure rates are homogeneous, Proposition 5 shows that the equilibrium not only exists but also is unique. The bound on ε implies that $c > \max\{\frac{1}{\alpha}, s^{\frac{1}{\alpha+1}}\}$. Moreover, Proposition 6 implies that the equilibrium is congested if and only if $F_\nu(c) > \max\left\{\frac{s}{c}, \frac{s}{s+rc}\right\}$. In this example, we have

$$F_\nu(c) = c^\alpha\,, \quad \frac{s}{c} < \frac{c^{\alpha+1}}{c} = c^\alpha\,, \quad s + rc \leq c\,,$$

therefore the equilibrium is indeed congested. Proposition 6 then implies that

$$T_{\text{risky}} = -\frac{1}{\delta}\log\left(\frac{s}{cF_\nu(c)}\right)\,, \quad T_{\text{safe}} = -\frac{1}{\delta}\log\left(\frac{s}{F_\nu(c)}\right)\,.$$

Plugging the form of $F_\nu(\cdot)$ into the above expression, we have

$$e^{-\delta T_{\text{risky}}} = \frac{s}{c^{\alpha+1}}, \quad e^{-\delta T_{\text{safe}}} = \frac{s}{c^{\alpha}}.$$

Note that $\frac{dc}{ds} = \frac{c}{s} - \frac{c^2}{s}$, we have

$$\frac{de^{-\delta T_{\text{risky}}}}{ds} = \frac{c^{\alpha+1} - s(\alpha+1)c^{\alpha}\left(\frac{c}{s} - \frac{c^2}{s}\right)}{c^{2\alpha+2}} = \frac{1-(\alpha+1)(1-c)}{c^{\alpha+1}} \leq \frac{1-(\alpha+1)(1-\frac{1}{\alpha})}{c^{\alpha+1}} < 0\,.$$
$$\frac{de^{-\delta T_{\text{safe}}}}{ds} = \frac{c^{\alpha} - s\alpha c^{\alpha-1}\left(\frac{c}{s} - \frac{c^2}{s}\right)}{c^{2\alpha}} = \frac{1-\alpha(1-c)}{c^{\alpha}} < \frac{1-\alpha(1-\frac{1}{\alpha})}{c^{\alpha}} < 0\,,$$

where the last step of each line uses $\alpha > 2$. This concludes the proof. □

For intuition, when the system is congested, an increase in safe organs results in more types waiting for safe organs. This can exacerbate congestion so severely that it increases waiting times for everyone.

While our example is stylized, it illustrates the severity of the congestion externality. The externality can be so severe that a benevolent social planner who is forced to used a queuing system might in principle prefer to destroy safe organs. This result shows that the queuing system is constrained inefficient under congestion, in that the inefficiency could be mitigated by taking a feasible action without changing the underlying structure of the queuing system (or priorities). In the remainder of the paper, we show how the congestion externality is relevant to other features of the market and studying the welfare effects of alternative waiting list policies.

7 Moral Hazard and the Congestion Externality

In practice, acceptance and rejection decisions are delegated to the patient's hospital, which may have less incentives to transplant risky organs. To allow for the possibility of moral hazard in this delegation problem, we suppose that acceptance and rejection of offered organs is instead decided based on the (Bernoulli) utility function

$$u_{\text{decision}} = \begin{cases} 1 & \text{if get safe organ} \\ \lambda(\theta) & \text{if get risky organ} \\ 0 & \text{if do not get an organ} \end{cases}.$$

where $\lambda(\theta) \in [0,1]$ is the effective willingness to accept a risky organ. Throughout, we assume that $\lambda(\theta)$ has distribution that is absolutely continuous on $(0,1]$.

Similar logic to Section 3.2 shows that the transplant center's willingness to wait for quality for a patient of type θ is

$$W_\lambda(\theta) = \frac{-\log \lambda(\theta)}{\delta}.$$

As transplant centers are penalized for adverse post-transplant outcomes but not deaths on the waiting list (Mehrotra et al., 2022; Sepulveda et al., 2023; Chan and Roth, 2024; Mildebrath et al., 2024), we assume that hospitals are less willing to accept risky organs than patients would be themselves. That is, we assume that $\nu(\theta) > \lambda(\theta)$. We then show that if the steady state with patient decisions is congested, decisions being made by hospitals hurts all patients, and even hospitals in aggregate. Intuitively, under this assumption, hospitals' risk aversion worsens congestion enough to cause substantial harm.

Proposition 9. *Suppose $\nu(\theta) > \lambda(\theta)$ for almost all θ. If the steady state is congested under patient decisions:*

(a) The steady state will also be congested under hospital decisions.

(b) All types of patients have lower expected utility under hospital decisions, and strictly lower expected utility if the distribution of ν has full support on $[0,1]$.

(c) The total expected utility of hospitals is weakly lower under hospital decisions than under patient decisions, with equality only if $\Pr_\theta[\lambda(\theta) = 0] = 1$.

Proof. We first check by moving from patient decisions to hospital decisions, how do the curves move.

The first constraint, i.e., the one corresponding to safe organs. For fixed Δ, since the willingness-to-wait increases, T_{safe} need to increase to keep the equality hold. Hence the curve moves upward. $\overline{\Delta T}$ also moves upward.

We ignore the second constraint since it is not binding, because we consider a congested equilibrium.

For the third constraint, i.e., the congestion constraint. For fixed ΔT, the integration area increases. As a result, the integrand has to decrease, meaning both T_{safe} and T_{risky} needs to increase, the curve $\varepsilon - H_3$ moves downward.

Using similar analysis as earlier, we look at the Pareto-optimal equilibrium. Because we are looking at a congested equilibrium, so $g(\overline{\Delta T}) < 0$. As a result, when the curve moves

downward, its largest zero decreases. As a result, Δ decreases, hence T_{safe} increases. Because the T_s curve moves up, the new T_{safe} increases. Because ΔT decreases, T_{risky} also moves up.

For patients that choose the same type of organs under both patient choice and hospital choice, they are weakly worse off. For the remaining patients that choose safe organs under hospital choice and risky organs under own choice, their change of utility is:

$$\frac{u'}{u} = \frac{e^{-\delta(\theta)T'_{\text{safe}}}}{\nu(\theta)e^{-\delta T_{\text{risky}}}} \leq \frac{e^{-\delta(\theta)T'_{\text{safe}}}}{e^{-\delta(\theta)T_{\text{safe}}}} \leq 1\,.$$

The first inequality follows from the fact that the patient's utility is maximized when choosing risky organs when under their own choice. The second inequality is strict if ν has full support on $[0,1]$.

