Hurwicz meets Veatch: Rationing deceased-donor transplants under dynamic asymmetric information*

Edwin Muñoz-Rodriguez†

April 14, 2022

Is it possible to identify just institutions for a society without making them contingent on actual behavior (not necessarily the same as 'just' or 'reasonable' behavior)?

—Amartya Sen. *The Idea of Justice.*

The primary moral tension in allocating organs among those waiting for them is the conflict between their efficient use and fair distribution.

—Robert Veatch. *Transplantation Ethics.*

Abstract

Since the late 80's, there is a heated debate on the principles of distributive justice for rationing transplants. At the same time, it is well-known that the U.S. transplantation authority has recurrently faced a pervasive problem of asymmetric information about transplant candidates' medical urgency. I investigate the optimal design of prioritization rules under different social welfare functions while taking patients' incentives to misrepresent medical needs into account, and analyze their long run stability. While the history of reports of medical urgency could always be used to incentivize truth-telling, it is not necessarily optimal to do so. When the social objective is to minimize the mass of unserved sick patients, the optimality of screening is ambiguous and depends on the parameter region. In sharp contrast, when the objective is utilitarian,

*First publicly presented: 05-21-2019. First draft publicly available: 02-02-2021. This version superseeds a previous draft circulated with the title "Incentive compatible triaging in deceased-donor transplantation." I own special gratitude to Brett Gordon, Daniel Martin, Luis Rayo and James Schummer. I benefited from multiple discussions with Henrique Castro-Pires, Daniel Barron, David Dranove, Piotr Dworczak, Laura Doval, Andres Espitia, Nitin Katariya, Peter Klibanoff, Daniela Ladner, Joshua Mollner, Sanket Patil, Alessandro Pavan, Michael Powell, Elena Prager, Alvaro Sandroni, Bruno Strulovici, and participants at Seminars at Colmex, Kellogg, Ottawa, U Tokyo, Virginia Tech, and the 2021 China and Asian Meetings of The Econometric Society. This work was supported in part by HRSA contract 234-2005-370011C. The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the DHHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

†Centro de Estudios Economicos, El Colegio de Mexico A.C., eamunoz@colmex.mx

screening is not optimal in general. Moreover, while the prescribed optimal policies for the two objectives are in general different, there is a region of parameters where they coincide, in which case, once the incentive problem is taken into account, the two principles of distributive justice are not in conflict anymore.

Keywords: Organ allocation, queue design, recursive contracts, dynamic mechanism design, social choice, medical ethics.

1 Introduction

With the exception of kidney transplantation, the design of organ allocation policies and algorithms involve decisions about *whom to save from death*.[1] In fact, in the U.S. the large majority of lung, liver and heart transplant candidates have no other alternative but hope to survive long enough to get an organ from a deceased donor through the national waiting list.[2] Since the organ supply is scarce, it is a fact that many of these patients will die while waiting. Given the complex moral issues in question in the assignment of these life-saving resources, medical ethics have played a key role in the design of the U.S. organ allocation system.

A fundamental design challenge has been to determine justifiable principles for rationing these lifesaving public resources, and translate them into material criteria that can be the basis for allocation policies. In 1986, the U.S. Task Force on Organ Transplantation recommended that organ allocation decisions consider equity, medical utility, and autonomy, but also recognized that the principles of equity and medical utility might result in competing claims to organs between transplant candidates (Organ Transplantation 1986): it is a difficult ethical issue to determine whether the patient with the better outcome or the most urgent one should receive an organ. This conflict between two major principles of distributive justice has resulted in heated debates among stakeholders, policymakers and ethicists.

In these debates, the informational demands of different principles and the actual behavior their implementation incentivize on patients and caregivers have been mostly overlooked. It is implicitly assumed that whichever is the morally correct balance between equity and transplant utility, the appropriate institutions or allocation rules follow from a “reasonable” behavior on the part of the players in the system, and the availability of accurate medical

[1] I borrow the expression “whom to save from death” from the subtitle of an authoritative book on the subject matter (Kamm 1993).

[2] To the date there is no reliable long-term treatment for end stage hepatic, cardiac or pulmonary disease, other than transplantation. In contrast, in the case of renal disease, dialysis is a long-term reliable alternative treatment, which differentiates kidney allocation from the other organ allocation problems. In addition, living-donor liver and lung transplants are very complex medical procedures only viable for patients with particular diagnoses.

metrics. For instance, in the allocation of hearts and livers, the lack of backup treatments led to favoring equity over medical utility since the system's inception. Historically, this has been translated into the requirement for allocation policies to give absolute priority to the medically worst-off or most urgent transplant candidate.

A major unintended consequence of this approach has been a recurrent pervasive problem of asymmetric information, particularly prominent in liver and heart allocation. Most of the transplant candidates' covariates intended to reflect medical urgency, such as diagnosis, medical therapies, and lab tests, have been recurrently manipulated through unnecessary medical interventions so that patients appear to have greater urgency than they actually do.[3] Since getting a liver or heart is a matter of life an death, patients and physicians face high-powered incentives to engage in such strategic manipulations.[4] In addition, transplant centers and providers also face financial and reputational incentives to increase the volume of transplants performed. Thus, it is not surprising that despite multiple revisions to the allocation rules, the ethics committee of the U.S. transplantation authority recognizes that

> current safeguards do not sufficiently mitigate this risk. Thus, the OPTN, OPTN/UNOS Committees, and the transplant community should consider refining current and/or developing additional safeguards to mitigate the risk of manipulation of candidates' waitlist priority (OPTN/UNOS 2018).

This gaming of priorities has several undesirable consequences which can be exacerbated in equilibrium. On top of organ misallocation, which jeopardizes the system's objective itself, the manipulation using medical interventions destroys information by pooling patients at the higher urgency statuses. Moreover, the mimicking requires overtreatment which unnecessarily increases healthcare expenditures. It also undermines the public perception and trust in the system, which is critical for a program which largely depends on both, public funding and the public willingness to donate organs.

In this paper, I use tools from mechanism design to study the implementation of different social welfare objectives in the allocation of deceased-donor organs in the presence of asymmetric information. While the delicate interplay between social objectives and their informational requirements is well-known in Economics since the pioneering work of Leonid Hurwicz, it has been recurrently overlooked in the discussions on the design of the national organ allocation system. Given the state of affairs above described, such a perspective seems

[3]See section 2 for a summary of the type of manipulations.

[4]For individual physicians, the commitment to their patients' well-being through the Hippocratic ethics can justify the gaming of the organ allocation system, a possibility envisioned in another context by one of the funding fathers of medical ethics and architect of the U.S. transplantation system, Robert M Veatch (2000).

necessary and can provide new elements to the debate on the principles of distributive justice regarded as desirable.

Because in the U.S. organ allocation system monetary transfers are forbidden, authorities are heavily limited in the instruments they can use to incentivize truth-telling. Since the information about transplant candidates' medical needs is intrinsically dynamic (see section 2), it is natural to explore a well-known idea from the literature on repeated games and dynamic incentives: the social planner could overcome the incentive problem by prioritizing patients using the history of medical urgency reports.[5] I derive the optimal prioritization rule in steady state for two different social welfare functions that capture in a stylized way the conflicting principles of equity and transplant utility, and analyze the long run stability of these optimal steady states under plausible long-run objectives.

I find that while a planner could always use the history of reports of medical urgency to incentivize truth-telling, it is not necessarily optimal to do so: how the trade-off between efficiency and information rents resolves depends on the social objective and the primitives of the environment. When the social objective is to minimize the mass of unserved sick patients, screening is optimal only in a sub-region of parameters; given that information rents are costly, outside this sub-region a uniform lottery among senior or junior transplant candidates is optimal. In sharp contrast, when the objective is utilitarian, screening is not optimal in general. Moreover, while the prescribed optimal policies for the two objectives are in general different, there is a region of parameters where they coincide, in which case, once the incentive problem is taken into account, the two principles of distributive justice are not in conflict anymore.

I set up a model of overlapping generations in discrete time. In every period, a unit mass of junior patients arrives to the system, and a social planner with full commitment receives a mass of perishable organs. Patients live for at most two of periods and have privately known binary health states that vary stochastically over time with some degree of persistence. Patients leave the market either by undergoing transplantation, when they reach their maximum longevity, or because they pass away while waiting. Each individual patient is required to report a health state to the planner every period. The set of health state messages for patients depend on their actual health state, as in the literature on mechanism design with partially verifiable information (Green et al. 1986). I assume that misrepresentation is not costly in the sense that it does not impose increased risk or lower quality of life to the patients. This assumption is conservative: when gaming is costly, the incentives to manipulate are partially offset by the cost of doing so.

From the two social welfare functions I consider, one, which I name 'equitable', captures

[5] literature history incentives

in a stylized way the objective of prioritizing the medically worst off and aims to minimize the mass of untreated sick patients in the market. The other is a utilitarian one which aims to maximize the total sum of patients' expected lifetime utility.

I find that the trade-off between the costs and benefits of screening patients faced by the clearinghouse resolves differently depending on social preferences for organ allocation. When the social objective is the 'equitable' one, Theorem 1 shows that the optimal allocation rule can take very different forms depending on the environment parameters. Optimal allocation rules can differ in two dimensions: whether screening is optimal or not, and who waits. Take as a benchmark a uniform lottery among junior patients, a simple incentive compatible allocation rule which does not screen and creates no delay. Such a rule allocates some organs to a fraction of junior healthy patients who would be either healthy in the future or dead, which is very inefficient from the perspective of removing sick patients from the market. This cost, however, is offset in some cases by the benefit from removing sick junior patients from the market who anyways would become sick. If the probability of a patient becoming sick in the future is large and the survival probabilities are small, the cost can offset the benefit of allocating organs to junior patients, hence a lottery among seniors performs better than a uniform lottery among junior patients.

In stark contrast, Theorem 2 shows that a utilitarian social planner never incentivize junior healthy patients to wait, *despite the fact* that healthy junior patients benefit less from an organ as compared to junior sick patients. The information rents that would be required to have junior healthy patients waiting are too high, and the utilitarian social planner internalizes their utility loss. Given this, if screening is optimal, it is achieved by forcing sick junior patients to wait. This is doable for free, because sick patients cannot manipulate their health state, but it is optimal only if demanding conditions are satisfied. One of these conditions is that a junior sick patient benefits more (in expectation) from getting an organ when senior than a healthy junior patient would from getting an organ upon arrival. A utilitarian social planner pools patients more often, and in this case serves them upon arrival with a uniform lottery.

My findings have clear policy implications. First, the current allocation rule, which is a naive implementation of the medical urgency principle under the assumption that individual reports of medical urgency are sincere, is not incentive compatible. To incentivize truth-telling, patients in position to "game" their medical urgency should not be unconditionally relegated to the lowest priority tiers. Second, in the face of recurrent manipulation of the triaging policies, the typical institutional response has been either to ignore "gameable" covariates, or to institute stricter conditions for considering a treatment as a valid signal of medical urgency.[6] A revision of how the current allocation policies use the history of reports of medical urgency to triage transplant candidates can help to alleviate the problem of asymmetric information.

Third, I bring a new perspective to the debate on the principles for the allocation of organs for transplantation: the trade-offs imposed by the need of eliciting the information relevant for allocation purposes. Policymakers and medical ethicists have largely ignored the problem of asymmetric information created by the attempt to implement different ethical principles. The two principles considered here, that have been the protagonists of a longtime debate, interact in a complex way with the informational constraint. In general, the prescribed optimal policies for the two objectives are different, but there is a region of parameters where they coincide, in which case, once the incentive problem is taken into account, the two objectives are not in conflict.

1.1 Contributions to the literature

The problem of queue design for allocating organs has a long tradition in Operations Research (Ata et al. 2018), but typically this literature does not consider problems of asymmetric information. In Economics, the allocation through waiting lists has been studied in settings without asymmetric information (Condorelli 2012; Bloch et al. 2017; Schummer 2020; Che et al. 2021) and with asymmetric information (Margaria 2016; Leshno 2017; Thakral 2016). In contrast to the latter, I compare the optimal allocation rule for two different social welfare functions in a setting where agents are short-lived, with dynamic and partially verifiable information.

