Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Alleviating the Patient's Price of Privacy Through a Partially Observable Waiting List

Burhaneddin Sandıkçı, Lisa M. Maillart, Andrew J. Schaefer, Mark S. Roberts,

To cite this article:
Burhaneddin Sandıkçı, Lisa M. Maillart, Andrew J. Schaefer, Mark S. Roberts, (2013) Alleviating the Patient's Price of Privacy Through a Partially Observable Waiting List. Management Science 59(8):1836-1854. http://dx.doi.org/10.1287/mnsc.1120.1671

MANAGEMENT SCIENCE
Vol. 59, No. 8, August 2013, pp. 1836–1854
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1120.1671

Alleviating the Patient's Price of Privacy Through a Partially Observable Waiting List

Burhaneddin Sandıkçı
Booth School of Business, University of Chicago, Chicago, Illinois 60637,
burhan@chicagobooth.edu

Lisa M. Maillart, Andrew J. Schaefer
Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261
{maillart@pitt.edu, schaefer@ie.pitt.edu}

Mark S. Roberts
Department of Health Policy and Management, University of Pittsburgh, Pittsburgh, Pennsylvania 15261,
mroberts@pitt.edu

In the United States, end-stage liver disease patients join a waiting list and then make accept/reject decisions for transplantation as deceased-donor organs are offered to them over time. These decisions are largely influenced by the patient's prospect for future offers, which can be ascertained most accurately by knowing the entire composition of the waiting list. Under the current transplantation system, however, the United Network for Organ Sharing (UNOS), in an effort to strike a balance between privacy and transparency, only publishes an aggregated version of the waiting list. However, it is not clear whether the published information is good enough (compared with perfect information) to help patients make optimal decisions that maximize their individual life expectancies. We provide a novel model of this accept/reject problem from an individual patient's perspective using a partially observed Markov decision process (POMDP) framework, which incorporates the imperfect waiting list information as published currently into the patient's decision making. We analyze structural properties of this model. In particular, we establish conditions that guarantee a monotone value function and a threshold-type optimal policy with respect to the partially observable rank state that captures the imperfect waiting list information. Furthermore, we develop an improved solution methodology to solve a generic POMDP model. This solution method guarantees, for any fixed grid, the best possible approximation to the optimal value function by solving linear programs to compute the optimal weights used for the approximation. Finally, we compare, in a clinically driven numerical study, the results of this model with those of an existing Markov decision process model that differs from our model in assuming the availability of perfect waiting list information. This comparison allows us to assess the quality of the published imperfect information as measured by a patient's so-called price of privacy (i.e., the opportunity loss in expected life days due to a lack of perfect waiting list information). Previous work estimates a significant loss in a patient's life expectancy, on average, when the patient has no waiting list information compared with full information. In this paper, we find that the currently published partial information is nearly sufficient to eliminate this loss, resulting in a negligible price of privacy and supporting current UNOS practice.

Key words: dynamic programming; partially and completely observable Markov decision process models; medical decision making; liver transplantation; value of information
History: Received February 9, 2011; accepted September 30, 2012, by Assaf Zeevi, stochastic models and simulation. Published online in *Articles in Advance* March 4, 2013.

1. Introduction

Organ transplantation, which significantly improves physical and social functioning, mental health, and overall quality of life of the transplant recipients (Dew et al. 1997), is the only therapy for many patients with end-stage liver disease (ESLD), such as cirrhosis and hepatitis, with 10-year survival rates exceeding 60% (U.S. Department of Health and Human Services (HHS) 2008). Despite efforts to increase the number of donated livers, there remains a wide gap between the supply and demand for such organs. Currently, there are more than 16,000 candidates in the United States waiting for a donor liver, yet only 6,000 liver transplants are performed in a typical year (U.S. Department of Health and Human Services 2008). More than 96% of these transplants use deceased-donor livers rather than living-donor livers. This paper considers the decision faced by ESLD patients without a living donor, namely, the decision of whether to accept or reject deceased-donor livers offered to them for transplantation. (Modeling the decisions of patients with an available living donor is possible through an

extension similar to that of Alagoz et al. 2007a.) Furthermore, unlike any prior research in this area, but consistent with current practice, we analyze this decision process assuming a partially observable waiting list. We focus on adult ESLD patients only.

As we describe in §2.1, the current allocation policy gives the patient (or the surgeon acting on her behalf) freedom to refuse any offer without penalty. Despite the scarcity of donated organs, organ acceptance rates are low; only 19% of offers are accepted (U.S. Department of Health and Human Services 2008). These decisions are made primarily based on patient health, the quality of the donated organ, and the patient's prospect for future organ offers.

In assessing the patient's prospect for future offers, the composition of the waiting list is important. Historically, this information has been unavailable to decision makers. Now, however, an aggregated version of the waiting list is published on the Internet (United Network for Organ Sharing (UNOS) 2010b). Intuitively, UNOS might prefer to avoid publishing the entire waiting list to prevent potential violation of patient privacy rights. The primary motivation of this paper is to analyze the effect of using this imperfect information when making the accept/reject decisions.

Sandıkçı et al. (2008) introduce the concept of a patient's price of privacy (see §6.2) to measure the value of accessing perfect waiting list information and estimate the average price (reported as the percent increase in the patient's life expectancy) to be around 5% when compared with no information. Utilizing the published partial description of the waiting list should reduce their estimates; however, the question of whether the published partial information is good enough to effectively eliminate a patient's price of privacy is still open.

In this paper, we address this open question by developing a partially observed Markov decision process (POMDP) model that incorporates the aggregated waiting list information available on the Internet. Our model builds upon the completely observable model of Sandıkçı et al. (2008). However, whereas they assume perfect waiting list information, we relax this assumption and consider imperfect waiting list information, which is consistent with current practice. As a result, whereas they deal with exact state representations of the actual system, we deal with probability measures defined over those (partially observable) states of the system. This distinction makes our model a better representation of reality, but it also makes our problem significantly more difficult to analyze and solve. Furthermore, we provide a more accurate probability model than the one available in Sandıkçı et al. (2008).

We analyze several structural properties of the resulting POMDP model. Among others, we establish conditions that guarantee a monotone value function and a threshold-type optimal policy with respect to the partially observable state of the model. These conditions are easy to check and our numerical study indicates that although the conditions are sometimes (mildly) violated, the optimal policy maintains its threshold structure. However, previous results (including those in Sandıkçı et al. 2008) do not generalize to the results proven in this paper.

Furthermore, solving a POMDP model is significantly harder than solving its Markov decision process (MDP) counterpart. For this purpose, given a fixed grid (i.e., finite set of vectors), we propose a solution methodology that guarantees the best possible approximation to the optimal value function. To find the optimal interpolation weights used in the approximation, this methodology requires solving linear programs at each iteration of a solution algorithm (such as value iteration algorithm). Moreover, in a partially observable setting, we demonstrate how one can take advantage of special problem structure when constructing a concise grid, which is particularly important as a standard uniform grid considers vectors that are likely to never be visited throughout the stochastic decision process.

Finally, we compare the results of our model with those of the completely observable model of Sandıkçı et al. (2008). This comparison allows us to assess the quality of the currently available waiting list information as measured by a patient's price of privacy (i.e., the opportunity loss in expected life-days due to a lack of perfect information). We find through our comparisons that the price of privacy can be almost completely eliminated through the aggregated waiting list information published in current practice.

Our findings have a number of implications for both policy makers and individual patients. We present compelling evidence that policy makers may continue revealing aggregated descriptions of the transplant waiting list without compromising the privacy of individual patients. Furthermore, we demonstrate that the published partial information can be used to effectively eliminate any potential loss in patients' life expectancies due to a lack of perfect waiting list information. Individual decision makers can also use our model as a decision support tool to help them make their choices when faced with an organ offer. Our results indicate that individuals who would benefit the most from using our model are younger patients, and those with less severe diseases as well as those that join the transplant waiting list from more densely populated regions.

The rest of this paper is organized as follows. We provide a brief overview of the current liver allocation system along with a review of the relevant literature in §2. We formulate our POMDP model in §3 and analyze its structural properties in §4. In §5, we

discuss the formation of a concise grid and propose a grid-based solution methodology for the POMDP model formulated in §3. We summarize extensive numerical results in §6 and conclude the paper in §7 with a summary of the paper along with its limitations and some future research directions.

2. Background

The liver accept/reject decision problem is a feature of the organ allocation policy employed in the United States. Accepting an offer provides the benefit of transplanting that particular liver, but forgoes the possibility of a higher-quality liver. Rejecting, on the other hand, allows for the possibility of obtaining a higher benefit through a higher-quality liver, but forgoes the benefit of transplanting the current liver, prolongs the waiting process, and exposes the patient to the risk of death and degraded quality of life while waiting for another offer. The focus of this paper is the evaluation of these trade-offs to make optimal accept/reject decisions for an individual risk-neutral patient.