For the third statement, look at the hospital utility:

$$\int_{\theta:W_\nu(\theta)<T_{\text{safe}}-T_{\text{risky}}} e^{-\delta(\theta)T_{\text{risky}}}\lambda(\theta)f(\theta)d\theta\,.$$

When ν becomes λ, ΔT decreases, $W_\cdot(\theta)$ increases, so the integration area decreases. Since T_{risky} increases, the integrand also decreases. As a result, the total utility decreases. □

The proof shows that the first two parts of the proposition only rely on ν (first-order) stochastically dominating λ: i.e., that hospitals are more risk-averse than patients in a distributional sense.

The congestion externality plays a crucial role in Proposition 9. Suppose ε is large enough such that there is no congestion. Then the expected utility of patients who (still) choose to wait for risky organs even under hospital decision-making will increase from the reduction in demand for risky organs that arises from hospitals' risk aversion.

8 Market Design Implications

Our framework and findings can help understanding congestion externalities under different market designs.

8.1 Market Thickness

In March 2021 the Organ Procurement and Transplantation Network (OPTN) has increased the regions in which patients are given a (local) geographic priority.[19] At the same year, the

[19]Instead of a "small" fixed service area patients within 250NM from the donor hospital are given priority.

discard rate increased by 3%. Motivated by this, we are interested in comparative statics with respect to the market thickness p. Mathematically, this is equivalent to reducing ε, so Proposition 7 provides comparative statics.

Proposition 10. *(a) Increasing p weakly lowers the steady state expected utility of all types of patients.*

(b) If the steady state is congested, then increasing p strictly lowers the expected utility of all types of patients.

We note that following the modification of local prioritization in March 2021, the discard rate has increased especially for low quality organ (high KDPI) as depicted in Figure 5 and as aligned with our theory.

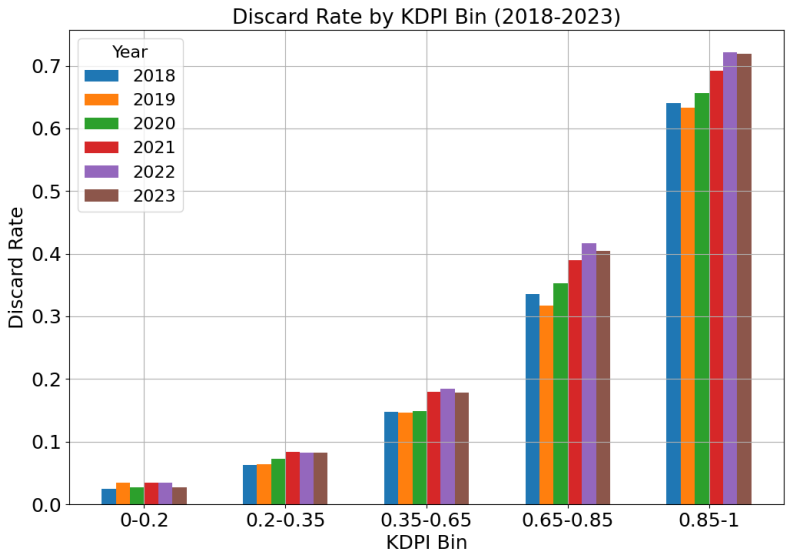


Figure 5: Annual trends of discard rate by KDPI

8.2 Offering Processes

In the U.S. a deceased donor kidney is offered sequentially to patients based on a fixed priority order. While high quality kidneys are accepted quickly, kidneys with marginal quality are declined by many patients (see Figure 3). Guan et al. (2025) find that refusals due to long cold-ischemic time (CIT) increase substantially 4 hours after recovery and becomes the main refusal reason 10 hours after recovery (see also Figure 4).[20] The large non-utilization rate

[20]Guan et al. (2025) also find that 4 hours after recovery, an organ has been offered to patients from only 11 transplant centers (out of 220).

has led Organ Procurement Organizations (OPOs) to expedite the offering process and in 2023 more than 20% of accepted organs have been placed out-of-sequence (Liyanage et al., 2025).

It is straightforward to see that in partial equilibrium (i.e., not taking into account effects on patients' decisions and equilibrium waiting times), expediting raises welfare by allocating risky organs to patients who would otherwise not receive an organ. However, expediting many organs will also have general equilibrium effects due to patients' decisions and equilibrium waiting times. Our model allows us to formally study these broader impacts.

Consider the following modification for the FCFS offering policy, denote by Π: Safe organs are offered in an FCFS manner and risky organs are offered in sequence to patients who have yet to declined a safe organ.[21] Observe that this policy achieves the same waiting times for any ϵ for safe and risky organs. Since patients do not change their decisions in the fluid limit:

Proposition 11. *The offering policy* Π *achieves the same welfare as in the FCFS policy when* $\epsilon = \infty$*:*

9 Discussion

This paper finds that the perishability of organs (due to long cold ischemic times) leads to congestion on deceased donor waiting list, causing substantial inefficiencies and externalities. In a congested waiting list not all risky organs are not utilized since they do not reach fast enough patients who are willing to accept them. Our findings shed light on policy recommendations.

When the waitlist is congested, thickening the market by increasing the service region may increase congestion even more. In such situations expediting offers of low quality organs allows increase their utilization.[22]Castro et al. (2021) show the effectiveness of this approach can improve the utilization of unattractive rides in ride-sharing. In the UK organs that have been declined more than 3 times, are offered to all patients simultaneously (Callaghan et al., 2017). In the US OPOs have been expediting organs in an ad-hoc manner (Jadlowiec et al., 2023) by making open offers to transplant centers. This approach has raised equity concerns and the Health Resources and Services Administration (HRSA) has directed the OPTN to create a standard policy for expedited placements.

[21]This is similar to a buffer queue as studied in (Leshno, 2022).

[22]We note that thickness may benefit patients when there is preferences exhibit large horizontal heterogeneity, for example to due to level of compatibility (Agarwal et al., 2021).