The *deceased-donor* organ allocation problem, which is one of allocating socially owned organs among many waiting patients is starting to attract the attention of economists. Concurrent papers study the allocation of deceased-donor organs from empirical (Dickert-Conlin et al. 2019; Agarwal, Ashlagi, et al. 2021; Agarwal, Hodgson, et al. 2020) and experimental perspectives (Genie et al. 2020; Sullivan 2021). My work contributes to this rapidly evolving literature by introducing state-of-the-art techniques from dynamic contracting and dynamic mechanism design to provide key insights on the design of the organ allocation rules.

My work also relates to the literature on dynamic contracting without transfers (Atkeson et al. 1992; Farhi et al. 2007; Olszewski et al. 2020; Guo et al. 2018; Li et al. 2017; Lipnowski et al. 2018). Most of this literature assume that agents or households are infinitely lived and demand more than one object, or that there are inter-generational linkages which lead

[6] The latter has been found to be controversial, for these stricter conditions are not necessarily meet by actual urgent patients (Parker, Garrity Jr, et al. 2017), while the former is likely to disregard information that would have been informative in the absence of gaming (Stewart et al. 2007).

to dynasties which consume forever.[7] In contrast, the organ allocation setting is one where agents demand one and only one object, which is technically challenging because it generates endogenous population dynamics. To make progress on the pressing policy issues studied here, I fist restrict attention to steady-state analysis, which is the standard approach in queuing theory and has been adopted in recent papers in Economics to overcome technical difficulties (Margaria 2016; Leshno 2017; Che et al. 2021). Moreover, this approach allows me to bypass the absence of clearly formulated long run social objectives in organ allocation. Second, I conduct numerical analysis for plausible long run social objectives and a large class of parameters which suggests the system converges to a steady-state over the optimal path, regardless initial conditions.

2 Background on heart and liver allocation

In the U.S., the organization in charge of coordinating the organ placement process is the Organ Procurement and Transplantation Network (OPTN), a public-private partnership that by law includes all the agents in the national system of donation and transplantation.[8] Part of the responsibilities of OPTN/UNOS is to develop policies for organ recovery, allocation, and transportation.

Since the late 80's, individual medical urgency was adopted as the main criteria in liver and heart allocation. Additional criteria used for allocation purposes are geography (relative position of donor and recipient), blood compatibility and waiting time. Except for medical urgency and geography, which are interwoven, the ordering is lexicographical with urgency and geography as the main criteria, followed by blood compatibility, and finally waiting time.

2.1 Manipulation in liver allocation

In liver allocation medical urgency was initially given by patients' hospitalization status, with patients' located in the ICU getting the largest priority for a liver. At the end of the 90's, evidence of widespread manipulation of patients' hospitalization status to raise their priority for a liver lead to a revision of the system and the development of the Model of

[7] To my knowledge, only Kovac et al. (2013) has studied the optimal stopping case in a principal-agent relationship, which can be interpreted as a single unit consumption. However, the settings are irreducible, mainly because the competition for organs.

[8] It was created by the U.S. Congress by means of the National Organ Transplant Act (NOTA) in 1984, and since 1986 it is managed by the private contractor United Network of Organ Sharing (UNOS). It is part of the Department of Health and Human Services

End-stage Liver Disease (MELD) score (Freeman et al. 2004; Merion et al. 2011).[9] MELD is based on a survival model and predicts three-month mortality in the waiting list based on blood tests and patient's dialysis status.[10] For allocation purposes, the score was restricted to take integer values, capped at 40, and to give a large value of creatinine to patients on dialysis (Freeman et al. 2004).

While the adoption of MELD score reduced the risk for manipulation, the OPTN points out that it "has not been eliminated completely," and cites the example of starting a patient with "mild to moderate renal impairment on dialysis without actual indication" as a way of gaming priorities (OPTN/UNOS 2018). The medical literature indicates that lab tests can also be manipulated towards higher MELD score through the use of medications like diuretics, warfarin or vasopressin receptor antagonists without affecting prognosis (Moore et al. 2012; Rowe et al. 2007; Asrani et al. 2010).

2.2 Manipulation in heart allocation

In the case of heart allocation, treatment aggressiveness determines patients' medical urgency, under the implicit assumption that critically ill candidates require more aggressive life-sustaining measures. The original scale of medical urgency, in place since 1989, had only two tiers. Evidence of gaming through unnecessary therapy escalation (Scanlon et al. 2004) prompted action from policymakers, leading to a refinement of the prioritization rules in 1999.

However, physicians and transplant centers adapted their behavior to the new rule. It became common knowledge among practitioners that high-dosages of inotropes and pulmonary artery catheters were being overused as a way of increasing the patients likelihood of undergoing transplantation (Stevenson 2013; Stevenson et al. 2016; Rao et al. 2018). This issue triggered a new revision of the prioritization scheme by the OPTN/UNOS Heart Transplantation Committee (Committee 2016 (accessed June 30, 2020)), which culminated with the adoption of a new allocation rule in 2018. In the new rule, inotropes only guarantee the third and fourth urgency tiers, while only patients on very aggressive therapies like mechanical circulatory or ventilatory support are given the highest priority status (Colvin-Adams et al. 2012; Rao et al. 2018).

Recent empirical evidence is consistent with a striking strategic response from transplant centers to these changes. Parker, Chung, et al. (2020) found that while the use of inotropes

[9] Snyder (2010) leverages the adoption of MELD in 2002 to show that ICU usage dropped significantly after its elimination as triaging criteria.

[10] Up to 2016, MELD was based on serum creatinine, international normalized ratio of prothrombin (INR) and bilirubin. It was afterwards updated to include serum sodium (MELD-Na).

	Liver	Heart
Patients at begining of year	16,539	3,380
Anual patients arrivals	11,199	3,638
Anual patients departures (different from transplantation)	5,239	1,131
Organs transplanted	6,890 (294 living)	2,487

Table 1: Dynamics of transplant waiting pool, annual averages 2002-2018. Source: OPTN STAR Files as of March, 2019

decreased after the policy change, that of other, more aggressive therapies like IABP and ECMO increased. Ran et al. (2021) found that

> "The odds of high-priority listing was more than five times greater than expected in the post-policy period, without accompanying explanatory changes in candidate characteristics. Transplant centers all over the country listed more candidates than expected at high-priority status."

2.3 Stylized facts

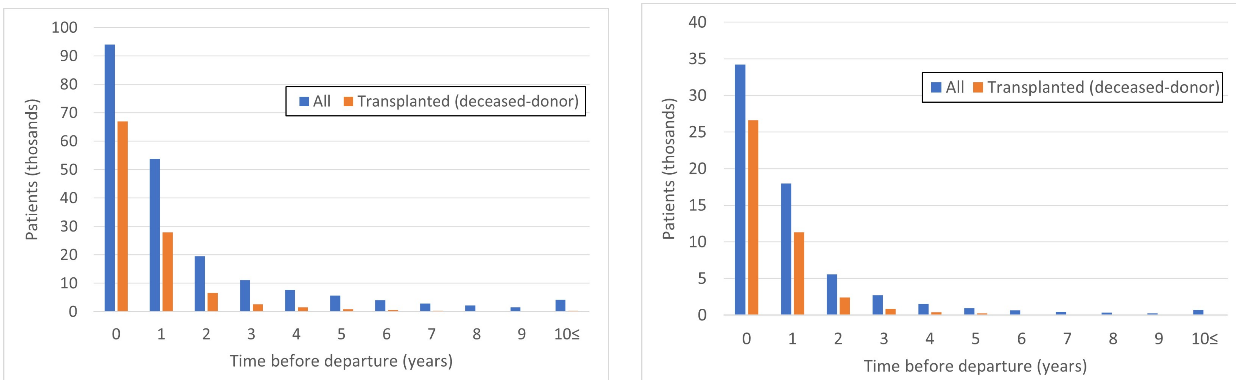
Based on the above exposition, data from Standard Transplant Analitical Files (STAR), medical literature and talks with practitioners, I highlight the most important stylized facts that motivate key features of the model.

Fact 1: *The ability of misrepresenting medical urgency depends on the actual medical urgency of the patient.*

Truly very sick patients in the highest degree of urgency cannot misrepresent their health state, for the medical urgency scale is bounded. In the case of hearts, the most invasive therapies like ECMO or Non-dischargable LVAD cannot be escalated further. In the case of livers, patients who medically require dialysis cannot use dialysis as a tool to misrepresent their degree of medical need. More nuance in the case of taking pills to increase MELD score, but the general idea is still true: since the MELD score is caped at 40 for allocation purposes, patients already in the highest score cannot increase it beyond.

Fact 2: *Organ arrival rate is smaller than patient arrival rate*

Between 2002 and 2018, the annual amount of organs transplanted was smaller than the arrival of new patients (Table 1). In the case of livers, there was on average 54 organs transplanted per each 100 new patients. In the case of hearts, there were 67 transplant per each 100 new patients registered each year.



(a) Liver (b) Heart

Figure 1: Time from arrival to departure, all patients 2002-2018

Fact 3: *Heterogeneity of tenures in patient mix*

Table 1 shows a considerable number of patients spend more than one year on the list waiting for an organ. The entire distribution of the time elapsed between arrival and departure for all patients in the waiting list between 2002 and 2018 is depicted in Figure 1. In the case of livers (Figure 1a) the average waiting time before departure is 1.49 ± 2.44 years, and 71 % of the candidates depart in the first two years after arrival. From this latter group of patients, 64% departs due to deceased donor transplantation. In the case of hearts (Figure 1b), the average waiting time before departure is considerably lower (1.05 ± 1.93 years), 79 % of the patients depart in the first two years, and from them 73% undergo transplantation.

Fact 4: *Patients' medical urgency status change over time*

The design of the allocation rules offers evidence of the natural variation in the level of medical urgency of a patient. In fact, the allocation policies for both, liver and heart feature an update schedule that requires transplant centers to update their patients' urgency status regularly with a frequency that depends on the last report of medical urgency. [11]

Moreover, using non-allocation data–hence not subject to manipulation–Alagoz et al. (2005) document within patient variation in MELD score.

Fact 5: *Transplantation improves quality of life*

Studies on quality of life before and after transplantation consistently find that transplantation improves patients' quality of life relative to the preoperative status in several dimensions including physical functioning, as well as cognitive abilities (Trzepacz et al. 2000; Yang et al. 2014; Mabrouk et al. 2012).

[11] For livers, see Policy 9.2 in OPTN Allocation policies. For hearts, see Policy 6.1.

Fact 6: *Pre-transplant candidates quality of life varies with severity of organ failure*

Sicker patients tend to have lower pre-transplant quality of life. In the case of end stage liver disease, it is well known among transplant specialists that pre-transplant quality of life varies across patients, and patients in advanced stage tend to present complications like hepatic encephalopathy that impairs transplant candidates' functional status and cognitive abilities (Orr et al. 2014).[12]

3 Setting

Time is discrete and indexed by $s = 0, 1, 2, \ldots$ There is a social planner (SP) with full commitment. In every period s, a mass $x < 1$ of identical and perishable organs and a unit mass of patients–or transplant candidates–arrive to the system. Upon arrival, candidates live for at most two periods. I index the patient's tenure in the system by $t \in \{1, 2\}$. Departures happen because either a patient pass away, reaches the maximum longevity or receives an organ.

While in the waiting list, each transplant candidate has a time-evolving binary health state which is either good (G) or bad (B) with probabiliries $q(G) + q(B) = 1$. This type determines patient's *pre-transplant* survival, health state evolution and flow payoffs.