2.1. Current Liver Allocation System

Deceased-donor liver transplantation in the United States is administered by UNOS, a nonprofit organization. The current liver allocation system partitions the country into 11 regions. These regions are further divided into smaller areas, called donation service areas (DSAs), each of which is coordinated by an organ procurement organization (OPO). Currently, there are 58 OPOs serving DSAs of varying sizes, demographic characteristics, and donation rates.

Figure 1 provides an overview of the deceased-donor liver allocation policy (see United Network for Organ Sharing 2010a for more details). This policy classifies the candidates into three categories in terms of geography: local (OPO) patients from the OPO that harvested the liver, regional patients from a different OPO within the harvesting region, and national patients outside the harvesting region. This policy further classifies the candidates into two categories of disease severity: status 1 patients (i.e., patients with a life expectancy of less than seven days without a transplant) and model for end-stage liver disease (MELD) patients. MELD patients are assigned an integer-valued score between 6 and 40,[1] where higher scores indicate more severe conditions. MELD score is a function of the patient's serum bilirubin, serum creatinine, and international normalized ratio (United Network for Organ Sharing 2010a). A patient's MELD score can go up or down over time.

[1] UNOS uses a slightly modified version of the original MELD formula introduced by Malinchoc et al. (2000) and its later refinements.

Figure 1 An Overview of Deceased-Donor Liver Allocation Policy

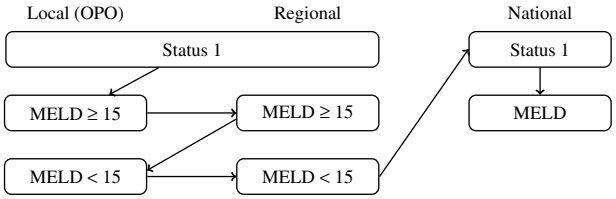


Table 1 Liver Waiting List Information for the Organ Procurement Organization Serving Pittsburgh

Disease severity	Number of registrations
MELD $\leq$ 10	182
11 $\leq$ MELD $\leq$ 18	188
19 $\leq$ MELD $\leq$ 24	51
MELD $\geq$ 25	46
All	467

Source. Based on OPTN (Organ Procurement and Transplantation Network) data as of January 1, 2010.

When a liver is harvested, it is first offered to status 1 patients in the harvesting region. If no patients accepts the organ, then the liver is offered to local MELD patients with a score of at least 15. If the liver is still not transplanted in this category, then the search continues as depicted by the arrows in Figure 1 until a patient accepts the offer or the liver exceeds its usable life (usually 12–18 hours) and is discarded. Within each MELD box in Figure 1, organs are offered to patients in order of decreasing MELD scores, where ties are broken first by blood type compatibility and then by the cumulative waiting time of patients at their current or higher MELD scores. In this paper, we take the allocation policy as given and assume that patients are offered organs using the given policy. Under this assumption, we investigate the impact of using the partially observable waiting list information on patients' lives when they are optimizing their accept/reject decisions for organs offered to them.

To illustrate the information published by UNOS, Table 1 shows a recent condensed[2] snapshot of the liver transplant waiting list for the OPO serving

[2] One can obtain more information from UNOS's website than the one displayed in Table 1. For instance, within each MELD category in Table 1, one can learn the number of waiting list registrations by blood type and waiting time among other categories. Intuition suggests that these additional pieces of information would potentially lead to better assessments of patient rank. The model we present in §3 is general enough to handle a wide range of such observations. However, in our computational study, we restrict the observations to be patient's health status only as measured by her MELD score. Obviously, such a restriction ignores some of the available information and hence would lead to overestimating the patient's price of privacy. However, we find, in §6, that using information based solely on MELD scores, even though restricted, dramatically reduces the price of privacy to negligible amounts. Therefore, the value of adding extra information would be negligibly small.

the greater Pittsburgh area. For clarity, we ignore the regional effects in the allocation policy and hence consider only the information provided in Table 1. Assuming a liver is harvested in this OPO during the time of this snapshot, it is only possible to determine a *range* of ranks for any patient. For instance, a patient with a MELD score of 30 can be no lower than 46th on the prioritized waiting list, whereas a patient with a MELD score of 18 could be ranked as high as 98th or as low as 285th.

In addition to the information provided over the Internet, patients may also obtain information about their waiting list rank by other means such as the information they receive through their surgeons. Our POMDP model is general enough to capture such additional sources of information. However, because such information is subjective and hard to obtain, for clarity we restrict our attention to two sources of information: the Internet and the quality of the offered liver (if any).

2.2. Organ Acceptance Literature

Problems in organ transplantation, including the accept/reject problem, have received considerable attention from the operations research community. One stream of research, similar to this paper, focuses on optimizing the individual decisions of each patient (e.g., Ahn and Hornberger 1996; Alagoz et al. 2007a, b; Sandıkçı et al. 2008), whereas another stream focuses on the societal aspects of alternative allocation policies either by simulation-based evaluation of given policies (e.g., Harper et al. 2000, Shechter et al. 2005, Zenios et al. 1999) or by designing an allocation policy that optimizes some measure of societal benefit (e.g., David and Yechiali 1995, Su and Zenios 2005, Zenios 1999, Zenios et al. 2000). A third stream investigates the system design problem while simultaneously considering the individual patient decisions (Su and Zenios 2004, 2006). Although many of the results in this body of literature are analytically interesting, they unfortunately rely on unrealistic assumptions such as static patient health, no new patient arrivals, at most one offer per organ, all organ offers are accepted, no patient deaths, and homogeneous patients. In this study, as in Alagoz et al. (2007b) and Sandıkçı et al. (2008), we relax such assumptions. Moreover, unlike any prior work in this area, we incorporate the partial observability of the waiting list in the accept/reject decisions.

The solution to an individual's accept/reject problem depends on several dynamic factors including (1) *patient health*, because the stage of a disease drives the treatment choices and different liver diseases are known to have different progressions; (2) the *quality of the offered liver*, because each potential recipient may benefit differently from each harvested liver; and (3) the *composition of the waiting list* (i.e., the characteristics of the specific patient under consideration as well as those of the other patients on the waiting list), because knowledge of this composition determines a patient's prospect for future offers. The first two factors were modeled in Alagoz et al. (2007b); however, they only implicitly capture the effects of the waiting list by defining the organ offer probabilities as a function of the patient's health. Sandıkçı et al. (2008) extend this work to explicitly include the third factor by assuming that the composition of the waiting list is perfectly known. They compare their findings with those of Alagoz et al. (2007b) to estimate the value of perfect waiting list information for an individual patient. They estimate this value, which they call the price of privacy, by computing the increased life expectancy associated with using perfect information. Their numerical results estimate this price (reported as percent increase in life expectancy) to be around 5%, on average, with some patients losing up to 15% of their life expectancy due to the concealed waiting list. In particular, they found that higher prices are typically paid by patients who (1) have less severe diseases, (2) are younger, and (3) join the waiting list in OPOs serving larger populations. In this paper, we are interested in estimating the losses associated with current UNOS practice, which reveals some but not all of the waiting list information.

2.3. POMDP Literature

POMDPs provide a natural framework to model decisions in a stochastic and dynamic environment where some features of the environment are not completely observable. Rather than directly observing these so-called core states, a decision maker receives a signal called an observation that (partially) informs the decision maker about the core states. The widely accepted modeling approach for such systems requires maintaining a belief state about the underlying core states of the system and updating this belief upon receiving observations. A belief state is a probability distribution over the set of core states, and is a sufficient statistic for the entire history of the process (Striebel 1965). As a result, one can retain the Markov property by writing the optimality equations as a function of the belief states. The downside, however, is that the set of belief states forms a continuum, namely, the $(n-1)$-dimensional probability simplex when the number of core states is n. Smallwood and Sondik (1973) prove a key result by showing that the optimal value function for a POMDP model is piecewise linear and convex. They also provide the first known exact algorithm for solving POMDPs, which has been expanded by several subsequent exact algorithms. Detailed reviews of POMDP methodology and available solution algorithms can be found elsewhere (e.g., Cassandra 1998, Lovejoy 1991b, Poupart 2005).

Despite the well-documented difficulties in solving and analyzing POMDPs, they continue to draw attention as a powerful modeling tool well suited for a wide variety of real-world problems. We refer the reader to Cassandra (1997) for a review of applications of POMDPs. POMDPs have also been used in the domain of healthcare, but only to a limited extent. Hu et al. (1996) use the POMDP framework for choosing an appropriate drug infusion plan in administering anesthesia; Hauskrecht and Fraser (2000) use it for ischemic heart disease treatment planning. Both of these analyses use heuristic methods to solve the resulting POMDPs. Maillart et al. (2008) formulate a partially observed Markov chain model to assess mammography screening policies and identify a set of efficient policies within a range of predefined policy set through sample-path enumeration of the Markov chain. Ayer et al. (2012) solve a POMDP model to optimize personalized mammography screening decisions based on individual risk of breast cancer. Zhang et al. (2012) use a similar framework to optimize biopsy referral decisions for prostate cancer patients.