Another approach to reduce congestion is eliciting more refined preferences from doctors and patients in order to avoid unnecessary offers. While currently it is possible to indicate which types of offers to filter, there number of refusals remains very large (Guan et al., 2025).

To reduce the inefficiencies arising due moral hazard, one can attempt to involve patients in rejection decisions (Mohan and Husain, 2022). Patients who recently joined the list may be interested in accepting low quality organs (Mehrotra et al., 2022; Sepulveda et al., 2023). One may create dedicated lists for risky organs (for example patients have to consent for receiving offers of KDPI larger than 85%). Another approach is to relax the disincentives for doctors to accept risky organs.

The paper analyses the inefficiencies that can arise due to the possibility of congestion. A detailed investigation of policies that alleviate congestion requires further research.

A Proofs in Section 4

A.1 Proof of Proposition 1

Define

$$T_{\text{safe}} \triangleq \sup_{t \in T} \left\{ t : \int_{\theta \in \Theta} a(t, \theta)\, d\theta > 0 \right\},$$

the waiting time of the longest-waiting agent.

Claim. At time T_{safe} all safe organs are allocated and no risky organ is:

$$\int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, s)\, d\theta = sp\,, \text{ and } \int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, r)\, d\theta = 0\,.$$

Proof of the claim. For safe organs, suppose, to the contrary, that

$$\int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, s)\, d\theta < sp\,.$$

Because $a(\cdot, \theta)$ is weakly decreasing for all $\theta \in \Theta$, there can be no $t' < T_{\text{safe}}$ such that $\int_{\theta \in \Theta} a(t', \theta)\, d\theta = 0$; otherwise T_{safe} would not be the supremum.

It follows from the assumption that $\exists \gamma > 0$ such that $b(t, s) > 0$ for all $t \in (T_{\text{safe}} - \gamma, T_{\text{safe}})$. Indeed, if this were not the case, then $b(t', s) = 0$ for all $t' < T_{\text{safe}}$ because $b(\cdot, s)$ is weakly increasing, which will violate the conservation law for organs at T_{safe}:

$$\lim_{t' \to T_{\text{safe}}^-} b(t', s) = b(T_{\text{safe}}, s) - \int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, s)\, d\theta > 0\,.$$

It follows that for all $t \in (T_{\text{safe}} - \gamma, T_{\text{safe}})$, $\int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, s)\, d\theta > 0$. Thus, for every such t, there is a jump discontinuity in $b(\cdot, s)$ at point t. This, however, contradicts the fact that $b(\cdot, s)$ is monotone, since any monotone function has countably many jump discontinuities, by Froda's theorem. This proves that $\int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, s)\, d\theta = sp$.

For risky organs, assume, for contradiction, that

$$\int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, r)\, d\theta > 0\,.$$

Then, there exists $\theta \in \Theta$ such that $m(T_{\text{safe}}, \theta, r) > 0$. This contradicts the optimality of the decision rule D_θ for agents of such type θ; such an agent can increase her payoff by switching

to a new decision rule D'_θ defined by

$$D'_\theta(t) = \begin{cases} D_\theta(t) & \text{for } t \leq T_{\text{safe}} - \delta, \\ s & \text{for } t > T_{\text{safe}} - \delta, \end{cases}$$

where δ is a sufficiently small positive constant. The claim follows.

It follows from the claim that all safe organs are allocated at T_{safe}: by the conservation law for organs:

$$\lim_{t' \to T_{\text{safe}}^-} b(t', s) = b(T_{\text{safe}}, s) - \int_{\theta \in \Theta} m(T_{\text{safe}}, \theta, s)\, d\theta = 0\,,$$

and because $b(t, s)$ is weakly increasing, it must be that $b(t, s) = 0$ for all $t \in [0, T_{\text{safe}})$.

Moreover, there exists $\delta > 0$ such that no risky organ is allocated on the interval $(T_{\text{safe}} - \delta, T_{\text{safe}}]$. Define the longest waiting time for risky organs by

$$T_{\text{risky}} \triangleq \sup_{t \in T} \left\{ t : \int_{\theta \in \Theta} m(t, \theta, r) > 0 \right\},$$

which satisfies $T_{\text{risky}} \leq T_{\text{safe}} - \delta$. There are three cases.

- *Case 1*: $T_{\text{risky}} = 0$. By definition, no risky organ is allocated at a positive waiting time.
- *Case 2*: $T_{\text{risky}} > 0$ and the cold-ischemic time limit is binding for risky organs at T_{risky}:

  $$\int_{t=T_{\text{risky}}}^{T_{\text{safe}}} b(t, r) dt = \epsilon\,.$$

  If $b(t, r)$ were positive on interval $(T_{\text{risky}} - \delta, T_{\text{risky}})$, the cold ischemic time limit would be violated; hence risky organs are allocated only at T_{risky}.
- *Case 3:* $T_{\text{risky}} > 0$ and the cold-ischemic time limit is not binding at T_{risky}:

  $$\int_{t=T_{\text{risky}}}^{T_{\text{safe}}} b(t, r) dt < \epsilon\,.$$

  The same argument used at T_{safe} implies $\int_{\theta \in \Theta} m(T_{\text{risky}}, \theta, r) d\theta = rp$, and that $b(t', r) = 0$ for all $t' < T_{\text{risky}}$. Thus risky organs are allocated exclusively at T_{risky}.

These three cases establish that no risky organ is allocated outside time T_{risky}, completing the proof of Proposition 1.

A.2 Proof of Proposition 2

Proof of "if" part. Assume a pair of waiting times T_{safe} and T_{risky} satisfying these conditions. We construct a tuple $(A, b, m, \{D_\theta\}_{\theta\in\Theta})$ and verify that it fulfills Definition 1.