Before transplant, a junior patient in health state h_1 survives with probability δ_{h_1} to the next period; conditional on not getting an organ and surviving, her health state in next period is $h_2 \in \{B, G\}$ with probability $p(h_2|h_1)$. Patients in a poor health condition are more likely to die without an organ, and if they survive to the next period, they are more likely to be sick than patients in good health condition, this is

$$\delta_B < \delta_G, \qquad p(B|B) > p(B|G) \tag{1}$$

While waiting for an organ, every patient receives a flow payoff $u(h_t)$. This per-period payoff is intended to capture the quality of life a patient enjoys while in need for an organ, and depends on her health state. According to Fact 6, sick patients waiting for organs have a lower quality of life than not very sick patients, this is

$$u(B) < u(G) \tag{2}$$

A transplant has two effects: extend life (increased survival) and improve quality of life.

[12] Hepatic encephalopathy is a "disturbance in the central nervous function because of hepatic insufficiency"; symptoms range from short-term memory loss and lack of concentration to disorientation and lethargy (Sharma et al. 2005).

Once they receive an organ, and for their remaining lifetime, patients are assumed to be identical. More specifically, patients who receive an organ stay alive for their remaining lifespan, and enjoy a better per-period quality of life, $\bar{u} > u(G)$. This assumes there is no heterogeneity in post-transplant patient survival, nor in quality of life across recipients, regardless of their pre-transplant medical history.[13]

3.1 Information

Every period patients privately learn their health state. For any patient with tenure t, denote by $m^t \in M^t = \{B, G\}^t$ a patient's history of realized health states since she was born and up to period t, inclusive. I also refer to this as the patient's medical history. Let $\mathcal{M}$ be the set of all medical histories.

Every period s a patient is in the waiting list, they are required to report their health state to the SP. Denote the report at time s of a patient with tenure t by $\hat{m}_{s,t}$, and the corresponding history of reports by $\hat{m}_s^t$.

The set of messages that can be sent by patients depends on their actual individual health state at the corresponding period. Specifically, when health state is B, the candidate can only report B, but when health state is G the candidate can report anything in $\{B, G\}$. This is in line with the manipulation using medical interventions described above. I assume misrepresentation is not costly to the healthy patient. In the case it is costly, the gain from manipulation is partially offset by its cost. So the costless case is conservative.

3.2 Mechanism or allocation rule

Let $h^s = (h_0, h_1, \ldots, h_s)$ be the history of the system up to calendar period s. h_s denotes what happened in period s. A sufficient statistic for h_s are $\mu_s(\hat{m}^t), \hat{m}^t \in \mathcal{M}$, the masses of patients who were in the market by the end of the period, after allocation happened, after reporting a given medical history in period s. Let H^s the set of possible histories of the system at period s, and $\mathcal{H} = \bigcup_s H^s$ the set of all histories.

In period s, based on the reports from all candidates and the history transpired so far, SP makes organ allocation decisions, which are given by a mechanism or allocation rule

$$\psi_s : \mathcal{H} \times \mathcal{M} \mapsto [0, 1] \tag{3}$$

that specifies the probability of a transplant candidate receiving an organ for transplantation

[13]This assumption is clearly strong, but it helps with tractability and allows to concentrate on the informational aspect of the tension between competing ethical principles. More on this in Section 3.4.

as a function of her own reported medical history.

Several remarks are in place.

In this setting there is population dynamics, which arises from the fact that patients only consume one object and can die in the waiting list. Moreover, this population dynamics is endogenous, for it is partially determined by the allocation rule itself. In fact, suppose all information is public. Thus, in period s, the mass of senior patients in the market with medical history m^2 before allocations have taken place is

$$\mu_s(m_1(m^2)) = q(m_1(m^2))(1 - \psi_{s-1}(h^{s-1}, m_1(m^2))) \tag{4}$$

which clearly depends on the mechanism. This adds a major difficulty to my analysis. In dynamic contracting and mechanism design either population or information is constant, but not both.

3.2.1 Incentive compatibility

Only patients in a good health state can misrepresent it. At any period s, those are the junior patients who arrive in good health state to the market, or those senior whose medical history is of the form $m^2 = (m_1, G)$. Thus, A mechanism is incentive compatible if, for all s, the G patients prefer to reveal their health state.

This is, a mechanism $\{\psi_s\}_{s=0}^{\infty}$ is incentive compatible if, for every s, h^s:

$$\psi_s(h^{s-1}, G, G) \geq \psi_s(h^{s-1}, G, B), \quad \psi_s(h^{s-1}, B, G) \geq \psi_s(h^{s-1}, B, B) \tag{5}$$

let

$$w^G_{s+1}(h^s, G) = \sum_{m \in \{G,B\}} p(m|G)[\psi_{s+1}(h^s, G, m)\bar{u} + (1 - \psi_{s+1}(h^s, G, m))u(m)] \tag{6}$$

and

$$w^G_{s+1}(h^s, B) = \sum_{m \in \{G,B\}} p(m|G)[\psi_{s+1}(h^s, B, m)\bar{u} + (1 - \psi_{s+1}(h^s, B, m))u(m)] \tag{7}$$

be the expected second period payoff of a healthy junior patient who does not get an organ in the first period, after a truthful and non-truthful first period report. Then, incentive compatibility for young patients is

$$\begin{aligned}&\psi_s(G)2\bar{u} + (1 - \psi_s(G))[u(G) + \delta_G w^G_{s+1}(G)] \\ &\qquad \geq \psi_s(B)2\bar{u} + (1 - \psi_s(B))[u(G) + \delta_G w^G_{s+1}(B)]\end{aligned} \tag{8}$$

Equations 5 -8 summarize the informational constraints faced by the social planner due to the private nature of health states.

3.2.2 Feasibility

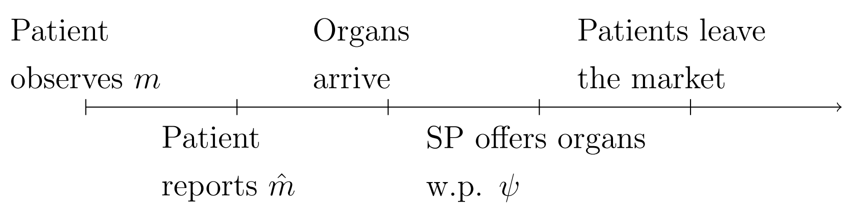
An incentive compatible allocation rule is said to be feasible if

$$\sum_{m\in\{G,B\}} q(m)\psi_s(h^{s-1},m) + \sum_{m'\in\{G,B\}} \mu_s(m')\delta(m') \sum_{m\in\{G,B\}} p(m|m')\psi_s(h^{s-1},m',m) \leq x \quad (9)$$

Since organs are perishable, the budget constraint holds per period. The first term corresponds to the organs allocated to junior patients in period s, while the second is the organs allocated to senior transplant candidates. The mass of senior transplant candidates in the market is $\sum_m' \mu_s(m')\delta(m')$, with μ_s given in (4).

3.3 The stage game

New patients and organs arrive to the market every period. All patients are required to report their health states to the social planner. Using these reports, the planner allocates organs. If a patient gets an organ, or pass away, leaves the waiting list. Surving junior patients stay in the list. The problem of the planner is to allocate organs among patients in the market in a way that they have incentives to report truthfully their health state.



3.4 Social welfare functions

I now introduce two different social welfare functions that formalize the two conflicting principles.

Principle of equity In the design of priorities for organs, the mainstream interpretation of equity is a version of the *maximin* principle: among all the patients in condition to undergo non-futile transplantation, those who are medically worst off deserve first consideration (Robert M Veatch 2000). The measure of medical need for a transplant is given by the pre-transplant survival probability. This is why the Final Rule (Title 42 § 121.8 (b) 2) mandates that organ allocation must be conducted

> Setting priority rankings [...] for patients or categories of patients who are medically suitable candidates for transplantation to receive transplants. These rankings shall be ordered from most to least medically urgent [...]

Since the current allocation rules of hearts and livers give absolute priority to the sickest transplant candidates, with homogeneous organ quality healthier patients only receive organs when there are no sick candidates in the system. I formalize this as follows.

For any incentive compatible mechanism, the mass of untreated sick patients by the end of period s, after allocation happened, is

$$q(B)\left(1-\psi_s(h^s,B)\right)+\sum_{m\in\{B,G\}}\mu_s(m)p(B|m)\left(1-\psi_s(h^s,m,B)\right) \tag{10}$$

The term $q(B)\left(1-\psi_s(h^s,B)\right)$ corresponds to the mass of junior patients who arrived to the market in bad health state who did not received an organ in period s after history h^s; on the other hand, the term $\sum_{m=B,G}\mu_s(m)p(B|m)\left(1-\psi_s(h^s,m,B)\right)$ is the mass of senior patients who survived to the second period $\mu_s(m)$, whose health state is bad when seniors (which happens with probability $p(B|m)$) who did not received an organ in period s.

The objective of an equitable social planner is to find an allocation rule ψ that minimizes the mass of unserved patients in poor health state (eq. 10). In addition, to prevent organ wastage, I assume that the budget constraint must bind.

Principle of utility There are multiple interpretations of utilitarianism in the context of organ transplantation and more broadly speaking in the ethics of allocation of scarce medical resources (R. Veatch 2004). At least two measures of transplant utility are relevant for the design of the organ allocation system: pure length of life and patients well-being. We consider the second one, which is the standard in Economics and is more general (set $u(m)=\bar{u}=1$ and the model reduces to a pure survival one).

Let

$$\begin{aligned}U_s(\psi_s,\mu_s,h^s)=&\sum_{m\in\{B,G\}}q(m)[u(m)+\psi_s(h^s,m)(\bar{u}-u(m))]+\\ &\sum_{m\in\{B,G\}}\left(\mu_s(m)\sum_{m'\in\{B,G\}}p(m'|m)[u(m')+\psi_s(m,m')(\bar{u}-u(m'))]+\left(q(m)-\frac{\mu_s(m)}{\delta_m}\right)\bar{u}\right)\end{aligned}$$

the total sum of the utility attained by all patients who are alive in period s, which includes all the junior patients, as well as all the senior patients who remain in the waiting list and those who received transplantation when junior, in period $s-1$. The first summation is the total utility gained in period s by junior patients, including the additional marginal benefit obtained by those who received an organ. The second summation is the total utility attained by senior patients in period s, which is the sum of two terms: the first, is utility achieved by those senior patients who are in the waiting list in period s, $\mu_s(m)$, and the second is the total payoff collected by the senior patients who received an organ when junior, $\left(q(m) - \frac{\mu_s(m)}{\delta_m}\right)$.

Long run social preferences The ethical principles as well as the current allocation rules are completely silent regarding the long run horizon. Thus, any formalization of the long run objectives is to some extent arbitrary and not grounded in explicitly established policy goals. In the next section I concentrate on steady-state analysis, so the lack of a long-run objective is not as crucial. In a subsequent section, I investigate the transitional dynamics and convergence to steady state under reasonable long run objectives.

4 Optimal allocation rules

In this section, I assume the planner is constrained to use stationary mechanisms that only condition on the patients' individual reports (but not, for instance, the mass of senior transplant candidates). Moreover, I restrict attention to the steady-states associated with this type of allocation rules. This allows me to overcome technical difficulties associated with the solution of an infinite dimensional optimization problem, and conceptual difficulties related to the lack of explicit long run objectives. Both will be revisited latter.

These assumptions simplify notoriously the incentive constraints. Senior patients' incentive compatibility constraints become

$$\psi(G,G) \geq \psi(G,B), \quad \psi(B,G) \geq \psi(B,B) \tag{11}$$

and for the junior patients,

$$\psi(G)2\bar{u} + (1-\psi(G))[u(G) + \delta_G w^G(G)] \geq \psi(B)2\bar{u} + (1-\psi(B))[u(G) + \delta_G w^G(B)] \tag{12}$$

with the corresponding obvious expressions for $w^G(B)$ and $w^G(G)$ which follow from eqns (7)-(6). Moreover, for any incentive compatible mechanism ψ, the per-period budget constraint reduces to

$$\sum_{m\in\{G,B\}} q(m)\psi(m) + \sum_{m'\in\{G,B\}} \mu(m')\delta(m') \sum_{m\in\{G,B\}} p(m|m')\psi(m',m) \leq x \tag{13}$$

with $\mu(m) = q(m)(1 - \psi(m))$.