2.4. Other Relevant Literature

Rational patients would consider the decisions of other patients on the waiting list while making their own accept/reject decisions. To calculate the price of privacy exactly, we would ideally compare the results of two game-theoretic models: one assuming perfect information and the other one assuming partial information. To our knowledge, Su and Zenios (2004) provide the only analysis that considers game-theoretic issues within the context of organ transplantation. They model the kidney waiting list as an $M/M/1$ queue and provide a detailed equilibrium analysis of different queuing disciplines by assuming a homogeneous patient population. Queuing-based approaches, in general, fall short of modeling the dynamics of the liver allocation system. Typically, most of the queuing models in the literature are concerned with kidney allocation (see Zenios 2004 for a review of this literature), which is modeled as a first-in, first-out queue due to the importance of waiting time in allocating deceased-donor kidneys. Moreover, the homogeneous population assumption in Su and Zenios (2004) is not justifiable in liver transplantation. Relaxing such assumptions, however, would lead to an intractable asymmetric stochastic game with thousands of players. Incorporating the partial observability would render the model even less tractable, given that POMDPs are significantly more difficult to analyze than their completely observed MDP counterparts. Therefore, we suppress the effects of competition among patients and, instead, focus on modeling the partial information available to patients to quantify the value of perfect waiting list information.

Finally, this paper also relates to research on the value of information. The value of information literature is vast; we refer the reader to the recent comprehensive survey by Yokota and Thompson (2004) for a review of applications in healthcare. In this context, this study is the first to explore the quality of imperfect waiting list information (as measured by the patient's opportunity loss in maximizing her life expectancy without perfect information) in liver transplantation.

3. The POMDP Model

In this section, we formulate a POMDP model to optimize a patient's accept/reject decisions. Generally speaking, decisions are made whenever an organ is offered as a function of (i) the patient's current health, (ii) the quality of the organ being considered, and (iii) the probability distribution of the patient's current position on the waiting list. Item (iii) is updated over time as the patient observes the quality of her organ offers, changes in her health status, and aggregated online waiting list information. The problem ends when the patient dies or accepts an organ.

More specifically, we formulate a discrete-time partially observed Markov decision process that uses the currently available imperfect waiting list information to optimize the accept/reject decisions of an ESLD patient. We assume that a risk-neutral patient is trying to maximize her total expected discounted reward. If there is no offer made at the time of a decision, the patient must wait until the next decision epoch. If, on the other hand, the patient receives a liver offer, then she chooses between rejecting or accepting the offer. A rejection provides the patient an immediate reward (e.g., the quality-adjusted time between decision epochs) and forces the patient to wait until the next epoch to face the same decision during which health changes, including death, are possible. Accepting the offer earns the patient a lump-sum terminal reward (e.g., the expected discounted posttransplant life days) and terminates the decision process.

Our POMDP model extends the completely observable MDP formulation of Sandıkçı et al. (2008). They define the state space of the process as $\mathcal{S} = \mathcal{S}' \cup \{\Delta\} \cup \{\nabla\}$, where Δ and ∇ are absorbing states that represent death and the state of being transplanted with a liver, respectively, and $\mathcal{S}' = \{(h, l, k) \mid h \in \Omega, l \in \Phi, k \in \Psi\}$, where h denotes the patient's health status, l denotes the quality of the offered liver, and k denotes a measure of the patient's priority for the given offer of type l. They define $\Omega = \{1, 2, \ldots, H\}$ with higher values corresponding to more deteriorated levels of health; $\Phi = \{1, 2, \ldots, L+1\}$ with higher values corresponding to lower-quality offers and $L+1$ indicating no liver offer; and $\Psi = \{1, 2, \ldots, K\}$ with higher values corresponding to lower patient priorities. As they note, modeling the priority of a patient for every type

of liver offer would require an enormous state space. As a compromise, they define k as the rank of the patient among all patients' expected priorities, where the expectation is taken over all possible offer types.

We take $\mathcal{S}$ to be the set of partially observable *core* states for our POMDP model. Furthermore, we assume that only the first two dimensions, h and l, of $\mathcal{S}$ are completely observable, whereas the third dimension, k, is not; but the patient can obtain probabilistic information about her true rank by observing the *signals* (i.e., the *ranges of ranks*) published on UNOS's website. Note that the *actual* range of ranks observed by a patient at any given time t, denoted ω^t, is subject to uncertainty because the evolution of the liver waiting list is a stochastic process. One possible way to model the observed range of ranks would be to define each possible realization of ω^t over all t as a signal. However, such an approach would clearly be intractable because of the formidable number of realizations of ω^t. We adopt an alternative modeling approach, which not only leads to a more tractable model but also provides a robust framework to deal with the uncertainty of ω^t.

Note that, under the current UNOS system, as illustrated in the discussion surrounding Table 1, ω^t depends on the patient's health at time t, denoted h^t, and the *ranges of health states*, denoted Ω_i for $i = 1, \ldots, r$, as predefined by UNOS, where r is the number of revealed health categories by UNOS. For notational convenience, let ω_i^t denote the range of rank states corresponding to Ω_i at time t. Hence, if $h^t \in \Omega_i$, then $\omega^t = \omega_i^t$. To illustrate, consider, for example, the time t for which the aggregated waiting list information is as depicted in Table 1. For this example, $r = 4$ and $\Omega_1 = \{25, \ldots, 40\}$, $\Omega_2 = \{19, \ldots, 24\}$, $\Omega_3 = \{11, \ldots, 18\}$, and $\Omega_4 = \{6, \ldots, 10\}$, which imply

$$\omega^t = \begin{cases} \omega_1^t = \{1, \ldots, 46\} & \text{if } h^t \in \Omega_1, \\ \omega_2^t = \{47, \ldots, 97\} & \text{if } h^t \in \Omega_2, \\ \omega_3^t = \{98, \ldots, 285\} & \text{if } h^t \in \Omega_3, \\ \omega_4^t = \{286, \ldots, 467\} & \text{if } h^t \in \Omega_4. \end{cases}$$

Also note that we assume r and Ω_i for all i are fixed a priori, as UNOS has not modified the Ω_i's noted above since inception. Despite fixing r and Ω_i for all i, the range given by ω_i^t for any i changes with t because of the stochastic nature of the liver waiting list. To avoid the obvious difficulties associated with this behavior, we take a *worst-case* approach in the sense that we introduce a time-invariant signal Y_i, which we define as the *widest* observable range of ranks associated with Ω_i. Given the stochastic nature of the waiting list, under this approach, the Y_i's will almost certainly overlap, but that does not cause difficulty in estimating and interpreting the associated parameters in our model. For notational convenience, we let the observed signal $y \in \Upsilon = \{Y_1, Y_2, \ldots, Y_r\}$ for given r and Y_i for all i. In §6.1, we discuss in detail how we estimate a signal Y_i. In summary, the patient actually occupies one of the core states (h, l, k), Δ, or ∇; she completely observes h, l, Δ, and ∇, but constructs a *belief* about k through signal y.

The set of possible actions in core state $s \in \mathcal{S}$ is

$$\mathcal{A}_s = \begin{cases} \{W\} & \text{if } s = \Delta \text{ or } s = \nabla \\ & \quad \text{or } s = \{(h, l, k) \mid l = L + 1\}, \\ \{T, W\} & \text{otherwise}, \end{cases}$$

where "W" indicates rejecting the offer and waiting for one more period; and "T" indicates accepting the offer and proceeding with the transplantation. The immediate expected reward for the state-action pair (s, a) is independent of the rank state and is given by

$$r(s, a) = \begin{cases} 0 & \text{if } s = \Delta \text{ or } s = \nabla, \\ r_W(h) & \text{if } s \in \mathcal{S}' \text{ and } a = W, \\ r_T(h, l) & \text{if } a = T. \end{cases}$$

The patient's decision process stops once she is dead or transplanted, therefore she does not accumulate any additional rewards in these states. We also assume that the reward of waiting, $r_W(h)$, only depends on the patient's current health, whereas the reward associated with transplanting, $r_T(h, l)$, depends not only on the patient's health, but also on the offered liver.