Agent densities.

$$a(t,\theta) \triangleq \begin{cases} f(\theta)\cdot e^{-\delta(\theta)}\cdot \mathbb{1}\{0 \le t \le T_{\text{safe}}\}\,, & \text{if } W_\nu(\theta) \ge T_{\text{safe}} - T_{\text{risky}}\,, \\ f(\theta)\cdot e^{-\delta(\theta)}\cdot \mathbb{1}\{0 \le t \le T_{\text{risky}}\}\,, & \text{if } W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}\,. \end{cases}$$

Offer rates.

$$b(t,s) \triangleq sp\cdot \mathbb{1}\{t \ge T_{\text{safe}}\}\,,$$
$$b(t,r) \triangleq rp\cdot \mathbb{1}\{t \ge T_{\text{risky}}\}\,.$$

Allocation rates.

$$m(t,\theta,s) \triangleq \begin{cases} f(\theta)\cdot e^{-\delta(\theta)t}\,, & \text{if } W_\nu(\theta) \ge T_{\text{safe}} - T_{\text{risky}} \text{ and } t = T_{\text{safe}}\,, \\ 0\,, & \text{otherwise}\,, \end{cases}$$

$$m(t,\theta,r) \triangleq \begin{cases} f(\theta)\cdot e^{-\delta(\theta)t}\,, & \text{if } W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}} \text{ and } t = T_{\text{risky}}\,, \\ 0\,, & \text{otherwise}\,, \end{cases}$$

Agents' decision rules.

$$D_\theta(t) \equiv \begin{cases} s\,, & \text{if } W_\nu(\theta) \ge T_{\text{safe}} - T_{\text{risky}}\,, \\ r\,, & \text{if } W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}\,. \end{cases}$$

Verifying that the constructed tuple meets every requirement of Definition 1 is straightforward, and this completes the sufficiency part of the proof.

Proof of "only if" part. Let $(A, b, m, \{D_\theta\}_{\theta\in\Theta})$ be a steady-state equilibrium as defined in Definition 1. Proposition 1 then guarantees a pair of cutoff waiting times $(T_{\text{safe}}, T_{\text{risky}})$,

which can be characterized as follows. Define $T_{\text{safe}}, T'_{\text{risky}}, T_\varepsilon$ as

$$T_{\text{safe}} \triangleq \inf\left\{T \geq 0 : \int_{\theta: W_\nu(\theta) \geq T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta)T} f(\theta)\, d\theta \leq s\right\},$$

$$T'_{\text{risky}} \triangleq \inf\left\{T \geq 0 : \int_{\theta: W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta)T} f(\theta)\, d\theta \leq r\right\},$$

$$T_\varepsilon \triangleq \inf\left\{T \geq 0 : \int_{\theta: W_\nu(\theta) \geq T_{\text{safe}} - T_{\text{risky}}} \int_T^{T_{\text{safe}}} e^{-\delta(\theta)t} f(\theta)\, dt d\theta \leq \varepsilon\right\},$$

and $T_{\text{risky}} \triangleq \max\{T'_{\text{risky}}, T_\varepsilon\}$.

We comment on the characterization of waiting times. If the number of patients opting for the safe organ exceeds the available supply, a positive waiting time T_{safe} will arise, so that the supply exactly matches the demand after a wait of T_{safe}. A similar rationale applies to the definition of T'_{risky}. However, due to congestion, there exists another cutoff, T_ε, representing the maximum depth a risky organ can reach in the waiting list (where a shorter time in the system corresponds to a deeper position in the list). The actual waiting time T_{risky} is the larger of T'_{risky} and T_ε.

Before proceeding with the proof, first note that we must have $T_{\text{risky}} \leq T_{\text{safe}}$ in equilibrium. Assume, instead, that $T_{\text{risky}} > T_{\text{safe}}$. Then everyone will choose safe organ to maximize utility, resulting in $T_{\text{safe}} > 0$ (because $s > 1$), while in the mean time $T_{\text{risky}} = 0$ since no one chooses risky organ. This contradicts the assumption. This ensures the integrals in the statement of the proposition are well defined.

As argued in Section 3.2, patient θ will accept risky organ if and only if $T_{\text{safe}} - T_{\text{risky}} \geq W_\nu(\theta)$.

We now prove (1). We can prove by contradiction. Assume not all safe organs are allocated, then $T_{\text{safe}} = 0$. Therefore, all patients choose safe organs. Since there is shortage of safe organs ($s < 1$), we have $T_{\text{safe}} > 0$. There is contradiction.

(2) and (3) simply follows the definitions of T'_{risky} and T_ε.

Finally, to prove (4), assume that $T_{\text{risky}} > 0$, then either $T'_{\text{risky}} > 0$ or $T_\varepsilon > 0$. If $T'_{\text{risky}} > 0$, then (2) holds with equality, otherwise it contradicts the definition of T'_{risky}. If $T_\varepsilon > 0$, then (3) holds with equality, otherwise it contradicts the definition of T_ε. This concludes the proof.

A.3 Proof of Proposition 3

Proof. For notational simplicity, define:

$$f_r(\theta) \triangleq f(\theta) \cdot \mathbb{1}\{W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}\}, \quad f_s(\theta) \triangleq f(\theta) \cdot \mathbb{1}\{W_\nu(\theta) \geq T_{\text{safe}} - T_{\text{risky}}\}.$$

Using the notation in the proof of Proposition 2, denote the mapping from strategy profile to waiting times as $\mathbf{T}(f_r, f_s)$. Note that (f_r, f_s) constitutes an equilibrium if patients' best response to $\mathbf{T}(f_r, f_s)$ results in the same waiting time $\mathbf{T}$. We define $\Gamma(\mathbf{T})$ as the waiting time pair induced by patients' best response to $\mathbf{T}$:

$$\begin{aligned}
\gamma_s(\mathbf{T}) &\triangleq \inf\left\{T \geq 0 : \int_{\theta: W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta)T} f(\theta) d\theta \leq s\right\}, \\
\gamma'_r(\mathbf{T}) &\triangleq \inf\left\{T \geq 0 : \int_{\theta: W_\nu(\theta) < T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta)T} f(\theta) d\theta \leq r\right\}, \\
\gamma_\varepsilon(\mathbf{T}) &\triangleq \inf\left\{T \geq 0 : \int_{\theta: W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}} \int_T^{T_{\text{safe}}} e^{-\delta(\theta)t} f(\theta) dt d\theta \leq \varepsilon\right\},
\end{aligned}$$

and we define $\gamma_r(\mathbf{T}) \triangleq \max\{\gamma'_r(\mathbf{T}), \gamma_\varepsilon(\mathbf{T})\}$, $\Gamma(\mathbf{T}) \triangleq (\gamma_r(\mathbf{T}), \gamma_s(\mathbf{T}))$.