4.1 Principle of equity

In a steady-state, the equitable objective (10) becomes to minimize

$$q(B)\,(1 - \psi(B)) + \sum_{m \in \{B,G\}} \mu(m) p(B|m)\,(1 - \psi(m, B)) \tag{14}$$

Because organs are scarce, it follows from eq. (8) that $\psi(B) < 1$. In fact, it is not feasible for the social planner to offer organs with certainty to sick young patients, for this would give a strong incentive to healthy young patients to misreport which would only be offset by the certainty of getting an organ when reporting the true health state, which requires more organs than what is available.

On the other hand, while guaranteeing organs to young healthy patients (i.e. setting $\psi(G) = 1$) solves the informational problem, it passes too many organs to healthy patients, leaving too many untreated sick patients as well. In fact, if a healthy patient gets an organ for sure, a young healthy patient receives interim utility $2\bar{u}$, the largest any patient can attain, and in particular strictly larger than the utility he would get by lying, this is (8) is slack. Hence, the social planner strictly prefers to decrease $\psi(G)$ by an small amount, and allocate those organs to junior sick patients by increasing $\psi(B)$ while respecting incentive compatibility.

Similarly, if ψ is optimal incentive compatibility for senior patients (11) should be binding: Since the social planner gets no benefit from passing organs to senior healthy patients, it would be optimal to set this inequality to hold with equality. This is actually doable, but some care needs to be put to not modify the continuation value of junior patients, $w^G(m)$, so that they do not have incentives to lie. Analogously, in the optimum the incentive compatibility constraint (12) binds.

These necessary properties of the optimal allocation rule when the planner's objective is to prioritize the medically worst-off patients are summarized in the following lemma.

Lemma 1 (Necessary properties-equity optimal) *Let ψ^* be an optimal allocation rule. Therefore,*

1. $\psi^*(B) < 1$.

2. $\psi^*(G) < 1$.

3. *Incentive compatibility constraints for senior patients (5) and junior patients (8) are binding.*

This lemma reduces the design problem to the choice of the probability of patients getting an organ as a function of their report when junior, this is $\psi(G), \psi(B), \psi(B, \cdot), \psi(G, \cdot)$.

Without further restrictions in the primitives, it is not possible to pin down the exact optimal allocation rule, nor general properties of it. In fact, the next theorem establishes that the optimal allocation rule can take very different forms depending on the environment parameters.

In what follows, a key parameter is how a healthy junior patient trade offs receiving an organ in the present against receiving an organ in the future. Formally, let $\underline{w}^G = \sum_{h \in \{B,G\}} p(h|G)u(h)$, the lowest continuation utility a junior healthy patient can expect. Let $L^G \triangleq 2\bar{u} - u(G) - \delta_G \underline{w}^G$ be the lifetime marginal benefit of a healthy young patient from getting an organ when young relative to never. Also, let $K^G \triangleq \bar{u} - \underline{w}^G$ expected gain of getting an organ in second period when healthy in first, relative to never. The marginal rate of intertemporal organ substitution is therefore

$$\rho = \frac{L^G}{\delta_G K^G} \tag{15}$$

Theorem 1 *The following statements are true:*

1. *For any of the allocation rules defined in Table 2, there exist parameters such that the chosen allocation rule minimizes (10) subject to incentive compatibility and feasibility constraints.*

2. *For any parameters satisfying the assumptions made so far, the optimal allocation rule is one of those defined in Table 2*

The proof is in the Appendix. The six allocation rules defined in Table 2 are those that satisfy the necessary Karush-Khun-Tucker conditions for optimality, and conditions in Lemma 1. For all these allocation rules, there exists admissible parameters such that are optimal.

The six allocation rules identified in Theorem 1 constitute the full set of possible solutions to the social planner problem. It highlights that the optimal allocation rule for a social welfare function that aims to minimize the mass of worst-off patients left untreated is highly dependent on the organ supply as well as parameters governing health evolution, patients preferences and survival. Broadly speaking, there are three types of optimal allocation rules:

Definition 1 (Taxonomy of optimal allocation rules)

(I) Sick junior patients served immediately, healthy juniors wait

Table 2: Possible optimal allocation rules for just criterion under asymmetric information. See Theorem 1.

Solution no.	$\psi(B)$	$\psi(B,\cdot)$	$\psi(G)$	$\psi(G,\cdot)$
1	x	0	x	0
2	A	0	0	ρA
3	B	0	$\frac{\rho B-1}{\rho-1}$	1
4	C	1	C	1
5	0	D	0	D
6	E	$\frac{1-\rho E}{1-E}$	0	1

$A = \frac{x}{(1-q)+q\delta_G\rho}$, $B = \frac{x(\rho-1)+q(1-\rho\delta_G)}{(\rho-1)+q(1-\rho\delta_G)}$, $C = \frac{x-(1-q)\delta_B-q\delta_G}{1-(1-q)\delta_B-q\delta_G}$, $D = \frac{x}{(1-q)\delta_B+q\delta_G}$, $E = -\frac{x-(1-q)\delta_B-q\delta_G}{(1-q)[\delta_B\rho-1]}$

(II) Uniform lottery, immediate service

(III) Uniform lottery, delayed service

Screening (I) may or may not be optimal. When screening is optimal, the planner leverages the history of reports made by patients when junior. When it is not optimal, a social planner who cares about minimizing the mass of untreated sick patients can find optimal to make patients to wait regardless of their report, despite the fact that screening takes no place (III).

To gain intuition, assume that organs are very scarce. Specifically, assume $x < \min\{(1-q)\delta_B, q\delta_G\}$. Even under this assumption of extreme organ shortage, all the types of optimal allocation the taxons are represented: Screening (solution 2), pooling with delay (solution 5) and without it (solution 1) are optimal allocation rules for some specification of health evolution, survival and preferences.[14]

Since, given the very small organ supply x, these three allocation rules are feasible and incentive compatible, which one is the optimal depends on the trade-off between efficiency and information rents in a specific economy.

[14]Solution 3 is not feasible, $\psi(G) < 0$, 4 because $\psi(B) < 0$, and 6 because it requires more than $q\delta_G$ organs.

Using binding BC, for any ψ incentive compatible the objective can be rewritten as:

$$
\begin{aligned}
&(1-q)[1+\delta_B P(B|B)] + q\delta_G P(B|G) \\
&- \{x - [q\psi(G) + q(1-\psi(G))\delta_G P(G|G)\psi(G,G) + (1-q)(1-\psi(B))\delta_B P(G|B)\psi(B,G)]\} \\
&\qquad\qquad - (1-q)\psi(B)\delta_B P(B|B) - q\psi(G)\delta_G P(B|G)
\end{aligned}
\tag{16}
$$

The first two terms, $(1-q)[1+\delta_B P(B|B)] + q\delta_G P(B|G)$, correspond to the total mass of total mass of sick patients in the market if no organs were allocated. The expression in curly brackets is the total amount of organs allocated to sick patients, which is the organs left after assigning to healthy patients the organs required to incentivize them to reveal their information, i.e. the information rents that must be paid, which in optimality are bounded away from zero by the incentive constraints. The last two terms are the additional dynamic effects of allocating organs to junior patients, which removes from the market patients who can be sick in the future.

Consider first the two pooling scenarios. On the one hand, pooling seniors improves efficiency by removing an additional mass $q\delta_G P(B|G)(D-1)x$ of senior sick patients who where healthy when junior, compared to pooling juniors. On the other hand, pooling seniors instead of juniors losses all the multiplicative dynamic effect of allocating organs to junior patients; for each organ allocated to senior there is an efficiency loss of

$$
\frac{1}{(1-q)\delta_B + q\delta_G}\delta_B P(B|B) - (1+\delta_B P(B|B)) < 0 \tag{17}
$$

relative to allocating the same organ to a junior. How the trade-off between these two effects resolves depends on the parameters.

To be more concrete, suppose $p(B|B) = p(B|G) = 1$, this is, every senior patient is sick. Moreover, assume $\delta_B = \delta_G = 1$. Hence, the mass of sick patients before organs are allocated is $1-q$ units of juniors plus 1 unit of senior patients. In such a case, the planner strictly prefers to allocate organs to juniors, for every allocated organ has an impact on the mass of sick patients in the market (those allocated to juniors healthy remove a patients who will be sick in the future), while pooling seniors misses this multiplicative effect. However, as the survival probabilities δ_G, δ_B decrease, pooling juniors "wastes" a fraction $q(1-\delta_G)x$ of organs, in the sense that their allocation does not remove sick patients from the market in the present nor in the future, but there is still a benefit of allocating organs to junior patients, which is that it removes an additional $\delta_B(1-q)x$ senior sick patients from the market. When the cost is larger than the benfit, i.e., when $q(1-\delta_G) > \delta_B(1-q)$ it is more efficient to allocate organs among seniors.

4.2 Principle of utility/utilitarian criterion

The incentive and feasibility constraints faced by the utilitarian planner are identical to those faced by the just social planner studied in the previous section. As before, I study the steady-states generated by stationary mechanisms that only condition on the history of reports of patients in the market, this is $\psi_s(m) = \psi_{s-1}(m) = \psi(m)$. Thus, the design problem for an utilitarian social planner becomes to maximize

$$\sum_m q(m) \left\{ \psi(m)2\bar{u} + (1-\psi(m)) \left[\delta_m \sum_{m'} p(m'|m)[u(m') + \psi(m,m')(\bar{u} - u(m'))] \right] \right\} \quad (18)$$

subject to the respective incentive constraints (11), feasibility constraint (13) and law of motion.

This is, the ex-ante total lifetime expected utility of patients arriving to the system in any period. Equivalently, this is the expect utility of a patient arriving to the system before their type is realized.

Because the per-period marginal utility a sick patient gets from an organ is larger than the one a healthy patient gets, and patients are assumed to be identical after undergoing transplantation, a version of Lemma 1 can be obtained for the utilitarian social planner.

Lemma 2 (Necessary properties of an utilitarian optimal allocation rule)

1. $\psi^*(B) < 1$.
2. $\psi^*(G) < 1$.
3. *Incentive compatibility constraints for senior (5) and junior patients (8) are binding.*

In comparison to the case of a social planner who maximizes the mass of organs passed to the sick patients, more can be said about the optimal allocation rule for a utilitarian social welfare function. The next lemma establishes that an utilitarian social planner never finds optimal to incentivize healthy junior patients to wait.

Lemma 3 $\psi^*(G,G) = 0$.

The intuition behind this lemma is that if the planner is going to allocate certain mass of organs to senior patients who declared to be healthy when junior, regardless of their report in the second period, both the planner and the patients are better off if the same mass of organs is allocated to them when they are junior. Alternatively, it is cheaper for the social planner to deliver interim utilities to junior healthy patients by allocating organs sooner than latter. The organs saved this way can be used to increase the utility a healthy junior patient gets from telling the truth.

The utilitarian social planner does not make healthy patients wait to be served. The utilitarian planner internalizes the disutility from not allocating organs to healthy patients when they are young, which comes from three different sources: lower quality of life, possible death and potentially becoming sick.

It remains to determine whether it is optimal for the utilitarian social planner to screen junior patients by making wait junior patients who claim to be sick. The next theorem establishes under which conditions it may be advantageous from a social perspective to do so.

Theorem 2 *Let* $S = -\left[\rho\delta_B(L^B - K^B) + (1-\rho\delta_B)q(L^B - L^G)\right]$ *with* K^G*,* L^G *and* ρ *as defined in Theorem 1, and* K^B*,* L^B *the analogous for a truly sick junior patient. Then, the optimal allocation rule for an utilitarian social planner is:*

- *If the weighted average* $S < 0$*,*

$$\psi^*(G) = \psi^*(B) = x, \quad \psi^*(G, \cdot) = \psi^*(B, \cdot) = 0$$

- *If the weighted average* $S > 0$*, and*
  - *organ supply is small, i.e.,* $x < \frac{1}{\rho}q + \delta_B(1-q)$

$$\psi^*(G) = \frac{x}{q + \delta_B(1-q)\rho}, \quad \psi^*(G, \cdot) = 0,$$
$$\psi^*(B) = 0, \quad \psi^*(B, \cdot) = \frac{x\rho}{q + \delta_B(1-q)\rho}$$

  - *organ supply is large, i.e.,* $x > \frac{1}{\rho}q + \delta_B(1-q)$

$$\psi^*(G) = \frac{[x - \delta_B(1-q)]\rho - q + 1 - x}{[1 - \delta_B(1-q)]\rho - q}, \quad \psi^*(G, \cdot) = 0$$
$$\psi^*(B) = \frac{[x - \delta_B(1-q)]\rho - q}{[1 - \delta_B(1-q)]\rho - q}, \quad \psi^*(B, \cdot) = 1$$

The following lemma provides a necessary condition for screening to be optimal.