When the transplant action is taken, the process transitions to state ∇ with probability 1. To define the transition probabilities for the wait action, let $\mathcal{K} = [\mathcal{K}\{k' \mid h', k, h\}]_{\forall h, h' \in \Omega, k, k' \in \Psi}$ denote the *rank probability matrix*, where $\mathcal{K}\{k' \mid h', k, h\}$ is the probability that the patient's rank will be k' at time $t + 1$ given that her health and rank at time t are, respectively, h and k, and that her health at time $t + 1$ is h'; $\mathcal{H} = [\mathcal{H}\{h' \mid h\}]_{\forall h, h' \in \Omega}$ denote the *health probability matrix*, where $\mathcal{H}\{h' \mid h\}$ is the probability that the patient's health status will be h' at time $t + 1$ given that her health at time t is h; and $\mathcal{L} = [\mathcal{L}\{l' \mid k'\}]_{\forall k' \in \Psi, \forall l' \in \Phi}$ denote the *liver offer probability matrix*, where $\mathcal{L}\{l' \mid k'\}$ is the probability that the patient will be offered an organ of quality l' at time $t + 1$ given that her rank at time $t + 1$ is k'. Finally, given $h, h' \in \Omega$, let $\mathcal{K}^{h,h'} = [\mathcal{K}\{k' \mid h', k, h\}]_{\forall k, k' \in \Psi}$ be the square submatrix of $\mathcal{K}$. Note that $\sum_{k' \in \Psi} \mathcal{K}\{k' \mid h', k, h\} = \sum_{l' \in \Phi} \mathcal{L}\{l' \mid k'\} = 1$ for all $h, h' \in \Omega, k, k' \in \Psi$ and $1 - \sum_{h' \in \Omega} \mathcal{H}\{h' \mid h\}$ is the probability of death during the current period if the patient is in health state $h \in \Omega$. Given these parameters, we define the core state transition probabilities for the wait action as

$$P\{s' \mid s = (h, l, k)\} = \begin{cases} p(h', l', k' \mid h, l, k) & \text{if } s' = (h', l', k'), \\ 1 - \sum_{h' \in \Omega} \mathcal{H}\{h' \mid h\} & \text{if } s' = \Delta, \\ 0 & \text{if } s' = \nabla, \end{cases}$$

where $p(h', l', k' \mid h, l, k) = \mathcal{H}\{h' \mid h\} \cdot \mathcal{K}\{k' \mid h', k, h\} \cdot \mathcal{L}\{l' \mid k'\}$. The product form of the $\mathcal{H}$, $\mathcal{K}$, and $\mathcal{L}$ terms in this expression follows from simple probability arguments. One of the distinguishing features of this model from that of Sandıkçı et al. (2008) is that we provide a more accurate expression for $p(h', l', k' \mid h, l, k)$, whereas they used an approximation by replacing $\mathcal{K}\{k' \mid h', k, h\}$ with $\mathcal{K}\{k' \mid k\}$. To see the intuition behind why k' depends on h', k, and h, consider the following example. If $h = h'$, then one would infer that rank has likely not changed much, that is, the probability distribution over the rank states at time $t+1$ is roughly the same as that at time t. On the other hand, if h is low (relatively healthy) and h' is high (relatively sick), then one would infer that rank has likely improved, that is, the probability distribution over the rank states at time $t+1$ would be more concentrated on the better rank states than that at time t.

Finally, the linkage between observations and core states is modeled by the observation probabilities, denoted $\mathcal{O}\{y' \mid k'\}$ (i.e., the probability of observing $y' \in \Upsilon$ at time $t+1$ when the system actually occupies the core state (h', l', k') during this time after taking the wait action at time t). Given that the observation y' indicates a range of ranks, adding an extra dependency on h' would not change the observation probabilities as k' encapsulates all of the information available in h'. For a given rank $k' \in \Psi$ and signal $y' \in \Upsilon$, the worst-case approach that we have adopted to model the signals leads to the following observation probability

$$\mathcal{O}\{y' \mid k'\} = \begin{cases} 0 & \text{if } k' \notin y', \\ 1 & \text{if } k' \in y' \text{ and } k' \notin \Upsilon \setminus \{y'\}, \\ (0,1) & \text{otherwise.} \end{cases} \quad (1)$$

Because the elements of Υ may have overlapping rank states, a particular rank state k' may belong to multiple elements of Υ. When k' is a member of multiple elements, say, y' and y'', of Υ (i.e., the third case in Equation (1)), the observation probabilities $\mathcal{O}\{y' \mid k'\}$ and $\mathcal{O}\{y'' \mid k'\}$ are both nonzero and indicate the relative fraction of time k' belongs to y' and y'', respectively. We provide a numerical illustration of estimating the observation probabilities in §6.1.

For a given set $\mathbb{X}$ with n elements, a *belief space*, denoted $\Pi(\mathbb{X})$, is defined as the probability simplex over $\mathbb{X}$, that is, $\Pi(\mathbb{X}) = \{\boldsymbol{\pi} \in \mathbb{R}^n\colon \sum_{i=1}^n \pi_i = 1,\ \pi_i \geq 0\ \forall i\}$. Given that h and l are completely observed, we define a *belief state* $\boldsymbol{\pi}$ to be a probability distribution over Ψ, the set of rank states, i.e., $\boldsymbol{\pi} \in \Pi(\Psi)$.

Let h and $\boldsymbol{\pi}$ be the current health and belief states, respectively, and assume that the wait action is taken, after which signal $y' \in \Upsilon$ and states $h' \in \Omega$ and $l' \in \Phi$ are observed at the next decision epoch. The prior belief $\boldsymbol{\pi}$ is updated to $\boldsymbol{\pi}'$ using Bayes' rule, where π'_j, the jth component of the updated belief, denotes the probability that the new rank is j and is computed by

$$\pi'_j = \frac{\mathcal{O}\{y' \mid j\} \cdot \mathcal{L}\{l' \mid j\} \cdot \sum_k \mathcal{K}\{j \mid h', k, h\} \cdot \pi_k}{\sum_{k'} \mathcal{O}\{y' \mid k'\} \cdot \mathcal{L}\{l' \mid k'\} \cdot \sum_k \mathcal{K}\{k' \mid h', k, h\} \cdot \pi_k}. \quad (2)$$

Given discount rate $\lambda \in [0, 1)$, an optimal solution to this problem can be obtained by solving the optimality equations

$$\begin{aligned} v(h, L+1, \boldsymbol{\pi}) &= r_W(h) + \lambda \cdot \sum_k \sum_{(h', l', k')} \sum_{y'} \pi_k \cdot P\{h', l', k' \mid h, L+1, k\} \\ &\quad \cdot \mathcal{O}\{y' \mid k'\} \cdot v(h', l', \boldsymbol{\pi}'), \quad \forall h \in \Omega,\ \boldsymbol{\pi} \in \Pi(\Psi) \end{aligned} \quad (3)$$

and

$$\begin{aligned} v(h, l, \boldsymbol{\pi}) &= \max\Bigg\{ r_T(h, l);\ r_W(h) + \lambda \cdot \sum_k \sum_{(h', l', k')} \sum_{y'} \pi_k \\ &\quad \cdot P\{h', l', k' \mid h, l, k\} \cdot \mathcal{O}\{y' \mid k'\} \cdot v(h', l', \boldsymbol{\pi}') \Bigg\}, \\ &\quad \forall h \in \Omega,\ l \in \Phi \setminus \{L+1\},\ \boldsymbol{\pi} \in \Pi(\Psi). \end{aligned} \quad (4)$$

Note that the updated belief vector $\boldsymbol{\pi}'$ in Equations (3) and (4) is a function of the observations l' and y' as well as the current belief vector $\boldsymbol{\pi}$ (i.e., $\boldsymbol{\pi}' = \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y')$), but this dependency is suppressed for notational clarity. Also note that the value associated with the absorbing states of death and post-transplant, Δ and ∇, is zero by construction and these states are, therefore, excluded from Equations (3) and (4).

4. Structural Properties

POMDPs are well known for lacking computational tractability. In this section, we identify conditions that guarantee structure in the optimal solution of our model from §3. The structures that are of most interest are monotonicity of the value function and a threshold-type optimal policy. A threshold-type policy with respect to a state s is a policy for which there exists a state s^*, called the threshold, and one of the actions is prescribed for all states $s \leq s^*$ and the other action for all states $s > s^*$. Given that our model has a hybrid state space, composed of completely observable components h and l, and a partially observable component k, we defer properties for h and l dimensions to Appendix A and focus here on the most interesting dimension, k (rank). Sandıkçı et al. (2008) prove the existence of a threshold-type optimal policy with respect to a completely observable rank state for their perfect information model. However, neither the

conditions nor the proofs in those results apply to our partially observable model.