In the following, we show that the mapping $\mathbf{T} \mapsto \Gamma(\mathbf{T})$ has a fixed point using Kakutani's fixed point theorem. We verify the conditions in Kakutani's fixed point theorem one by one.

The set of $\mathbf{T}$ *is non-empty, compact, and convex subset of* $\mathbb{R}^2$. Recall that $\underline{\delta} \triangleq \min_{\theta \in \Theta} \delta(\theta) > 0$. Therefore, for $T \geq -\frac{\log s}{\underline{\delta}}$, we have $\int_{\theta: W_\nu(\theta) > T_{\text{safe}} - T_{\text{risky}}} e^{-\delta(\theta)T} f(\theta) d\theta < s$, hence $\gamma_s(\mathbf{T}) \in [0, -\frac{\log s}{\underline{\delta}}]$ for any $\mathbf{T}$. Similarly, we have $\gamma'_r(\mathbf{T}) \in [0, -\frac{\log r}{\underline{\delta}}]$ for any $\mathbf{T}$, $\gamma_\varepsilon(\mathbf{T}) \leq T_{\text{safe}}$. Therefore, we can choose the set of $\mathbf{T} \in [0, -\frac{\log s}{\underline{\delta}}] \times [0, -\frac{\log r}{\underline{\delta}}]$, which is non-empty, compact, and convex, and ensures the that $\Gamma(\mathbf{T})$ also falls into this set.

The mapping $\Gamma(\cdot)$ *has a closed graph.* This is true because: the integral in the definition of $\Gamma_s(\mathbf{T})$, the integral in the definition of $\Gamma'_r(\mathbf{T})$, and the integral in the definition of $\Gamma_\varepsilon(\mathbf{T})$ are continuous in $T_{\text{safe}}, T_{\text{risky}}$ since $\nu(\theta)$ is absolutely continuous and $f(\theta)$ is atomless.

$\Gamma(\mathbf{T})$ *is non-empty and convex.* Note that the integral in the definition of $\gamma_s(\mathbf{T})$ is non-increasing in T. Since we have derived above that when $T = -\frac{\log s}{\underline{\delta}}$, the inequality in the definition of $\gamma_s(\mathbf{T})$ holds, therefore there must exists $\gamma_s(\mathbf{T}) \in [0, -\frac{\log s}{\underline{\delta}}]$, and $\gamma_s(\mathbf{T})$ is a point, therefore convex. Similar argument holds for $\gamma'_r(\mathbf{T})$ and $\gamma_\varepsilon(\mathbf{T})$.

This concludes the proof.

□

A.4 Proof of Proposition 5

Proof. Note that by our definition of equilibrium, for any market there are three candidate sets of equilibria:

- *Candidate 1.* There are abundant risky organs, hence $T_{\text{risky}} = 0$.
- *Candidate 2.* $T_{\text{risky}} > 0$, but no risky organ is discarded due to congestion, hence constraint (2) is binding.
- *Candidate 3.* $T_{\text{risky}} > 0$, and some risky organs are discarded due to congestion, hence constraint (3) is binding.

We call the first two candidate set of equilibria the *uncongested equilibria*, and the third candidate set of equilibria the *congested equilibria*.

As a first step, we first calculate explicitly all three candidate sets of equilibria. It turns out each candidate set contains exactly one candidate equilibrium.

Candidate equilibria 1. The set of candidate equilibria is the solution of the following two equations:

$$\begin{aligned} T_{\text{risky}} &= 0\,, \\ F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})pe^{-\delta T_{\text{safe}}} &= sp\,. \end{aligned}$$

The solution is:

$$T_{\text{safe}} = -\frac{1}{\delta}\log(\phi(s))\,,$$

where $\phi(s)$ is the unique solution of equation:

$$F_\nu(y) = \frac{s}{y}$$

on $(0,1)$. The solution is unique because LHS is non-decreasing in y on $(0,1)$, RHS is strictly decreasing in y on $(0,1)$, and that LHS<RHS for y close enough to 0, LHS>RHS at $y = 1$.

Now we determine the conditions on model parameters under which this is indeed an equilibrium. It suffices to look at the conditions for inequalities (2) and (3) to hold. For (2), we need the input to satisfy

$$\left(1 - F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})\right) \cdot e^{-\delta T_{\text{risky}}} = 1 - F_\nu(\phi(s)) < r\,.$$

This is equivalent to $s > (1-r)\phi(s)$. Note that by definition of $\phi(s)$, we have $x > \phi(s) \Leftrightarrow F_\nu(x) > \frac{s}{x}$. Therefore the above condition is equivalent to

$$r \geq 1 \quad \text{or} \quad F_\nu\left(\frac{s}{1-r}\right) > 1-r\,. \tag{6}$$

For (3) to hold, we need

$$F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})(e^{-\delta T_{\text{risky}}} - e^{-\delta T_{\text{safe}}}) = F_\nu(e^{-\delta T_{\text{safe}}})(1-e^{-\delta T_{\text{safe}}}) < \frac{\varepsilon\delta}{p}\,,$$

which is equivalent to

$$\frac{s}{\phi(s)}(1-\phi(s)) < \frac{\varepsilon\delta}{p} \quad \Leftrightarrow \quad \phi(s) > \frac{sp}{sp+\varepsilon\delta}\,,$$

which is equivalent to

$$F_\nu\left(\frac{sp}{sp+\varepsilon\delta}\right) < \frac{sp+\varepsilon\delta}{p}\,. \tag{7}$$

Therefore, this candidate equilibrium is an actual equilibrium if and only if (6) and (7) hold.