Proposition 1 *If* $S > 0$ *then* $L^G - K^B < 0$*.*

For an utilitarian planner to find it optimal to screen and force sick patients to wait, it is necessary that a sick junior patient who survives and gets an organ when senior with certainty, experiments an expected marginal increase in his quality of life larger than the lifetime marginal benefit a healthy junior patient would get from undergoing transplantation when junior. This is, a sick patient who only enjoys a transplant for one period, should get more out of the transplant, in expectation, than a healthy patient who enjoys a transplant for two periods. Notice that this condition is necessary, but not sufficient.

4.3 Discussion

The current algorithm used by the OPTN to allocate organs is a naive attempt to materialize the principle of medical urgency. The desiderata, as in the allocation policies, is to prioritize currently sick patients, and among patients in the same health states, give priority to those who have been waiting the longest in the worst health state. With the algorithm, and if patients in good health state were sincere, they would receive organs only if sick patients in the same geographic location have declined them. In my setting, with undifferentiated organs, this translates to the following allocation rule:[15]

$$\psi(B,B) > \psi(B) = \psi(G,B) > \psi(B,G) > \psi(G,G) > \psi(G) \tag{19}$$

This allocation rule is not incentive compatible. To see why, note that $\psi(G,G) < \psi(G,B)$ and $\psi(B,G) < \psi(B,B)$, which violates the incentive constraint for senior patients. Hence, a senior patient has incentives to not revealing his information. More importantly, the utility a junior patient in state G gets from misreporting in the first period and tells the truth in the second is

$$\psi(B)2v+(1-\psi(B))\Big[u(G)+\delta_G[p(G|G)\psi(B,G)(v-u(G))+p(B|G)\psi(B,B)(v-u(G))+\underline{w}^G]\Big] \tag{20}$$

while the utility of telling the truth in both periods is

$$\psi(G)2v+(1-\psi(G))\Big[u(G)+\delta_G[p(G|G)\psi(G,G)(v-u(G))+p(B|G)\psi(G,B)(v-u(G))+\underline{w}^G]\Big] \tag{21}$$

Since $\psi(G,G) < \psi(B,G)$ and $\psi(G,B) < \psi(B,B)$, the term in squared brackets in expression (21) is smaller than the corresponding in expression (20) Therefore the utility of misreporting when junior and healthy is larger than the utility of telling the truth in both periods, and a

[15]The actual allocation is the equilibrium of these "intended" priorities with the strategic response of patients to them.

healthy young patient has incentives to lie.[16]

The key to prevent gaming of health states is to guarantee to patients who report to be healthy that they have a chance of getting an organ large enough so that they reveal their information, this is, left some information rents to healthy patients in order to incentivize them to reveal their information. This is the content of claims 1 and 2, and should be done to obtain incentive compatibility of the allocation rule regardless of the social welfare function.

However, the optimal way of providing these incentives depends on the social welfare function in a remarkable way.

For some parameters of the environment, the medical urgency objective may lead the planner to find optimal to force healthy patients to wait (no screening with delay and screening with healthy waiting in definition 1) so that they get sick, hence easing the trade-off between optimality and information rents. If the required information rents are too large, the planner can find that even forcing sick patients to wait is in his interest (no screening with delay).

In contrast, an utilitarian social planner never forces healthy patients to wait, for she internalizes the loss in well-being of a healthy patient that is forced to wait, which arises from three different sources: likelihood of death, lower quality of life, and the possibility of becoming sick.

5 Long-run objective and transitional dynamics

In this section, I study the transitional dynamics generated by the optimal allocation rule when the system starts from an arbitrary mass of senior patients. The purpose is to investigate the steady state stability. For economy of space, I restrict attention to the equitable objective, i.e. the one concerned with the worst-off patients. The analysis for the utilitarian one goes along the same lines.

A major challenge for studying the transitional dynamics of the transplantation system is the lack of a long run objective. In fact, both the policy guidelines (POLICY documentation)

[16]Note that if the patient lies in both periods gets

$$\psi(B)2v + (1-\psi(B))\Big[u(G) + \delta_G[p(G|G)\psi(B,B)(v-u(G)) + p(B|G)\psi(B,B)(v-u(G)) + \underline{w}^G]\Big]$$

while if he tells the truth in the first period and lies in the second,

$$\psi(G)2v + (1-\psi(G))\Big[u(G) + \delta_G[p(G|G)\psi(G,B)(v-u(G)) + p(B|G)\psi(G,B)(v-u(G)) + \underline{w}^G]\Big]$$

Since $\psi(G,B) < \psi(B,B)$, the same conclusion on the incentives to manipulate for junior patients follows if one factor in that the they have incentives to lie in the second period under the current allocation rule as given by (19).

and the medical ethics literature are silent on this crucial aspect, at least from a formal perspective. In principle, different long run objectives might lead to different transitional dynamics and properties of the optimal steady state. In view of this, In this section I adopt a standard social discount factor $\beta \in (0,1)$ and assume that the long run equitable objective is to minimize the total discounted sum of untreated sick patients in the market, starting from an initial state with some masses $\mu_0(B), \mu_0(G)$ of senior patients, this is,

$$\sum_{s=1}^{\infty} \beta^{s-1} \left(q(B)\left(1-\psi_s(h^s,B)\right) + \sum_{m\in\{B,G\}} \mu_{s-1}(m)\delta_m p(B|m)\left(1-\psi_s(h^s,m,B)\right) \right) \tag{22}$$

The planner's problem is to choose allocation probabilities $\{\psi_s\}_s$, to minimize the total discounted mass of untreated sick patients (22), subject to incentive compatibility (5)-(8) and feasibility (9). In general, since any mechanism ψ depends on the whole history of allocation decisions from time $s=0$ on (see subsection 3.2) , the problem of solving for the optimal mechanism ψ is an infinite dimensional optimization problem. To deal with it, I use dynamic programming and Lagrange multipliers in infinite dimensional spaces, two standard techniques in the literature on dynamic incentives (see for instance Spear et al. (1987), Abreu et al. (1990), and Golosov et al. (2016)).

This program underlines several crucial differences with the standard case, which in principle difficult an analytic approach. First, in the incentive compatibility constraint (8), present and future are not additive separable as usual. Similarly, eq. (7)-(6) are non-linear in the controls. Together, these two constraints also imply that the feasibility set is not convex. Moreover, the per-period payoff function is not convex nor concave. Hence, standard arguments cannot be used to show uniqueness of the optimal policy function, nor to study the transitional dynamics.[17]

The following Lemma gives necessary conditions for optimality which simplify the optimization problem significantly. It is equivalent to Lemmas 1 and 2; the proof is essentially the same, but additional care is required with book-keeping.

Lemma 4 *Let $\{\psi^*_s\}_{s=1}^{\infty}$ be an optimal allocation rule. For all s and h^s*

1. $\psi^*_s(h^s,B) < 1$.

2. $\psi^*_s(h^s,G) < 1$.

[17]These features also impose some numerical difficulties as well. They rule out using heuristics such as binary search that exploit the value function concavity to speed up the numerical solution. In addition, the non-linear constraints prevent one from using fast linear programming algorithms as those used in Doepke et al. (2006) and Phelan et al. (1991) to accelerate the solution of dynamic incentive problems.

3. Incentive compatibility constraints for senior (5) and junior patients (8) are binding.

4. For every s, the per-period budget constraint 9 binds.

I use these results and the definition of $\mu_s(m)$ to eliminate the continuation utilities w_t from the objective and constraints, so the above non-convex dynamic optimization problem reduces to a one with linear objective and constraints. More explicitly, let

$$F(\mu_{s-1}, \mu_s) = \mu_s(h^s, B) + \sum_m \mu_{s-1}(h^{s-1}, m)\delta_m p(B|m) - \frac{N^{se}}{N}\left[x - 1 + \sum_m \mu_s(h^s, m)\right] - \frac{\delta_B q(B)\delta_G q(G)}{N}\rho(p(B|B) - p(B|G))\left(\frac{\mu_{s-1}(B)}{q(B)} - \frac{\mu_{s-1}(G)}{q(G)}\right) \quad (23)$$

Thus, the objective becomes to choose $\{\mu_s\}_{s=1}^{\infty}$ to minimize

$$\sum_{s=1}^{\infty} \beta^{s-1} F(\mu_{s-1}, \mu_s) \quad (24)$$

subject to

$$0 < \mu_s(m) \leq q(m) \quad (25)$$

and

$$0 \leq x - 1 + \sum_m \mu_s(h^s, m) + \rho\delta_G q(G)\left(\frac{\mu_{s-1}(B)}{q(B)} - \frac{\mu_{s-1}(G)}{q(G)}\right) \leq \frac{N\mu_s(B)}{q(B)} \quad (26)$$

and

$$0 \leq x - 1 + \sum_m \mu_s(h^s, m) - \rho\delta_G\delta_B q(B)\left(\frac{\mu_{s-1}(B)}{q(B)} - \frac{\mu_{s-1}(G)}{q(G)}\right) \leq \frac{N\mu_s(G)}{q(G)} \quad (27)$$

The latter two correspond to the constraints on the probabilities of allocating organs to senior patients. For every pair $\mu_{s-1}(B), \mu_{s-1}(G)$, the latter four inequalities define a correspondence $\Gamma(\mu_{s-1}(B), \mu_{s-1}(G))$ of feasible pairs $\mu_s(B), \mu_s(G)$ which is nonempty, compact-valued and continuous. This, coupled with continuity and boundedness of F imply, through standard arguments, that a solution to the problem exists, and, that the sequential problem admits a recursive representation and that the Bellman operator that solves the recursive representation is unique (see Stokey 1989)

It is immediate that the optimal steady state calculated in section 4.1 is a steady state of this system. Moreover, using lagrange multipliers in infinite dimensional spaces (**luenberger1997optimization**; Golosov et al. 2016), one can verify that if this optimal steady state is unique in the steady state program, then it is the unique steady state satisfying the Euler-Lagrange equations for the infinite horizon program.

Next, I investigate the local stability of this steady state. The key question is whether small perturbations out the steady state drive the system back to the steady state. It is standard to analyze the stability of the nonlinear dynamical system generated by the necessary Euler-Lagrange equations by linearizing the system around the optimal steady state (Stokey 1989). Since the steady state is corner, special care should be placed to consider the complementary slackness conditions.

Theorem 3 *Suppose the optimal steady state is unique. Thus*

1. *If the optimal steady state is screening or pooling seniors, it is locally stable.*

2. *If the optimal steady state is pooling juniors, it is unstable.*

While this Theorem is informative about the stability of the optimal steady state near the steady state, it says nothing about its global stability. To investigate this issue, I explore numerically the optimal paths in the phase plane, which is a well known technique in the theory of dynamical systems. To do so, I exploit the numerical tractability of the recursive representation of the problem.

$$V(\mu) = \min_{\mu' \in \Gamma(\mu)} F(\mu, \mu') + \beta V(\mu') \tag{28}$$

with $\mu' \in \Gamma(\mu)$ if and only if

$$0 < \mu'(m) \leq q(m) \tag{29}$$

and

$$0 \leq x - 1 + \sum_m \mu'(m) + \rho \delta_G q(G) \left(\frac{\mu(B)}{q(B)} - \frac{\mu(G)}{q(G)} \right) \leq \frac{N\mu'(B)}{q(B)} \tag{30}$$

and

$$0 \leq x - 1 + \sum_m \mu'(m) - \rho \delta_G \delta_B q(B) \left(\frac{\mu(B)}{q(B)} - \frac{\mu(G)}{q(G)} \right) \leq \frac{N\mu'(G)}{q(G)} \tag{31}$$

I solve this dynamic program in Mathematica 12.0 using simple value function iteration in computational node with two Intel (R) Xeon (R) Gold 6132 2.6Ghz and 28 physical cores. Figures 2 to 4 display the phase plane corresponding to the (calculated numerically) optimal paths generated by the policies that solve the above dynamic program starting from different market compositions under various parameter configurations. Its construction is explained in detail in the Appendix. The parameters were chosen so that the assumption $x < \min\{(1-q)\delta_B, q\delta_G\}$ holds and the three types of optimal steady-states in Definition 1 are possible. The numerical solutions were computed using simple value function iteration.