Let $\mathbb{X} = \{1, 2, \ldots, n\}$. The vector $\boldsymbol{\pi}^2 \in \Pi(\mathbb{X})$ is *greater than* $\boldsymbol{\pi}^1 \in \Pi(\mathbb{X})$ *in monotone likelihood ratio* (*MLR*) *order*, denoted $\boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2$, if $\pi^1_{i'}\pi^2_i \le \pi^1_i\pi^2_{i'}$ for $i \le i' \in \mathbb{X}$. A stochastic matrix $P = [p_{ij}]$ is *totally positive of order* 2 (TP_2) if $p_{ij}p_{i'j'} \ge p_{ij'}p_{i'j}$ for $i \le i'$, $j \le j'$. An equivalent characterization of a TP_2 matrix is that its successive rows are in increasing MLR order (Karlin 1968).

One of the unique features of our POMDP model is the availability of two separate signals, namely, the Internet signals y and the liver offers l, which is also part of the hybrid state space. Lovejoy (1987) provides a set of sufficient conditions that guarantee the optimal value function of a POMDP to be monotone with respect to MLR-ordered belief states. However, these conditions do not directly apply to our setting because of the mentioned distinction. Therefore, before proving the monotonicity of the optimal value function, we introduce the concept of more informative signals to deal with multiple signals. For this purpose, consider N separate observation streams. For $n = 1, \ldots, N$, let S_n be a set of possible observations from stream n signaling about a core state $\xi \in \Xi$ and $A_n = [a^n_{\xi i}]_{\substack{\xi\in\Xi \\ i\in S_n}}$ be the associated observation matrix. We say that, for $n_1 \ne n_2$, A_{n_1} is *more informative* than A_{n_2} if $\sigma^{A_{n_1}}(\xi, \xi', i_1, i_1') \ge \sigma^{A_{n_2}}(\xi, \xi', i_2, i_2')$ for all $\xi \le \xi'$ in Ξ, $i_1 \le i_1'$ in S_{n_1}, and $i_2 \le i_2'$ in S_{n_2}, where

$$\sigma^{A_n}(\xi, \xi', i, i') = \frac{a^n_{\xi i} \cdot a^n_{\xi' i'}}{a^n_{\xi i'} \cdot a^n_{\xi' i}} \quad \text{for } n = 1, \ldots, N. \tag{5}$$

To illustrate, consider two sets of signals $\{l', l''\} \subseteq \Phi$ and $\{y', y''\} \subseteq \Upsilon$ for our POMDP model such that $l' < l''$ meaning that the lower-quality liver l'' is a worse (i.e., more indicative of a lower rank) signal than l' and $y' < y''$ meaning that y'' is a worse signal than y'. If $\sigma^{\mathcal{O}}(k', k'', y', y'') \ge \sigma^{\mathcal{L}}(k', k'', l', l'')$ for all $k' \le k''$ in Ψ, then we say $\mathcal{O}$ is more informative than $\mathcal{L}$ for the given signal sets $\{l', l''\}$ and $\{y', y''\}$. Intuitively, this means that the relative likelihood of receiving, from the more informative source, the better signal (y') while the system is actually in a better state (k') and the worse signal (y'') while the system is actually in a worse state (k'') is at least as much as the analogous likelihood from the less informative source. If $\mathcal{O}$ is more informative than $\mathcal{L}$ for all possible signals in Υ and Φ, then we say $\mathcal{O}$ is more informative than $\mathcal{L}$.

The previous definition requires an ordering of the received signals. The liver qualities $1, 2, \ldots, L+1$ are naturally ordered, but the Internet signals $Y_1, Y_2, \ldots, Y_r$ are each composed of a (possibly overlapping) range of rank states. However, we observe from the UNOS website that the rank states in any given Y_i are in natural order and that $\min\{Y_i\} < \min\{Y_j\}$ for any $i < j$, because $\min\{h \in \Omega_i\} > \min\{h \in \Omega_j\}$ for any $i < j$. Therefore, we order the Internet signals as $Y_1 \le Y_2 \le \cdots \le Y_r$ and interpret Y_i to be a better signal than Y_j for $i < j$, because Y_i is indicative of higher ranks.

Theorem 1 provides sufficient conditions under which the optimal value function is monotone with respect to MLR-ordered belief states, and Theorem 2 uses the same conditions to demonstrate the existence of a threshold-type optimal policy. We provide the proofs in Appendix A and a numerical investigation of the sufficient conditions in Appendix B.

Theorem 1. *If $\mathcal{O}$, $\mathcal{L}$, and $\mathcal{K}^{h,h'}$ are TP_2 for all $h, h' \in \Omega$, and $\mathcal{O}$ is more informative than $\mathcal{L}$, then $v(h, l, \boldsymbol{\pi})$ is MLR-nonincreasing in $\boldsymbol{\pi} \in \Pi(\Psi)$ for any $h \in \Omega$ and $l \in \Phi$.*

Theorem 2. *If $\mathcal{O}$, $\mathcal{L}$, and $\mathcal{K}^{h,h'}$ are TP_2 for all $h, h' \in \Omega$, and $\mathcal{O}$ is more informative than $\mathcal{L}$, then there exists a threshold belief vector $\boldsymbol{\pi}^* \in \Pi(\Psi)$ such that the optimal action, for any $h \in \Omega$ and $l \in \Phi$, is*

$$a^*(h, l, \boldsymbol{\pi}) = \begin{cases} W & \text{if } \boldsymbol{\pi} \preceq_{LR} \boldsymbol{\pi}^*, \\ T & \text{if } \boldsymbol{\pi} \succ_{LR} \boldsymbol{\pi}^*. \end{cases}$$

5. Solution Methodology

Evaluating the entire belief space for a POMDP model with more than a handful of core states via exact algorithms is computationally intractable; hence, grid-based approximations are a natural and widely used technique to find approximate solutions. Lovejoy (1991a) uses a uniform grid comprised of a finite subset of points from the belief simplex that do not vary throughout the iterative solution procedure. The advantage of this uniform grid method is its efficient interpolation, whereas the disadvantage is its exponential growth in the size of the grid as the number of core states or the resolution of the grid increases. Hauskrecht (1997) uses a nonuniform grid method that starts with an arbitrary set of points (including the corner points) from the belief simplex and enhances the grid at each iteration by including more points according to various heuristics. The advantage of this method is its more economic use of grid points in approximating the value function and its ability to refine the grid at each iteration, whereas its disadvantage is its use of heuristic rules in selecting the initial and successive grid points as well as in interpolating the values of nongrid points. Zhou and Hansen (2001) describe a variable-resolution uniform grid method that combines the strengths of the previous two methods. Further details on grid-based methods can be found in Lovejoy (1991b) and Spaan and Vlassis (2005). In this section, we describe a fixed-resolution nonuniform grid method that exploits the special structure of our problem.

We first describe how we exploit the special structure of our problem to form the grid. Intuitively, we

would want to dedicate our limited computational resources to the evaluation of reachable belief points. In other words, if the probability of reaching a belief point is negligibly small, then we would not include such a point in the grid. For the model presented in §3, a belief that assigns positive probabilities to rank states k_1 and k_2 ($k_1 < k_2$) while assigning zero probability for all ranks j, where $k_1 < j < k_2$, is highly unlikely to occur; it is counterintuitive for a patient to think that she could be in ranks k_1 and k_2 but not in between.

More formally, we say a vector $\boldsymbol{\pi}$ is a *band vector* if the nonzero values in $\boldsymbol{\pi}$ appear consecutively. Similarly, we define a *band belief vector* as a belief distribution that forms a band vector. Our solution approach builds a grid by selecting a finite subset of band belief vectors from the belief space. Furthermore, as in any grid-based approach, we select belief vectors so that all of the nonzero values in a selected vector are positive integer multiples of $1/q$, where q is a positive integer representing the grid resolution. For instance, when the number of core states is $m = 3$ and $q = 3$, we form a grid comprised of the belief vectors

$$\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \begin{bmatrix}0\\1\\0\end{bmatrix},\quad \begin{bmatrix}0\\0\\1\end{bmatrix},\quad \begin{bmatrix}\frac{1}{3}\\\frac{1}{3}\\\frac{1}{3}\end{bmatrix},$$

$$\begin{bmatrix}\frac{1}{3}\\\frac{2}{3}\\0\end{bmatrix},\quad \begin{bmatrix}0\\\frac{1}{3}\\\frac{2}{3}\end{bmatrix},\quad \begin{bmatrix}\frac{2}{3}\\\frac{1}{3}\\0\end{bmatrix},\quad \begin{bmatrix}0\\\frac{2}{3}\\\frac{1}{3}\end{bmatrix}.$$