Candidate equilibria 2. The set of candidate equilibria is the solution of the following two equations:

$$\begin{aligned} F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})pe^{-\delta T_{\text{safe}}} &= sp\,,\\ \left(1-F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})\right)pe^{-\delta T_{\text{risky}}} &= rp\,. \end{aligned}$$

Let $\gamma(s,r) \triangleq e^{-\delta(T_{\text{safe}}-T_{\text{risky}})}$. Then $\gamma(s,r)$ equals the unique solution of equation:

$$\frac{F_\nu(\gamma)}{1-F_\nu(\gamma)} = \frac{s}{r}\frac{1}{\gamma}\,.$$

The solution is unique because LHS is non-decreasing in γ on $(0,1)$, RHS is strictly decreasing in γ on $(0,1)$, and that LHS<RHS for small enough $y \in (0,1)$, LHS>RHS at for γ close enough to 1. The candidate equilibrium waiting times are:

$$T_{\text{safe}} = -\frac{1}{\delta}\log\left(\frac{s}{F_\nu(\gamma(s,r))}\right)\,, \quad T_{\text{risky}} = -\frac{1}{\delta}\log\left(\frac{r}{1-F_\nu(\gamma(s,r))}\right)\,.$$

For this candidate equilibrium to be an actual equilibrium, we need $T_{\text{safe}}, T_{\text{risky}} \geq 0$ and

inequality (3) to hold. First note that

$$\gamma(s,r) > x \quad \Leftrightarrow \quad \frac{F_\nu(x)}{1-F_\nu(x)} < \frac{s}{rx}\,.$$

Therefore

$$T_{\text{safe}} \geq 0 \quad \Leftrightarrow \quad F_\nu(\gamma(s,r)) \geq s \quad \Leftrightarrow \quad \frac{s}{1-s} < \frac{s}{rF_\nu^{-1}(s)}\,,$$

which is equivalent to

$$s < F_\nu\left(\frac{1-s}{r}\right)\,. \tag{8}$$

Similarly, we have

$$T_{\text{risky}} \geq 0 \quad \Leftrightarrow \quad 1-r > F_\nu\left(\frac{s}{1-r}\right) \text{ and } r < 1\,. \tag{9}$$

For (2) to hold, we need the input to satisfy

$$F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})(e^{-\delta T_{\text{risky}}} - e^{-\delta T_{\text{safe}}}) \leq \frac{\varepsilon\delta}{p} \quad \Leftrightarrow \quad F_\nu(\gamma(s,r)) \leq \frac{\varepsilon\delta + sp}{\varepsilon\delta + sp + rp}\,,$$

which is equivalent to

$$\frac{\varepsilon\delta + sp}{\varepsilon\delta + sp + rp} > F_\nu\left(\frac{sp}{\varepsilon\delta + sp}\right)\,. \tag{10}$$

Therefore, this candidate equilibrium is an actual equilibrium if and only if (8), (9), and (10) hold.

Candidate equilibria 3. The set of candidate equilibria is the solution of the following two equations:

$$F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})pe^{-\delta T_{\text{safe}}} = sp\,,$$
$$F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})p\int_{T_{\text{risky}}}^{T_{\text{safe}}} e^{-\delta t} = \varepsilon\,.$$

Simplifying the second equation leads to:

$$F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})(e^{-\delta T_{\text{risky}}} - e^{-\delta T_{\text{safe}}}) = \frac{\varepsilon\delta}{p}\,.$$

We apply a change of variables: $x \triangleq e^{-\delta T_{\text{risky}}}, y \triangleq e^{-\delta T_{\text{safe}}}$. Now we have:

$$F_\nu(y/x)y = s\,, \quad F_\nu(y/x)(x-y) = \frac{\varepsilon\delta}{p}\,.$$

This system can be solved explicitly. We have:

$$x = \frac{\varepsilon\delta + sp}{pF_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right)}\,, \quad y = \frac{s}{F_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right)}\,.$$

Therefore,

$$T_{\text{risky}} = -\frac{1}{\delta}\log\left(\frac{\varepsilon\delta + sp}{pF_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right)}\right)\,, \quad T_{\text{safe}} = -\frac{1}{\delta}\log\left(\frac{s}{F_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right)}\right)\,.$$

For this to be equilibrium, we need $T_{\text{risky}} \geq 0$, i.e.,

$$F_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right) > \frac{\varepsilon\delta+sp}{p}\,, \tag{11}$$

and inequality (2) to hold, i.e.,

$$\left(1 - F_\nu(e^{-\delta(T_{\text{safe}}-T_{\text{risky}})})\right)pe^{-\delta T_{\text{risky}}} \leq rp\,.$$

The latter is equivalent to

$$\left(1 - F_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right)\right)\frac{\varepsilon\delta+sp}{pF_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right)} \leq r \quad \Leftrightarrow \quad F_\nu\left(\frac{sp}{\varepsilon\delta+sp}\right) \geq \frac{\varepsilon\delta+sp}{\varepsilon\delta+sp+rp}\,. \tag{12}$$

Hence this candidate equilibrium is an actual equilibrium if and only if (11) and (12) hold.

By comparing the set of parameters for each candidate equilibrium to be an actual equilibrium, we have the following (denote $c \triangleq \frac{sp}{\varepsilon\delta+sp}$):

- When $r+s > 1$: If $F_\nu(c) < \frac{s}{c}$, there exists a unique uncongested equilibrium; otherwise, there exists a unique congested equilibrium.
- When $r+s < 1$:
  - If $(1-r-s)p > \varepsilon\delta$. Then if $F_\nu\left(\frac{s}{1-r}\right) > 1-r$, we have a unique congested equilibrium; if $F_\nu\left(\frac{s}{1-r}\right) < 1-r$ then we have a unique uncongested equilibrium if $F(c) < \frac{s}{s+rc}$, a unique congested equilibrium if $F(c) > \frac{s}{s+rc}$.

– If $(1-r-s)p < \varepsilon\delta$. Then if $F_\nu\left(\frac{s}{1-r}\right) < 1-r$, we have a unique uncongested equilibrium; if $F_\nu\left(\frac{s}{1-r}\right) > 1-r$ then we have a unique uncongested equilibrium if $F(c) < \frac{s}{c}$, a unique congested equilibrium if $F(c) > \frac{s}{c}$.

Therefore, equilibrium always exists and is unique. □

A.5 Proof of Proposition 6

Proof. Follows from the proof of Proposition 5. □

B Proofs in Section 5

Proof of claim in Proposition 7. We verify the following properties of the correspondence $h(\cdot)$.