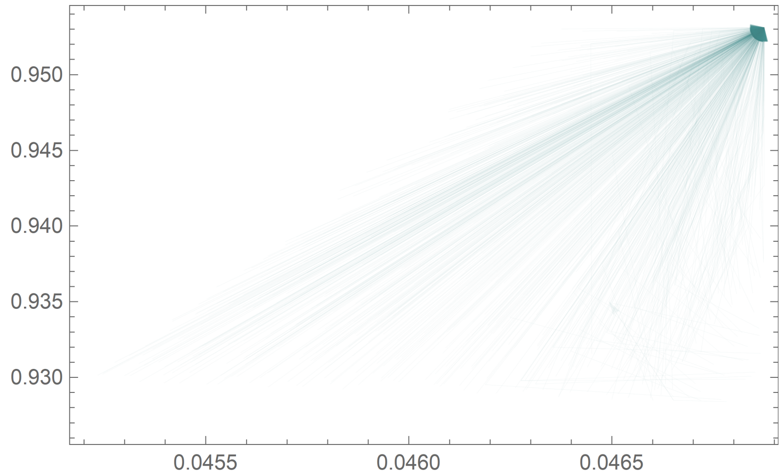


Figure 2: Global convergence to steady state pooling seniors.

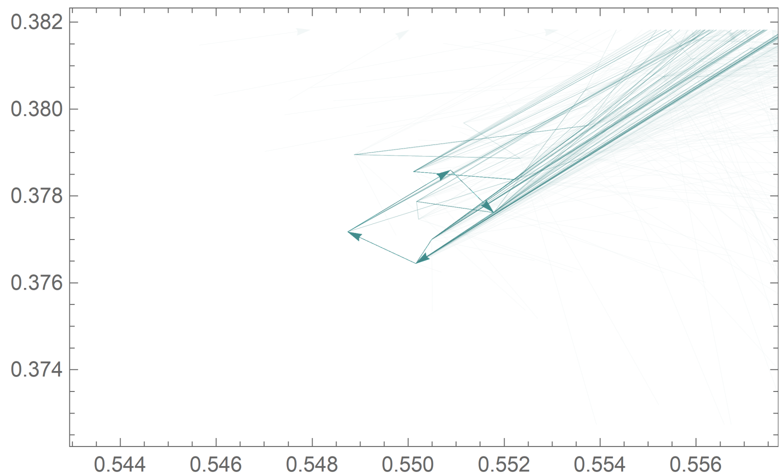


Figure 3: Long run instability of the steady state pooling juniors.

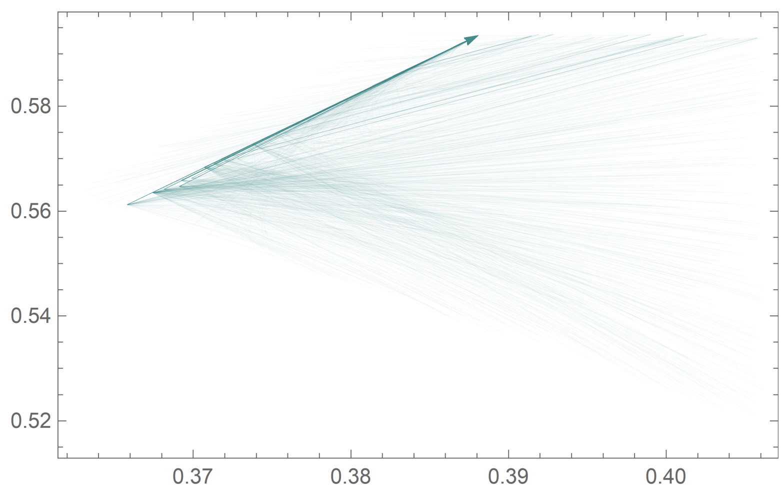


Figure 4: Global convergence to steady state screening.

Figures 2 and 4 clearly suggest that when the optimal steady-state is screening or pooling seniors, its globally asymptotically stable. For every initial market composition, the optimal path converges to the optimal steady state.

On the contrary, Figure 3 illustrates our finding in Theorem 3 on the unstability of pooling juniors when it is the unique optimal steady state: the optimal paths oscillate around the optimal steady state without reaching it.

6 Concluding remarks and further research avenues

I set up a model that formalizes the problem of waitlist priority manipulation through unnecessary medical therapies faced by the U.S. transplantation authority in the allocation of hearts and livers. The simple framework allows the derivation of optimal queue designs to implement different social welfare objectives in the presence of information asymmetries.[18]

The current allocation rule, which gives absolute priority to the most urgent patients and only considers patients' medical reports histories for breaking ties among equally sick patients, is not incentive compatible. It gives patients and physicians strong incentives to manipulate health conditions to raise priorities for transplants.

When the design problem reckons with the medical status manipulability, it becomes clear that any prioritization scheme which uses such information should incentivize individual patients and physicians to reveal it. Since transfers are repugnant in this setting (Roth 2007), one of the few tools available to policymakers to incentivize truthful reporting is conditioning allocation priorities on report histories.

Enlarging the set of allocation rules to include history-dependent ones enables the planner to incentivize truth-telling. When the social welfare objective is to maximize the number of organs allocated to medically urgent patients, providing these incentives is costly: organs must be given to non-urgent patients with some probability.

The trade-off between the cost and benefit of providing these incentives depends on the economy's parameters, so it is not evident that screening patients according to their medical urgency is always optimal. For a parameter region the costs offset the benefits; the planner gives up on incentivizing truthful reporting, and the optimal allocation policy is a uniform lottery among patients as a function of tenure. This lottery can be front-loaded or back-loaded. If the probability of patients becoming sick while waiting is large, back-loading the lottery and forcing patients to wait helps with the allocation problem, and minimizes the cost of screening.

[18]I use the term implementation in a layperson sense here. It has nothing to do with the sophisticated Implementation Theory, as should be evident at this point.

In sharp contrast, when the social welfare criterion is to maximize the total sum of patients' expected lifetime utility, the optimal allocation policy is a front-loaded uniform lottery among new arrivals to the system. The only exception is if a sick patient benefits more from waiting for an organ and getting it than what a healthy patient with the same tenure benefits from getting an organ. The cost of screening patients is not high, for truly urgent patients cannot manipulate their medical urgency, and healthy patients have no incentives to wait, so separation is achieved by serving junior healthy patients right away and forcing junior sick patients to linger.

From a policymaking perspective, two significant findings hold regardless of the social welfare objective. First, the optimality of screening patients and how to perform it depends on the economy's parameters. Second, when accounting for information asymmetries, a simple uniform lottery can be the optimal prioritization scheme. These two findings warrant the revision of the current allocation policies to the light of further empirical research.

For the sake of tractability, I made several simplifying assumptions. First, I made assumptions about patients' health evolution, age, lifetime horizon, choice and organ quality, and post-transplant survival. Second, I abstracted that transplant centers are long-run players who make treatment decisions for a set of patients instead of individual ones. Third, I assumed medical need information can be reported every period at no cost for the patients.

While the specific form of the optimal allocation rule is granted to depend on these assumptions, the general policy-making insights I found are robust to most of them. An exception that needs to be carefully assessed is the heterogeneity in the willingness to manipulate medical urgency across patients and physicians. If the fraction of transplant centers and patients willing to game the medical urgency is small, it may be the case that a uniform lottery renders not optimal compared with static screening.

References

Abreu, Dilip et al. (1990). "Toward a theory of discounted repeated games with imperfect monitoring". In: *Econometrica* 58.5, pp. 1041–1063.

Agarwal, Nikhil, Itai Ashlagi, et al. (2021). "Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys". In: *Econometrica, forthcoming.*

Agarwal, Nikhil, Charles Hodgson, et al. (2020). "Choices and Outcomes in Assignment Mechanisms: The Allocation of Deceased Donor Kidneys". In: *NBER Working Paper* 28064.

Alagoz, Oguzhan et al. (2005). “Incorporating biological natural history in simulation models: empirical estimates of the progression of end-stage liver disease”. In: *Medical Decision Making* 25.6, pp. 620–632.

Asrani, Sumeet K et al. (2010). “Organ allocation for chronic liver disease: model for end-stage liver disease and beyond”. In: *Current opinion in gastroenterology* 26.3, p. 209.

Ata, Baris et al. (2018). “Organ transplantation”. In: *Handbook of Healthcare Analytics: Theoretical Minimum for Conducting 21st Century Research on Healthcare Operations* 8.6, p. 4.

Atkeson, Andrew et al. (1992). “On efficient distribution with private information”. In: *The Review of Economic Studies* 59.3, pp. 427–453.

Bloch, Francis et al. (2017). “Dynamic assignment of objects to queuing agents”. In: *American Economic Journal: Microeconomics* 9.1, pp. 88–122.

Che, Yeon-Koo et al. (2021). “Optimal queue design”. In: *working Paper*, pp. 312–313.

Colvin-Adams, M et al. (2012). “Lung and heart allocation in the United States”. In: *American Journal of Transplantation* 12.12, pp. 3213–3234.

Committee, OPTN/UNOS Organ Transplantation (2016 (accessed June 30, 2020)). *Proposal to Modify the Adult Heart Allocation System. Board Briefing.* `http://engineering.purdue.edu/~mark/puthesis`.

Condorelli, Daniele (2012). “What money can’t buy: Efficient mechanism design with costly signals.” In: *Games Econ. Behav.* 75.2, pp. 613–624.

Dickert-Conlin, Stacy et al. (2019). “Allocating scarce organs: How a change in supply affects transplant waiting lists and transplant recipients”. In: *American Economic Journal: Applied Economics* 11.4, pp. 210–39.

Doepke, Matthias et al. (2006). “Dynamic mechanism design with hidden income and hidden actions”. In: *Journal of Economic Theory* 126.1, pp. 235–285.

Farhi, Emmanuel et al. (2007). “Inequality and social discounting”. In: *Journal of political economy* 115.3, pp. 365–402.

Freeman, Richard B et al. (2004). “Improving liver allocation: MELD and PELD”. In: *American Journal of Transplantation* 4, pp. 114–131.

Genie, Mesfin G et al. (2020). “The role of heterogeneity of patients’ preferences in kidney transplantation”. In: *Journal of Health Economics*, p. 102331.

Golosov, Mikhail et al. (2016). “Recursive contracts and endogenously incomplete markets”. In: *Handbook of Macroeconomics*. Vol. 2. Elsevier, pp. 725–841.

Green, Jerry R et al. (1986). “Partially verifiable information and mechanism design”. In: *The Review of Economic Studies* 53.3, pp. 447–456.

Guo, Yingni et al. (2018). *Dynamic allocation without money.* Tech. rep. Working Paper.

Kamm, Frances Myrna (1993). *Morality, mortality: Death and whom to save from it*. Vol. 1. Oxford University Press on Demand.

Kovac, Eugen et al. (2013). "Optimal stopping in a principal-agent model with hidden information and no monetary transfers". In: *Unpublished manuscript*.

Leshno, Jacob (2017). "Dynamic matching in overloaded waiting lists". In.

Li, Jin et al. (2017). "Power dynamics in organizations". In: *American Economic Journal: Microeconomics* 9.1, pp. 217–41.