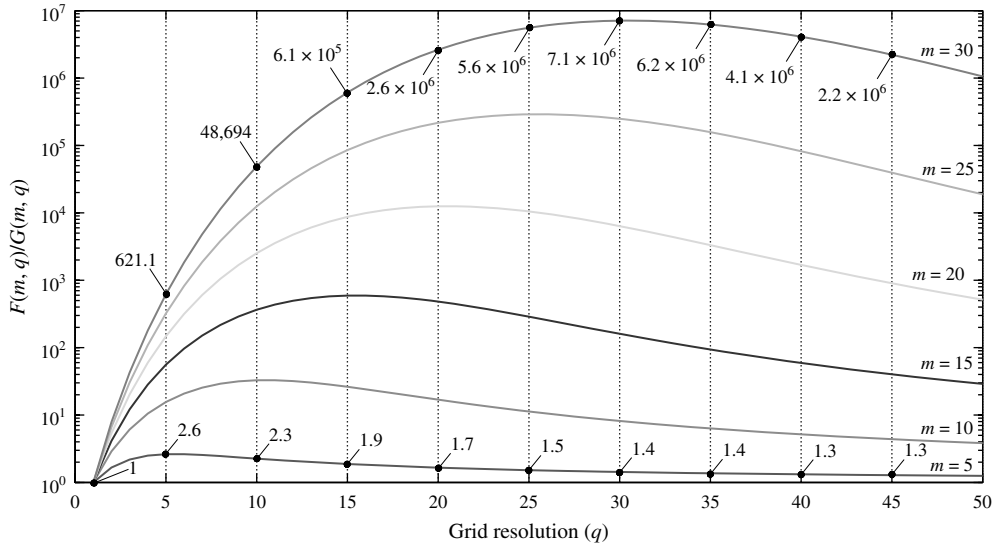
There are several advantages of building a grid in this manner. First and foremost, this approach, compared with a uniform grid, significantly reduces the size of the grid as m or q increases. The number of band belief points in a grid formed using this approach is given by $G(m, q) = \sum_{i=0}^{\min\{q-1,m\}} \binom{q-1}{i} \cdot (m-i)$, whereas the number of belief points using the uniform grid approach is given by $F(m, q) = \binom{m+q-1}{q}$. Figure 2 displays the ratio of the sizes of the grids formed by the uniform grid approach and by our approach. Clearly, when $q = 1$, both grids include only the degenerate distributions, hence this ratio is 1. For any given grid resolution q, as m increases, the size of the grid grows at a significantly slower rate using our approach. On the other hand, for any given number of core states m, the indicated ratio significantly increases with q until reaching its peak at $q = m$, but starts slowly decreasing from this peak point as q is increased further.

Moreover, we can establish conditions under which the Bayesian updating of a band belief vector preserves the band property. That is, the set of band belief vectors is closed under Bayesian updating, which implies that starting off the decision process with an initial band belief vector will retain the decision maker in the space of band belief vectors for all subsequent beliefs. Therefore, it suffices to store only band belief vectors in the grid. We numerically check the violation of these conditions in Appendix B.

THEOREM 3. *If $\mathcal{O}$ and $\mathcal{K}^{h,h'}$ are TP_2 for all $h, h' \in \Omega$, and $\mathcal{L}\{l \mid k\}$ is monotone in k, then $\boldsymbol{\pi}'(\boldsymbol{\pi}, \cdot, \cdot)$ is a band belief vector for any band vector $\boldsymbol{\pi} \in \Pi(\Psi)$.*

Next, we describe how we solve the optimality Equations (3) and (4) using a grid $\mathcal{G}$ formed as described above. (The solution approach remains valid with any arbitrarily specified $\mathcal{G}$ as long as the corner points of the belief simplex, i.e., the degenerate distributions, are included in $\mathcal{G}$.) For any given

Figure 2 **Ratio of the Number of Belief Vectors in a Uniform Grid and in Our Nonuniform Grid**



$h \in \Omega$, $l \in \Phi$, and $\boldsymbol{\pi} \in \Pi(\Psi)$, we approximate the optimal value function $v(h, l, \boldsymbol{\pi})$ by

$$\hat{v}(h, l, \boldsymbol{\pi}) = \sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}) \cdot \hat{v}(h, l, \hat{\boldsymbol{\pi}}), \tag{6}$$

where the interpolation weights $\beta_{\hat{\boldsymbol{\pi}}}(\cdot) \geq 0$ for all $\hat{\boldsymbol{\pi}} \in \mathcal{G}$, $\sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}) = 1$ for any $\boldsymbol{\pi} \in \Pi(\Psi)$, and the values for the belief points $\boldsymbol{\pi} \in \mathcal{G}$ are found by recursively solving the equations

$$\begin{aligned}\hat{v}(h, L+1, \boldsymbol{\pi}) &= r_W(h) + \lambda \sum_k \sum_{(h', l', k')} \sum_{y'} \pi_k P\{h', l', k' \mid h, L+1, k\} \\ &\quad \cdot \mathcal{O}\{y' \mid k'\} \sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') \hat{v}(h', l', \hat{\boldsymbol{\pi}})\end{aligned} \tag{7}$$

and, for $l \neq L+1$,

$$\begin{aligned}\hat{v}(h, l, \boldsymbol{\pi}) = \max\Bigg\{ & r_T(h, l); r_W(h) \\ & + \lambda \sum_k \sum_{(h', l', k')} \sum_{y'} \pi_k P\{h', l', k' \mid h, l, k\} \\ & \cdot \mathcal{O}\{y' \mid k'\} \sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') \hat{v}(h', l', \hat{\boldsymbol{\pi}}) \Bigg\}.\end{aligned} \tag{8}$$

Equations (7) and (8) can be solved using iterative algorithms such as value iteration, policy iteration, or variations thereof. In each iteration of such an algorithm, the updated belief point $\boldsymbol{\pi}'$ is again obtained by using Equation (2). Clearly, at any iteration, if $\boldsymbol{\pi}' \in \mathcal{G}$, then $\beta_{\boldsymbol{\pi}'}(\boldsymbol{\pi}') = 1$ and $\beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') = 0$ for all $\hat{\boldsymbol{\pi}} \in \mathcal{G} \backslash \{\boldsymbol{\pi}'\}$. If, however, $\boldsymbol{\pi}' \notin \mathcal{G}$, we find the interpolation weights that yield the closest approximation to the optimal value function $v(h', l', \boldsymbol{\pi}')$, for any $h' \in \Omega$ and $l' \in \Phi$, by solving the following linear program[3] (LP) with $\{\beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}')\}_{\hat{\boldsymbol{\pi}} \in \mathcal{G}}$ as the decision variables:

$$\begin{aligned}\min\Bigg\{ & \sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \hat{v}(h', l', \hat{\boldsymbol{\pi}}) \cdot \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') : \sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \hat{\boldsymbol{\pi}} \cdot \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') = \boldsymbol{\pi}', \\ & \sum_{\hat{\boldsymbol{\pi}} \in \mathcal{G}} \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') = 1,\ \beta_{\hat{\boldsymbol{\pi}}}(\boldsymbol{\pi}') \geq 0 \text{ for } \hat{\boldsymbol{\pi}} \in \mathcal{G} \Bigg\}.\end{aligned} \tag{9}$$

For each updated belief $\boldsymbol{\pi}'$, this LP must be re-solved at the beginning of each iteration of the solution algorithm, where the parameters $\{\hat{v}(h', l', \hat{\boldsymbol{\pi}})\}_{\hat{\boldsymbol{\pi}} \in \mathcal{G}}$ are taken from the value function estimates at the end of the previous iteration of the solution algorithm. By solving these LPs, one guarantees that, for any given grid $\mathcal{G}$, no other interpolation approach can provide a better approximation. Of course, suboptimal interpolation approaches may be faster in specific cases.

[3] Zhou and Hansen (2001) also formulate a similar LP, but they conclude that solving these LPs is computationally prohibitive.

Finally, also note that it is common to approximate the optimal value function $v(h, l, \boldsymbol{\pi})$, for any $h \in \Omega$, $l \in \Phi$, and $\boldsymbol{\pi} \in \Pi(\Psi)$, by entirely ignoring imperfect information. That is, we could approximate $v(h, l, \boldsymbol{\pi})$ by

$$\hat{u}(h, l, \boldsymbol{\pi}) = \sum_{k \in \Psi} \pi_k \cdot u_E(h, l, k), \tag{10}$$

where $u_E(\cdot)$ is the optimal value function of the completely observable counterpart of the POMDP model in question (namely, the explicit waiting list model (EWLM) of Sandıkçı et al. 2008). Comparing the approximations given in Equations (6) and (10), we can show that

$$v(h, l, \boldsymbol{\pi}) \leq \hat{v}(h, l, \boldsymbol{\pi}) \leq \hat{u}(h, l, \boldsymbol{\pi}) \quad \text{for } h \in \Omega,\ l \in \Phi,\ \boldsymbol{\pi} \in \Pi(\Psi). \tag{11}$$

6. Numerical Results

We compare results from the partially observable model presented in §3 with those from the completely observable model of Sandıkçı et al. (2008) so as to quantify the quality of the imperfect information (as measured by a patient's opportunity loss in her life expectancy due to using imperfect as opposed to perfect waiting list information) available on the UNOS's website.