Existence on $[0, F_\nu^{-1}(s)]$. For $\Delta T = 0$, observe that

$$H(0,0) = \int_{\theta\in\Theta} f(\theta)\, d\theta = 1 > s\,, \qquad \lim_{T_{\text{safe}}\to\infty} H(T_{\text{safe}}, 0) = \ 0 < s\,.$$

By continuity of H in its first argument, the intermediate value theorem implies that there exists $T_{\text{safe}} > 0$ such that $H(T_{\text{safe}}, 0) = s$, i.e., $h(0) \neq 0$. Similarly, for $\Delta T = \bar{F}_\nu^{-1}(s)$, the measure of agents with $W_\nu(\theta) \geq \Delta T$ is exactly s, hence $H(0, \bar{F}_\nu^{-1}(s)) = s$ and $h(\bar{F}_\nu^{-1}(s)) = 0$. To ensure that $T_{\text{risky}} \geq 0$, we require $T_{\text{risky}} = T_{\text{safe}} - \Delta T \geq 0$, or $h(\Delta T) \geq \Delta T$. Since $h(\Delta T)$ is non-increasing and that $h(0) > 0$, $h(\bar{F}_\nu^{-1}(s)) = 0 < \bar{F}_\nu^{-1}(s)$, there must exists $\overline{\Delta T} \in (0, \bar{F}^{-1}\nu(s))$ such that $T_{\text{risky}} \geq 0$ holds on $[0, \overline{\Delta T}]$.

Uniqueness. Suppose to the contrary that $h(\Delta T)$ is multi-valued at some $\Delta T \geq 0$. Then there exists $T_1 < T_2$ such that $H(T_1, \Delta T) = H(T_2, \Delta T) = s$, But then:

$$\int_{\theta: W_\nu(\theta)\geq\Delta T} \left(e^{-\delta(\theta)T_1} - e^{-\delta(\theta)T_2}\right) f(\theta)\, d\theta = \ 0\,.$$

Because the integrand is strictly positive and $W_\nu(\theta)$ has positive density on $(0,\infty)$, the integral must be strictly positive – a contradiction. Hence $h(\Delta T)$ is single-valued.

Monotonicity. Fix $\Delta T' > \Delta T$. Suppose $T_{\text{safe}} = h(\Delta T)$. We show that $T'_{\text{safe}} \triangleq h(\Delta T') \leq T_{\text{safe}}$. By definition,

$$s = H(T_{\text{safe}}, \Delta T) > H(T_{\text{safe}}, \Delta T')\,.$$

But also, since H is decreasing in T_{safe},

$$H(T'_{\text{safe}}, \Delta T') = s \quad \Rightarrow \quad T'_{\text{safe}} < T_{\text{safe}} \,.$$

Therefore, $h(\cdot)$ is weakly decreasing.

Continuity. Given that H is continuously differentiable in both T_{safe} and ΔT, and strictly decreasing in T_{safe}, if follows from the implicit function theorem that $h(\cdot)$ is not only continuous but also differentiable on its domain.

This concludes the proof of the claim.

References

Agarwal, N., I. Ashlagi, M. A. Rees, P. Somaini, and D. Waldinger (2021). Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys. *Econometrica 89*(1), 37–76.

Arnosti, N. and P. Shi (2020). Design of lotteries and wait-lists for affordable housing allocation. *Management Science 66*(6), 2291–2307.

Ashlagi, I., F. Monachou, and A. Nikzad (2025). Optimal allocation via waitlists: Simplicity through information design. *Review of Economic Studies 92*(1), 40–68.

Aubert, O., P. P. Reese, B. Audry, Y. Bouatou, M. Raynaud, D. Viglietti, C. Legendre, D. Glotz, J.-P. Empana, X. Jouven, et al. (2019). Disparities in acceptance of deceased donor kidneys between the united states and france and estimated effects of increased us acceptance. *JAMA internal medicine 179*(10), 1365–1374.

Barah, M., V. Kilambi, J. J. Friedewald, and S. Mehrotra (2022). Implications of accumulated cold time for us kidney transplantation offer acceptance. *Clinical Journal of the American Society of Nephrology 17*(9), 1353–1362.

Bloch, F. and D. Cantala (2017). Dynamic assignment of objects to queuing agents. *American Economic Journal: Microeconomics 9*(1), 88–122.

Callaghan, C. J., L. Mumford, L. Pankhurst, R. J. Baker, J. A. Bradley, and C. J. Watson (2017). Early outcomes of the new uk deceased donor kidney fast-track offering scheme. *Transplantation 101*(12), 2888–2897.

Castro, F., H. Ma, H. Nazerzadeh, and C. Yan (2021). Randomized FIFO mechanisms. Working paper.

Chan, A. and A. E. Roth (2024). Regulation of organ transplantation and procurement: A market design lab experiment. *Journal of Political Economy 132*(11), 3827–3866.

De Mel, S. C., K. Munshi, S. Reiche, and H. Sabourian (2021). Herding with heterogeneous ability: An application to organ transplantation. Technical report, National Bureau of Economic Research.

Debout, A., Y. Foucher, K. Trébern-Launay, C. Legendre, H. Kreis, G. Mourad, V. Garrigue, E. Morelon, F. Buron, L. Rostaing, M. Kamar, Nassim adn Kessler, M. Ladrière, A. Poignas, A. Blidi, J.-P. Soulillou, M. Giral, and E. Dantan (2015). Each additional hour of cold ischemia time significantly increases the risk of graft failure and mortality following renal transplantation. *Kidney International 87*(2), 343–349.

Doval, L., F. E. W. Huang, and Y. Xin (2024). Social learning in lung transplant decision. *arXiv preprint arXiv:2411.10584*.

Gentry, S. E. and D. L. Segev (2013). Fast-tracking and fairness: Getting organ offers quickly to candidates who will accept them. *Liver Transplantation 19*(4), 358–359.

Guan, G., S. Neelam, J. Studnia, X. S. Cheng, M. L. Melcher, M. A. Rees, A. E. Roth, P. Somaini, and I. Ashlagi (2025). Insights from refusal patterns for deceased donor kidney offers. *Transplantation*, 10–1097.

Held, P. J., F. McCormick, A. Ojo, and J. P. Roberts (2016). A cost-benefit analysis of government compensation of kidney donors. *American Journal of Transplantation 16*(3), 877–885.

Jadlowiec, C. C., S. Y. Ohara, R. Punukollu, J. Wagler, B. Ruch, K. Kumm, P. Budhiraja, H. M. Me, A. K. Mathur, K. S. Reddy, et al. (2023). Outcomes with transplanting kidneys offered through expedited allocation. *Clinical Transplantation 37*(11), e15094.