Lipnowski, Elliot et al. (2018). "Repeated delegation". In: *Available at SSRN 2552926*.

Mabrouk, Mahasen et al. (2012). "Health-related quality of life in Egyptian patients after liver transplantation". In: *Annals of Hepatology* 11.6, pp. 882–890.

Margaria, Chiara (2016). *Queueing to learn*. Tech. rep. Working paper.

Merion, Robert M et al. (2011). "Evidence-based development of liver allocation: a review". In: *Transplant International* 24.10, pp. 965–972.

Moore, Christopher M et al. (2012). "The MELD Score: Limitations and consequences." In: *Transplantation* 242, pp. 0–0001.

Olszewski, Wojciech et al. (2020). "Efficient dynamic allocations in endowment economies with linear utilities and private information". In.

OPTN/UNOS (2018). *Manipulation of The Organ Allocation System Waitlist Priority through the Escalation of Medical Therapies*. Tech. rep. United Network of Organ Sharing.

Organ Transplantation, Task Force on (1986). *Organ transplantation: Issues and recommendations*.

Orr, James G et al. (2014). "Health related quality of life in people with advanced chronic liver disease". In: *Journal of hepatology* 61.5, pp. 1158–1165.

Parker, William, Kevin Chung, et al. (2020). "Practice changes at US transplant centers after the new adult heart allocation policy". In: *Journal of the American College of Cardiology* 75.23, pp. 2906–2916.

Parker, William, Edward R Garrity Jr, et al. (2017). "Trends in the use of inotropes to list adult heart transplant candidates at status 1A". In: *Circulation: Heart Failure* 10.12, e004483.

Phelan, Christopher et al. (1991). "Computing multi-period, information-constrained optima". In: *The Review of Economic Studies* 58.5, pp. 853–881.

Ran, Gege et al. (2021). "Between-center variation in high-priority listing status under the new heart allocation policy". In: *American Journal of Transplantation*.

Rao, Prashant et al. (2018). "Potential impact of the proposed revised UNOS thoracic organ allocation system". In: *Seminars in thoracic and cardiovascular surgery*. Vol. 30. 2. Elsevier, pp. 129–133.

Roth, Alvin E (2007). “Repugnance as a Constraint on Markets”. In: *Journal of Economic perspectives* 21.3, pp. 37–58.

Rowe, Ian et al. (2007). “Prognostic models in cirrhosis: an aid but not a replacement for clinical judgement”. In: *Liver International* 27.5, pp. 595–597.

Scanlon, Dennis P et al. (2004). “Does competition for transplantable hearts encourage ‘gaming’of the waiting list?” In: *Health Affairs* 23.2, pp. 191–198.

Schummer, James (2020). *Influencing waiting lists*. Tech. rep. Working Paper.

Sharma, Pratima et al. (2005). “Transplantation of the Liver”. In: ed. by Ronald Busuttil et al. Elsevier. Chap. Monitoring and Canre of the Patient Before Liver Transplantation, pp. 473–489.

Snyder, Jason (2010). “Gaming the liver transplant market”. In: *The Journal of Law, Economics, & Organization* 26.3, pp. 546–568.

Spear, Stephen E et al. (1987). “On repeated moral hazard with discounting”. In: *The Review of Economic Studies* 54.4, pp. 599–617.

Stevenson, Lynne Warner (2013). “The urgent priority for transplantation is to trim the waiting list”. In: *The Journal of Heart and Lung Transplantation* 32.9, pp. 861–867.

Stevenson, Lynne Warner et al. (2016). “Major advantages and critical challenge for the proposed United States heart allocation system”. In: *The Journal of Heart and Lung Transplantation* 35.5, pp. 547–549.

Stewart, Charmaine A et al. (2007). “Hepatic encephalopathy as a predictor of survival in patients with end-stage liver disease”. In: *Liver transplantation* 13.10, pp. 1366–1371.

Stokey, Nancy L (1989). *Recursive methods in economic dynamics*. Harvard University Press.

Sullivan, Colin D (2021). “Eliciting Preferences over life and death: experimental evidence from organ transplantation”. In.

Thakral, Neil (2016). *The public-housing allocation problem*. Tech. rep. Technical report, Technical report, Harvard University.

Trzepacz, Paula T et al. (2000). *The transplant patient: Biological, psychiatric and ethical issues in organ transplantation*. Cambridge University Press.

Veatch, RM (2004). “Justice, Utility, and Organ Allocation”. In: *Ethical, Legal, and Social Issues in Organ Transplantation. Berlin: Pabst Science Publishers*, pp. 57–67.

Veatch, Robert M (2000). *Transplantation ethics*. Georgetown University Press.

Yang, Linda S et al. (2014). “Liver transplantation: a systematic review of long-term quality of life”. In: *Liver international* 34.9, pp. 1298–1313.

Appendices

Proof of Lemma 1.

1. In text.

2. Suppose not, i.e. $\psi(G) = 1$. Thus, no healthy junior patient is left waiting and all the mechanisms with $\psi(G,G), \psi(G,B) \in [0,1]$ are payoff equivalent. It suffices to restrict attention to $\psi(G,\cdot) = 0$. Consider decreasing $\psi(G)$ from 1 to $1-\epsilon$, $\epsilon > 0$. This frees a mass $\epsilon q(G)$ of organs than can now be allocated to junior sick patients. This increases the mass of organs allocated to junior sick patients from $\psi(B)q(B)$ to $\psi(B)q(B) + \epsilon q(G)$, hence the chances of getting an organ would be now $\psi(B) + \epsilon\frac{q(G)}{q(B)}$. By doing so, the change in welfare is

$$\epsilon q(G)p(B|G)\delta_G - q(B)\epsilon\frac{q(G)}{q(B)}[1 + \delta_B(1 - \psi(B,B))P(B|B)] < 0$$

   implying a net welfare improvement. It is left to see if such change is incentive compatible.I want to pick $\epsilon > 0$ such that $(1-\epsilon)2\bar{u} + (1-(1-\epsilon))[u(G) + \delta_G \underline{w}] \geq (\psi(B)+\epsilon\frac{q(G)}{q(B)})2\bar{u} + (1-\psi(B)-\epsilon\frac{q(G)}{q(B)})[u(G)+\delta_G w^G(B)]$, this is $2\bar{u} - \epsilon[2\bar{u} - u(G) - \delta^G \underline{w}] \geq \psi(B)2\bar{u} + (1-\psi(B))[u(G)+\delta_G w^G(B)] + \epsilon\frac{q(G)}{q(B)}[2\bar{u} - u(G) - \delta_G w^G(B)]$. Since $\psi(B) < 1$, equation (12) is slack, i.e. $2\bar{u} > \psi(B)2\bar{u} + (1-\psi(B))[u(G)+\delta_G w^G(B)]$ so picking $\epsilon > 0$ this way is always possible by choosing it very small.

3. $\psi^*(G,G) = \psi^*(G,B)$. If not, SP can improve by passing more organs after history G, B and decreasing organs offered to patients with history G, G. This is always doable, for $\bar{u} - u(B) > \bar{u} - u(G)$, so that the period 2 ex-ante utility delivered to a patient after history G can still be delivered by increasing $\psi(G,B)$ by $\epsilon' = \frac{p(G|G)}{p(B|G)}\frac{\bar{u}-u(G)}{\bar{u}-u(B)}\epsilon$ where ϵ is the decrease in $\psi(G,G)$. The change in the organs needed is $q(G)(1 - \psi(G))\delta_G(\epsilon' p(B|G) - \epsilon p(G|G)) < 0$ so the change is always feasible. This obviously improves welfare by removing more sick patients from the market.

   $\psi^*(B,G) = \psi^*(B,B)$. First, notice that the argument above does not work. It breaks down because the organs needed to sustain the change are now proportional to

$$\epsilon' p(B|B) - \epsilon p(G|B) = \epsilon p(B|B)\left(\frac{p(G|G)}{p(B|G)}\frac{\bar{u}-u(G)}{\bar{u}-u(B)} - \frac{p(G|B)}{p(B|B)}\right)$$

   which has an undefined sign, for $\frac{p(G|G)}{p(B|G)} > \frac{p(G|B)}{p(B|B)}$.

   Assume not, so $\psi^*(B,G) > \psi^*(B,B)$ but consider instead decreasing $\psi(B,G)$ by

ϵ. Thus, the interim utility obtained by a junior healthy patient who lies decreases, which is fine from IC perspective and does not affect welfare. This frees an amount of organs $m = q(B)(1 - \psi(B))\delta_B p(G|B)\epsilon$. Split this mass of organs among junior patients randomly. Thus, junior patients get $\psi(G)q(G)+mq(G)$ and $\psi(B)q(B)+mq(B)$ respectively. This removes more sick patients from the market, which is obviously welfare improving:

$$-mq(B)[1 + \delta_B p(B|B)(1 - \psi(B, B)] - mq(G)\delta_G p(B|G)(1 - \psi(G, G)) < 0$$

Need to check if this is incentive compatible. Recall incentive compatibility

$$\psi(G)2\bar{u} + (1 - \psi(G))[u(G) + \delta_G w^G(G)] > \psi(B)2\bar{u} + (1 - \psi(B))[u(G) + \delta_G w^G(B)]$$

with the change,

$$(\psi(G)+m)2\bar{u}+(1-\psi(G)-m)[u(G)+\delta_G w^G(G)] > (\psi(B)+m)2\bar{u}+(1-\psi(B)-m)[u(G)+\delta_G w^G(B)]$$

i.e.

$$\psi(G)2\bar{u}+(1-\psi(G))[u(G)+\delta_G w^G(G)]-m\delta_G w^G(G) > \psi(B)2\bar{u}+(1-\psi(B))[u(G)+\delta_G w^G(B)]-m\delta_G w^G(B$$

i.e.

$$\psi(G)2\bar{u}+(1-\psi(G))[u(G)+\delta_G w^G(G)]-\psi(B)2\bar{u}-(1-\psi(B))[u(G)+\delta_G w^G(B)] > m\left[\delta_G w^G(G) - \delta_G w^G(\right.$$

which can be satisfied by picking m small enough, i.e. ϵ small enough.

To see that IC-youngs is binding, assume not. There are two cases. First, if $\psi^*(G) > 0$, decrease it by ϵ, this decreases organs allocated to healthy junior patients by $\epsilon q(G)$ increasing the mass of unserved junior healthy patients by the same amount, so to sustain the same $\psi(G, \cdot)$ a mass $\epsilon q(G)\delta_G\psi(G, \cdot)$ of organs are required (from previous numerals $\psi(G, \cdot) = \psi(G, G) = \psi(G, B)$), so a net mass of $M = \epsilon q(G)(1 - \delta_G\psi(G, \cdot))$ organs is available and can be allocated to sick junior patients.

$$\begin{aligned}
&- M[1 + \delta_B p(B|B)(1 - \psi(B, B)] + \epsilon q(G)\delta_G p(B|G)(1 - \psi(G, G)) \\
&- (1 - \delta_G\psi(G, \cdot))[1 + \delta_B p(B|B)(1 - \psi(B, B)] + \delta_G p(B|G)(1 - \psi(G, G)) \\
&- 1 + \delta_G[\psi(G, \cdot) + p(B|G)(1 - \psi(G, G))] - \delta_B p(B|B)(1 - \psi(B, B))(1 - \delta_G\psi(G, \cdot)) < 0
\end{aligned}$$

where the inequality follows from the fact that the term in squared brackets is strictly less than 1. It is left to see if this change respects incentive compatibility,

$$(\psi(G)-\epsilon)2\bar{u}+(1-\psi(G)+\epsilon)[u(G)+\delta_G w^G(G)] \geq \\ (\psi(B)+\frac{M}{q(B)})2\bar{u}+(1-\psi(B)-\frac{M}{q(B)})[u(G)+\delta_G w^G(B)]$$

rearranging,

$$\psi(G)2\bar{u}+(1-\psi(G))[u(G)+\delta_G w^G(G)]-\psi(B)2\bar{u}-(1-\psi(B))[u(G)+\delta_G w^G(B)] \geq \\ \frac{M}{q(B)}2\bar{u}-\frac{M}{q(B)}[u(G)+\delta_G w^G(B)]-\epsilon 2\bar{u}-\epsilon[u(G)+\delta_G w^G(G)]$$

since by assumption the left hand side of the inequality is strictly positive, one can always pick ϵ small enough so that the inequality holds.