6.1. Estimating the Observation Probabilities

For computational tractability in estimating the parameters of our models, we assume that the liver allocation system is in steady state, which can partially be justified by the fact that the size of the liver waiting list has been relatively stable since 2001 (United Network for Organ Sharing 2010b). Furthermore, the liver simulation model of Shechter et al. (2005), which we use only to simulate the dynamics of the liver waiting list to estimate the parameters $\mathcal{K}$, $\mathcal{L}$, and $\mathcal{O}$ for our models, implements the allocation policy prior to January 2005, in which the policy exhausts all the MELD patients in the local OPO prior to moving into the harvesting region followed by the nation.

To estimate an observation matrix $\mathcal{O}$, we first need to estimate Y_i for each Ω_i. As discussed in §3 and following the UNOS practice, we set $r = 4$ and $\Omega_1 = \{25, \ldots, 40\}$, $\Omega_2 = \{19, \ldots, 24\}$, $\Omega_3 = \{11, \ldots, 18\}$, and $\Omega_4 = \{6, \ldots, 10\}$. Because the relevant waiting list data are not available from UNOS, we estimate Y_i's using the national liver simulation model of Shechter et al. (2005) by tracking the daily range of ranks corresponding to each Ω_i throughout the simulation. To obtain the *widest* observable range of ranks Y_i for $i = 1, \ldots, r$, we define the lower limit of Y_i to be the minimum of the lower limits of the observed ω_i^t's (i.e., we set the lower limit of Y_i to $\min_t\{\min\{k: k \in \omega_i^t\}\}$) and the upper limit of Y_i to be the maximum of

the upper limits of the observed ω_i^t's (i.e., we set the upper limit of Y_i to $\max_t\{\max\{k\colon k \in \omega_i^t\}\}$).

To illustrate this worst-case approach, assume that there are 100 patients in the waiting list at some arbitrary time t and that the ω_i^t's are

$$\omega_1^t = \{1, \ldots, 10\}, \quad \omega_2^t = \{11, \ldots, 18\},$$
$$\omega_3^t = \{19, \ldots, 35\}, \quad \omega_4^t = \{36, \ldots, 100\},$$

and those at time $t+1$ are

$$\omega_1^{t+1} = \{1, \ldots, 9\}, \quad \omega_2^{t+1} = \{10, \ldots, 17\},$$
$$\omega_3^{t+1} = \{18, \ldots, 37\}, \quad \omega_4^{t+1} = \{38, \ldots, 99\}.$$

Further assume that $\omega_i^{t+2} = \omega_i^{t+1}$ for all i. We then define the Y_i's as

$$Y_1 = \{1, \ldots, 10\}, \quad Y_2 = \{10, \ldots, 18\},$$
$$Y_3 = \{18, \ldots, 37\}, \quad Y_4 = \{36, \ldots, 100\}.$$

As a result, a patient with a MELD score of 30, which falls within Ω_1, should expect her waiting list rank to be anything between 1 and 10.

Given Y_i's, we estimate the observation probability $\mathcal{O}\{Y_i \mid k\}$ associated with signal $Y_i \in \Upsilon$, for $i = 1, \ldots, r$, and rank $k \in \Psi$ by computing

$$\frac{\sum_{t=1}^{T} |k \cap \omega_i^t|}{\sum_{t=1}^{T} \sum_{j=1}^{r} |k \cap \omega_j^t|},$$

where T denotes the number of simulated days. Intuitively, the numerator of this fraction counts the number of days the patient in rank k has a MELD score in the range Ω_i (or, equivalently, the number of days the patient in rank k observes the signal Y_i). Similarly, the denominator of this fraction counts the number of days the patient in rank k has a MELD score in any one of the Ω_j's. For the three day example above, we would estimate the observation matrix $\mathcal{O}$ as in Figure 3 (note that this matrix has 100 rows). Figure 3 implies that a patient would observe Y_1 with probability 1 given that her rank is $k = 1, 2, \ldots, 9$. On the other hand, she is twice as likely to observe Y_2 than Y_1, if her rank is $k = 10$.

Figure 3 **Illustration of the Estimation of Observation Probabilities**

k	Y_1	Y_2	Y_3	Y_4
1–9	1	0	0	0
10	$\frac{1}{3}$	$\frac{2}{3}$	0	0
11–17	0	1	0	0
18	0	$\frac{1}{3}$	$\frac{2}{3}$	0
19–35	0	0	1	0
36–37	0	0	$\frac{2}{3}$	$\frac{1}{3}$
38–100	0	0	0	1

Figure 4 **Estimated Observation Matrix for a Typical Organ Procurement Organization**

Aggregated rank	Original rank	Y_1	Y_2	Y_3	Y_4
1	1	0.9997	0.0003	0	0
2	2	0.9920	0.0080	0	0
3	3–4	0.9396	0.0604	0	0
4	5–9	0.6264	0.3736	0	0
5	10–15	0.1202	0.8788	0.0010	0
6	16–22	0.0036	0.9675	0.0289	0
7	23–29	0	0.8430	0.1570	0
8	30–38	0	0.5866	0.4134	0
9	39–49	0	0.2257	0.7743	0
10	50–68	0	0.0200	0.9800	0
11	69–102	0	0	1.0000	0
12	103–391	0	0	0.5273	0.4727

Note. $Y_1 = \{1, \ldots, 22\}$; $Y_2 = \{1, \ldots, 67\}$; $Y_3 = \{12, \ldots, 321\}$; and $Y_4 = \{123, \ldots, 391\}$.

The waiting list data needed to estimate the Y_i's, and hence the $\mathcal{O}$'s, are not recorded by UNOS for public use. To create this data set, we use the simulation model of Shechter et al. (2005). We simulate the waiting list dynamics for a duration of four years and 30 independent replications. Using the results of these simulations, we estimate OPO-specific Y_i's and $\mathcal{O}$'s, because there is significant heterogeneity across OPOs (e.g., in terms of their size and the distribution of patient and donor characteristics). Furthermore, for computational tractability, we aggregate the core rank states using the methodology described in §4.1 of Sandıkçı et al. (2008). The resulting observation matrix $\mathcal{O}$ for a typical OPO appears in Figure 4.

6.2. Estimating the Quality of the Imperfect Waiting List Information

We estimate the quality of the imperfect information obtained via the partially observable waiting list for the same cohort of 200 ESLD patients studied in Sandıkçı et al. (2008) by comparing the approximations given in Equations (6) and (10). This sample of 200 patients consists entirely of MELD patients. Status 1 patients do not appear in this cohort mainly because they typically compose less than 0.1% of all patients nationwide at a given time.

For each patient in the cohort, we use the same parameter values as used in Sandıkçı et al. (2008): an annual discount rate of 0.97 that translates into $\lambda = 0.999917$ and $r_W(h) = 1$ day. Furthermore, we use the same methodologies used therein to estimate patient-specific posttransplant rewards (i.e., the Cox proportional hazards model of Roberts et al. 2004 to estimate $r_T(h, l)$), disease-specific health probability matrices (i.e., the natural history model of Alagoz et al. 2005 to estimate $\mathcal{H}$), and OPO-specific rank

transition and liver offer probability matrices (i.e., the national liver simulation model of Shechter et al. 2005 to estimate $\mathcal{H}$ and $\mathcal{L}$).

For computational tractability, we aggregate MELD scores into four health states (i.e., 6–10, 11–18, 19–24, 25–40), liver types into four qualities as determined by the donor's age (i.e., < 30, 30–44, 45–59, ≥ 60), which is an important factor in determining the quality of a donated liver (Roberts et al. 2004), and the ranks into 10 categories using the methodology described in §4.1 of Sandıkçı et al. (2008). This setup allows us to set the grid resolution parameter q to a maximum of 8, which cannot be increased further due to memory. Setting $q = 8$ leads to a grid $\mathcal{G}$ with 835 belief vectors, including an estimated initial belief state $\tilde{\boldsymbol{\pi}}$, in an 11-dimensional belief simplex. We solve each instance on a modern computational grid composed of 17 nodes. A typical node on this grid has 48 GB of RAM and 16 Intel processors running at 2.93 GHz. Portions of our code are paralellized to take advantage of the available processors. It takes, on average, about 10 minutes to solve each problem when $q = 8$.