King, K. L., S. A. Husain, J. D. Schold, R. E. Patzer, P. P. Reese, Z. Jin, L. E. Ratner, D. J. Cohen, S. O. Pastan, and S. Mohan (2020). Major variation across local transplant centers in probability of kidney transplant for wait-listed patients. *Journal of the American Society of Nephrology 31*(12), 2900–2911.

Leshno, J. D. (2022). Dynamic matching in overloaded waiting lists. *American Economic Review 112*(12), 3876–3910.

Liyanage, L. N., D. Akizhanov, S. S. Patel, D. L. Segev, A. B. Massie, D. E. Stewart, and S. E. Gentry (2025). Contemporary prevalence and practice patterns of out-of-sequence kidney allocation. *American Journal of Transplantation 25*(2), 343–354.

Medin, C., C.-G. Elinder, B. Hylander, B. Blom, and H. Wilczek (2000). Survival of patients who have been on a waiting list for renal transplantation. *Nephrology Dialysis Transplantation 15*(5), 701–704.

Mehrotra, S., J. M. Gonzalez, K. Schantz, J.-C. Yang, J. J. Friedewald, and R. Knight (2022). Patient preferences for waiting time and kidney quality. *Clinical Journal of the American Society of Nephrology 17*(9), 1363–1371.

Mildebrath, D., T. Lee, S. Sinha, A. J. Schaefer, and A. O. Gaber (2024). Characterizing rational transplant program response to outcome-based regulation. *Operations Research 72*(4), 1421–1437.

Milgrom, P. and J. Roberts (1994). Comparing equilibria. *American Economic Review 84*(3), 441–459.

Mohan, S., M. C. Chiles, R. E. Patzer, S. O. Pastan, S. A. Husain, D. J. Carpenter, G. K. Dube, R. J. Crew, L. E. Ratner, and D. J. Cohen (2018). Factors leading to the discard of deceased donor kidneys in the united states. *Kidney International 94*(1), 187–198.

Mohan, S. and S. A. Husain (2022). Improving the utilization of deceased donor kidneys by prioritizing patient preferences.

Mohan, S. and J. D. Schold (2022). Accelerating deceased donor kidney utilization requires more than accelerating placement.

Mohan, S., M. Yu, K. L. King, and S. A. Husain (2023). Increasing discards as an unintended consequence of recent changes in united states kidney allocation policy. *Kidney International Reports 8*(5), 1109.

Nikzad, A. and P. Strack (2023). Equity and efficiency in dynamic matching: Extreme waitlist policies. *Management Science 70*(8), 5187–5207.

Peters-Sengers, H., J. H. Houtzager, M. M. Idu, M. B. Heemskerk, E. L. van Heurn, J. J. H. van der Heide, J. Kers, S. P. Berger, T. M. van Gulik, and F. J. Bemelman (2019). Impact of cold ischemia time on outcomes of deceased donor kidney transplantation: an analysis of a national registry. *Transplantation direct 5*(5), e448.

Puttarajappa, C. M., S. Hariharan, X. Zhang, A. Tevar, R. Mehta, V. Gunabushanam, P. Sood, W. Hoffman, and S. Mohan (2023). Early effect of the circular model of kidney allocation in the united states. *Journal of the American Society of Nephrology 34*(1), 26–39.

Quiroga, I., P. McShane, D. D. Koo, D. Gray, P. J. Friend, S. Fuggle, and C. Darby (2006). Major effects of delayed graft function and cold ischaemia time on renal allograft survival. *Nephrology Dialysis Transplantation 21*(6), 1689–1696.

Sampaio, M. S., B. Chopra, A. Tang, and K. K. Sureshkumar (2018). Impact of cold ischemia time on the outcomes of kidneys with kidney donor profile index 85%: mate kidney analysis-a retrospective study. *Transplant International 31*(7), 729–738.

Schaubel, D. E., M. K. Guidinger, S. Biggins, J. D. Kalbfleisch, E. A. Pomfret, P. Sharma, and R. M. Merion (2009). Survival benefit-based deceased-donor liver allocation. *American Journal of Transplantation 9*(4), 970–981.

Sepulveda, J. M. G., S. Mehrotra, J.-C. Yang, K. J. Schantz, Y. Becker, R. Formica, D. P. Ladner, D. Kaufman, and J. Friedewald (2023). Physician preferences when selecting candidates for lower-quality kidney offers. *Clinical Journal of the American Society of Nephrology 18*(12), 1599–1609.

Shi, P. and J. Yin (2022). Eliminating waste in cadaveric organ allocation. Working paper, USC Marshall School of Business.

Su, X. and S. Zenios (2004). Patient choice in kidney allocation: The role of the queueing discipline. *Manufacturing & Service Operations Management 6*(4), 280–301.

Su, X. and S. A. Zenios (2006). Recipient choice can address the efficiency-equity trade-off in kidney transplantation: A mechanism design model. *Management Science 52*(11), 1647–1660.

Sweat, K. (2023). Endogenous priority in centralized matching markets: The design of the heart transplant waitlist. Technical report, Working Paper. Stanford University.

USRDS (2013). Us renal data system. usrds 2013 annual data report: Atlas of chronic kidney disease and end-stage renal disease in the united states. National Institutes of Health, National Institute of Diabetes and Digestive and Digestive and Kidney Diseases, Vol. 2014.

USRDS (2018). Us renal data system. usrds 2018 annual data report: Atlas of chronic kidney disease and end-stage renal disease in the united states. National Institutes of Health, National Institute of Diabetes and Digestive and Digestive and Kidney Diseases, Vol. 2019.

Waldinger, D. (2021). Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation. *American Economic Review 111*(8), 2660–2696.

Wolfe, R. A., V. B. Ashby, E. L. Milford, A. O. Ojo, R. E. Ettenger, L. Y. Agodoa, P. J. Held, and F. K. Port (1999). Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *New England journal of medicine 341*(23), 1725–1730.

Wood, N. L., D. N. VanDerwerken, A. B. Massie, D. L. Segev, J. J. Snyder, and S. E. Gentry (2025). Diagnosing the recent decrease in utilization of deceased donor kidneys. *Transplantation 109*(3), 496–503.

Zhang, J. (2010). The sound of silence: Observational learning in the us kidney market. *Marketing Science 29*(2), 315–335.