If $\psi^*(G)=0$, decrease $\psi^*(G,\cdot)$ by an small amount ϵ, and increase $\psi^*(B)$ by $\frac{q(G)\delta_G\epsilon}{q(B)}$. This increases the mass of sick patients served by $q(G)\delta_G\epsilon p(G|G)$, improving welfare, and obviously respects feasibility. ϵ can be chosen small enough so that incentive compatibility is respected.

4. If the budget constraint does not bind, a fraction of the unused organs can be allocated to healthy junior patients, which is always incentive compatible, and improves welfare by removing from the market a fraction of patients who will be sick when seniors.

■

Proof of Theorem 1. The lagrangian for the optimization problem is

$$(1-q)[1-\psi(B)+(1-\psi(B))\delta_B P(B|B)(1-\psi(B,B))]+q(1-\psi(G))\delta_G P(B|G)(1-\psi(G,G) \\ -\xi_1(U_1(B,\psi,G)-U_1(G,\psi,G)) \\ -\xi_2[(1-q)[\psi(B)+(1-\psi(B))\delta_B\psi(B,B)]+q[\psi(G)+(1-\psi(G))\delta_G\psi(G,G)]-x] \\ +\sum_{h=G,GG,B,BB}\lambda^0(h)\psi(h)-\lambda^1(h)(\psi(h)-1)$$

where ξ_1 is the lagrange multiplier for the junior's incentive constraint, ξ_2 is the one associated with the feasibility constraint, $\lambda^0(h)$ are the lagrange multipliers of the non-negativity constraints, and $\lambda^1(h)$ correspond to the constraints that $\psi(\cdot)$ are smaller than 1.

To the necessary first order conditions, add the following conditions implied by lemma 1:

$$\lambda^1(G)=\lambda^1(B)=0$$

In addition, binding feasibility requires $\lambda^0(G)\lambda^0(G,\cdot) = 0$. Thus, the first order conditions reduce to

$$\lambda_1 - (1-q)(-1 - P(B|B)\delta^B(1-\psi(B,B))) - (1-q)\xi_2(1-\delta^B\psi(B,B)) + \xi_1(-L^G + K^G\delta^G\psi(B,B)) = 0 \tag{32}$$

$$\lambda_5 + P(B|B)(1-q)\delta^B(1-\psi(B)) - (1-q)\delta^B\xi_2(1-\psi(B)) + \xi_1(-K^G\delta^G + K^G\delta^G\psi(B)) = 0 \tag{33}$$

$$\lambda_3 + P(B|G)q\delta^G(1-\psi(G,G)) - q\xi_2(1-\delta^G\psi(G,G)) + \xi_1(L^G - K^G\delta^G\psi(G,G)) = 0 \tag{34}$$

$$\lambda_7 + P(B|G)q\delta^G(1-\psi(G)) - q\delta^G\xi_2(1-\psi(G)) + \xi_1(K^G\delta^G - K^G\delta^G\psi(G)) = 0 \tag{35}$$

$$-K^G\delta^G\psi(B,B) - \psi(B)(L^G - K^G\delta^G\psi(B,B)) + K^G\delta^G\psi(G,G) + \psi(G)(L^G - K^G\delta^G\psi(G,G)) = 0 \tag{36}$$

$$x - (1-q)(\psi(B) + \delta^B(1-\psi(B))\psi(B,B)) - q(\psi(G) + \delta^G(1-\psi(G))\psi(G,G)) = 0 \tag{37}$$

In addition, binding incentive compatibility requires $\lambda^0(B)\lambda^0(B,\cdot) = 0$. I input this enlarged set of first order conditions to Mathematica, which finds 13 candidates to optimal solution. The conditions on λ's and ψ's, and comparison of payoffs, reduce this set to the six solutions as given by table 2.

1. Numerical inspection shows that there exists feasible parameters for which each of these eight solutions is in fact optimal.

2. Follows from the necessity of the first order conditions.

■

Proof of Lemma 2.

1. There are no enough organs to set $\psi^*(B) = 1$ and satisfy incetive compatibility for juniors.

2. Note that since $u(B) < u(G)$ and $\delta_B < \delta_G$, the increase in lifetime utility for a junior sick patient who gets an organ is the largest. This and previous numeral implies $\psi^*(G) < 1$.

3. $\psi^*(B,G) = \psi^*(B,B)$. Otherwise, one can increase $\psi^*(B,B)$ and decrease $\psi^*(B,G)$, in a way that the first period interim utility delivered to a healthy patient who lies in the first period is the same. Since $u(B) < u(G)$ any ex ante period 2 utility can be delivered with less organs by increasing $\psi^*(B,B)$. The organs saved can be allocated to junior patients wo report to be healthy, which is incentive compatible and strictly increases planner's payoff.

$\psi^*(G,G) = \psi^*(G,B)$. If not, SP can improve by passing more organs after history G,B and decreasing organs offered to patients with history G,G. As before, this is doable. In contrast to the medical urgency case, this is not strictly preferred by society, which is infact indifferent. However, the excess of organs can be allocated to young healthy patients, therefore increasing the social welfare.

IC-young is binding. If not, decrease $\psi^*(G)$ (if larger than zero) or $\psi^*(G,G)$ (if $\psi^*(G) = 0$) by an small amount, and increase $\psi^*(B)$, in a way that incentive compatibility and feasibility are still respected. This is always possible and increase the mass of organs allocated to sick junior patients, which increases the total social welfare.

■

Proof of Lemma 3. Let

$$w^G = \psi(G)2v + (1-\psi(G))[u(G) + \delta_G W^G(G)]$$

with $W^G(G) = \psi(G,G)[v - \underline{W}^G] + \underline{W}^G$. Define $\tilde{\psi}$ to coincide with ψ over the paths containing B as initial report, $\tilde{\psi}(G,G) = 0$, and $\tilde{G}$ s.t.

$$w^G = \tilde{\psi}(G)2v + (1-\tilde{\psi}(G))[u(G) + \delta_G \underline{W}^G]$$

Such $\tilde{\psi}$ always exists, for $w^G \in [u(G) + \delta_G \underline{W}^G, 2v]$. Since it delivers the same interim utility to young healthy patients, it respects incentive compatibility.

$$\tilde{\psi}(G) = \frac{w^G - u(G) - \delta_G \underline{W}^G}{2v - u(G) - \delta_G \underline{W}^G}$$

or

$$\tilde{\psi}(G) = \frac{\psi(G)[2v - u(G) - \delta_G \underline{W}^G] + (1-\psi(G))\delta_G \psi(G,G)[v - \underline{W}^G]}{2v - u(G) - \delta_G \underline{W}^G}$$

this is

$$\tilde{\psi}(G) = \psi(G) + (1-\psi(G))\delta_G \psi(G,G)\frac{v - \underline{W}^G}{2v - u(G) - \delta_G \underline{W}^G}$$

so one concludes $\tilde{\psi}(G) < \psi(G) + (1-\psi(G))\delta_G \psi(G,G)$ and therefore the quantity of organs used by $\tilde{\psi}$ is less that used by ψ.

The remaining organs can be distributed randomly among young patients. This strictly increases the payoff. ■

Proof of Theorem 2. Proceeding in a similar fashion than in the proof of Theorem 1, but enriching now the first order conditions with the results in claims 2 and 3, Mathematica finds three allocation rules are obtained that satisfy the conditions on λ's being positive and

ψ's in $[0,1]$. S, as defined in the Theorem as well as x determine the signs of the multipliers and therefore which allocation is the optimal one, in a unique way. ■

Proof of Proposition 1. Recall

$$S = -L^G\delta_B[L^B - K^B] + q(L^B - L^G)(L^G\delta_B - K^G\delta_G)$$

Since $L^B > L^G$, it follows that

$$\begin{aligned} S &< -L^G\delta_B[L^B - K^B] + q(L^B - L^G)L^G\delta_B \\ &< -L^G\delta_B[L^B - K^B] + (L^B - L^G)L^G\delta_B \\ &< -L^G\delta_B[K^B - L^G] \end{aligned}$$

Therefore, if $S > 0$ then $K^B - L^G > 0$, as desired. ■

Proof of Lemma 4.

1. $\psi_s^*(h^s, B) < 1$. Suppose not for some s. Thus, it has to be the case that $\psi_s^*(h^s, G) = 1$ to satisfy incentive compatibility of junior patients (8), which is not feasible because $x < 1$.

2. $\psi_s^*(h^s, G) < 1$. Assume not for some period s, i.e., it is optimal to set $\psi_s^*(h^s, G) = 1$. Along the lines of the proof for Lemma 1, it suffices to restrict attention to $\psi_{s+1}^*(h^{s+1}, G, \cdot) = 0$, consider the net change in welfare from transferring a mass $\epsilon q(G)$ organs from junior healthy patients to junior sick candidates in period s,

$$\epsilon q(G)p(B|G)\delta_G - q(B)\epsilon\frac{q(G)}{q(B)}[1 + \beta\delta_B(1 - \psi_{S+1}(B,B))P(B|B)] < 0$$

   so the change is welfare improving. The proof of the existence of such $\epsilon > 0$ such that this change is also incentive compatible, is identical to the one in Lemma 1.

3. That in optimality $\psi_s^*(h^s, G, G) = \psi_s^*(h^s, G, B)$ follows from exactly the same arguments as in the corresponding proof for Lemma 1.

   On the other hand, to see that $\psi_s^*(h^s, B, G) = \psi_s^*(h^s, B, B)$, assume it is not the case, so $\psi_s^*(h^s, B, G) > \psi_s^*(h^s, B, B)$. Decrease $\psi_s^*(h^s, B, G)$ by ϵ, this makes more slack the period $s-1$ incentive compatibility constraint for junior healthy patients, frees $m = q(B)(1 - \psi_{s-1})\delta_B p(G|B)\epsilon$ organs in period s and does not change net welfare. As before, these organs can now be split among period s junior patients randomly, improving welfare. With this change, the period s junior patients incentive constraint becomes

$$(\psi_s(G)+m)2\bar{u}+(1-\psi_s(G)-m)[u(G)+\delta_G w^G_{s+1}(G)] \geq \\ (\psi_s(B)+m)2\bar{u}+(1-\psi_s(B)-m)[u(G)+\delta_G w^G_{s+1}(G)]$$

which as before can be rewritten as

$$\psi_s(G)2\bar{u}+(1-\psi_s(G))[u(G)+\delta_G w^G_{s+1}(G)]-\psi_s(B)2\bar{u}-(1-\psi_s(B))[u(G)+\delta_G w^G_{s+1}(B)] \geq \\ m\left[\delta_G w^G_{s+1}(G)-\delta_G w^G_{s+1}(B)\right]$$

in contrast to Lemma 1 for the steady-state program, this time the constraint can be binding, for the organs taken from s period seniors do not decrease the expected interim utility of s period juniors. If it is strict, then the same argument goes through. However, if it is binding, the only way to satisfy this constraint is choosing $m=0$. This problem can be overcame easily. Instead of spliting m proportionally, pass a fraction ϕ to junior healthy and the rest to sick patients, or even throw them away. the IC constraint is thus

$$(\psi_s(G)+\phi m)2\bar{u}+(1-\psi_s(G)-\phi m)[u(G)+\delta_G w^G_{s+1}(G)] \geq \\ \psi_s(B)2\bar{u}+(1-\psi_s(B))[u(G)+\delta_G w^G_{s+1}(G)]$$

and this change improves welfare by decreasing the mass of unserved sick patients by $-\phi m q(G)\delta_G p(B|G)(1-\phi_S(G,B))$.

4. Forthcoming.

5. Forthcoming.

■

Proof of Theorem 3. Forthcoming. ■