Sandıkçı et al. (2008) first introduced the concept of a patient's price of privacy as the number of life days she would gain by acting optimally based on perfect waiting list information, as opposed to her optimal actions under the current allocation system. They used the implicit waiting list model (IWLM) of Alagoz et al. (2007b) as the best representation of the allocation system at the time. However, the IWLM does not incorporate partially observable waiting list information as available in the current system. For each patient in the cohort, they estimate this price by computing

$$\rho = \frac{u_E(\tilde{h}, L+1, K) - u_I(\tilde{h}, L+1, K)}{u_I(\tilde{h}, L+1, K)}, \qquad (12)$$

where $\tilde{h} \in \Omega$ is the patient's health at the time of listing and $u_I(\cdot)$ is the value of implementing the optimal policy of IWLM in the EWLM space. Rather than assuming a fixed rank $k = K$ for the patient's rank at the time of listing as in Equation (12), it is more realistic to estimate a belief $\tilde{\boldsymbol{\pi}} = [\tilde{\pi}_k]_{k\in\Psi}$ at the time of listing, representing a probability distribution across possible rank states given $\tilde{h}$. Hence, using the information embedded in $\tilde{\boldsymbol{\pi}}$, we provide an improved estimate as

$$\hat{\rho} = \sum_{k\in\Psi} \tilde{\pi}_k \cdot \left[\frac{u_E(\tilde{h}, L+1, k) - u_I(\tilde{h}, L+1, k)}{u_I(\tilde{h}, L+1, k)}\right]. \qquad (13)$$

We estimate the initial belief $\tilde{\boldsymbol{\pi}}$ for each patient in the cohort using the mentioned simulation model. The price of privacy as estimated by $\hat{\rho}$ in Table 2 ranges between 0.05% and 8.7%, which implies that the lack of waiting list information would cost some patients a significant fraction of their expected lives, although it may not impact others. Furthermore, we observe the same qualitative patterns (details not reported), as in Sandıkçı et al. (2008), among different patient categories with respect to age, disease severity, and geography. That is, we observe that younger patients, and those that have less severe diseases as well as those that join the waiting list from more densely populated geographies have a relatively higher price of privacy, and hence such patients can benefit more from accessing the waiting list information.

Table 2 **Perfect vs. No Information ($\hat{\rho}$) and Perfect vs. Partial Information ($\delta(\cdot)$)**

Statistic	$\hat{\rho}$	$\delta(1)$	$\delta(2)$	$\delta(3)$	$\delta(4)$	$\delta(5)$	$\delta(6)$	$\delta(7)$	$\delta(8)$
Min	0.05	—	—	—	—	—	—	—	—
Max	8.70	—	—	4×10^{-6}	—	—	4×10^{-6}	4×10^{-6}	—
Median	1.48	—	—	—	—	—	—	—	—
Mean	2.37	—	—	2×10^{-7}	—	—	5×10^{-7}	4×10^{-7}	—
St. dev.	2.11	—	—	6×10^{-7}	—	—	8×10^{-7}	8×10^{-7}	—

Notes. All of the values are reported as a percentage and "—" indicates 0%. Note that the values in the second column are different from those reported in Sandıkçı et al. (2008), because (i) we use $\hat{\rho}$ as opposed to ρ, (ii) we set $(H, L, K) = (4, 4, 10)$ as opposed to $(H, L, K) = (18, 14, 30)$, and (iii) we use $\mathcal{H}\{k' \mid h', k, h\}$ as opposed to $\mathcal{H}\{k' \mid k\}$.

In the next set of comparisons, we replace IWLM with our POMDP model to compute more accurate estimates of the price of privacy, because the POMDP model provides a better representation of the current system than IWLM. Specifically, we compute the new estimates, denoted $\delta(q)$ for grid resolution q, using

$$\delta(q) = \frac{\hat{u}(\tilde{h}, L+1, \tilde{\boldsymbol{\pi}}) - \hat{v}_q(\tilde{h}, L+1, \tilde{\boldsymbol{\pi}})}{\hat{v}_q(\tilde{h}, L+1, \tilde{\boldsymbol{\pi}})}, \qquad (14)$$

where $\hat{v}_q(\cdot)$ is the approximate value function for the POMDP model obtained by solving Equation (6) with grid resolution q. Intuitively, $\delta(q)$ measures the fraction by which a patient's life expectancy increases by acting optimally using perfect information as opposed to acting optimally under the current (partially observable) system.

The $\delta(q)$ column for $q = 1$ (coarser grid) to $q = 8$ (finer grid) in Table 2 summarizes the new estimates for the same cohort of patients. The key observation from these results is that one can eliminate the inefficiencies (as measured by a patient's price of privacy) caused by a concealed waiting list through the partially observable waiting list model. The fact that $\delta(1) = 0\%$ for all 200 patients is expected, because when $q = 1$, the grid is only composed of the corner points of the belief simplex, which is equivalent to assuming complete observability, hence $\hat{u}(h, l, \boldsymbol{\pi}) = \hat{v}_q(h, l, \boldsymbol{\pi})$ for any $\boldsymbol{\pi} \in \Pi(\Psi)$. As we

increase the grid resolution q, however, it is striking to observe that the price of privacy estimates remain essentially zero. Comparing these results to the mean $\hat{\rho}$ (2.37%), we conclude that a patient can benefit significantly from the published waiting list information.

6.3. A Detailed Comparison of the Policies

The results in §6.2 compare the optimal value functions of the three models (namely, IWLM, EWLM, and the POMDP model of this paper). To compare the optimal accept/reject policies from the three models, we simulate the underlying stochastic process for a given patient using the transition probability matrix and the initial belief state, both of which are estimated for the patient in question. We make 100,000 independent replications of the simulation. During each replication, as the stochastic process evolves over time, we superimpose the optimal accept/reject decisions prescribed for the patient by a given model. Hence, we repeat each replication three times; the first time superimposing the prescription from the partially observable model, the second and third times superimposing the prescriptions from EWLM and IWLM, respectively. We use the common random numbers technique (Law 2007) to facilitate comparisons across the three decision rules.

We repeat this simulation for each of the 200 patients in our cohort. Table 3 displays the average, across the 200 instances, of the expected value of several performance metrics estimated from simulating the three models when $(H, L, K) = (4, 4, 10)$. It is striking that all of the metrics under IWLM are statistically significantly different than those under EWLM, whereas there is no distinguishable difference between EWLM and the partially observable model. In particular, if patients use the optimal IWLM policy, they wait about 81.5 days (or 18%) longer, on average, to accept an offer or die compared with using the optimal EWLM policy. This extra waiting period allows the patients to receive, on average, 11% more offers, but they reject a higher portion of the offers made to them. As a consequence, the probability that a patient dies before receiving a transplant is significantly higher if she follows the policy from IWLM compared with using perfect information. Even worse is the conditional probability of death without a transplant, conditioned on receiving at least one offer. On the other hand, patients transplant marginally better organs from younger donors if they follow IWLM rather than EWLM, but deteriorate to a slightly sicker MELD score at the time of a transplant. All of these results further support our main finding from §6.2: a lack of waiting list information can have severe negative consequences for a patient, but these negative events can be remedied by using the aggregated waiting list information.

Table 3 **Comparing the Accept/Reject Decisions Suggested by the Three Models**

	IWLM	EWLM	POMDP
Number of days on the waiting list	370.383*	288.870	289.115
Number of offers received	1.761*	1.699	1.701
Fraction of offers rejected	0.323*	0.253	0.252
Probability of death before a transplant	0.192*	0.157	0.157
Probability of death before a transplant given received at least one offer	0.374*	0.295	0.297
Quality of transplanted liver (age of donor)	38.708*	39.842	39.828
Health (MELD score) at transplant	22.335*	21.380	21.385

*Statistically significant difference, at 1% level, than the corresponding value from EWLM.

6.4. Sensitivity Analysis

To test the robustness of the results discussed so far, we repeat the optimization study of §6.2 and the simulation study of §6.3 with various parameter settings. For this purpose, we include three levels for the number of health states (i.e., $H = 4, 9, 18$); two levels for liver quality states (i.e., $L = 4, 14$); and three levels for the number of rank states (i.e., $K = 10, 12, 30$).[4] Table 4 presents summary statistics for $\hat{\rho}$ across the 200 patients for each (H, L, K) combination. These results confirm that some patients may be severely affected if they ignore the waiting list information, although some may not be affected by much. On average, a patient can forfeit a significant fraction of her life expectancy by ignoring the waiting list information. Comparing, on the other hand, the partially observable model to the perfect information model for each (H, L, K) combination, we again find $\delta(q) \approx 0$ for $q = 1, \ldots, 5$. Furthermore, when we repeat the simulation study of §6.3 for each (H, L, K) combination, we find qualitatively identical patterns. Therefore, we conclude that ignoring the waiting list information can have significant negative consequences for patients, which can be remedied by using the published (imperfect) information.

Finally, the values in Table 4 indicate that the mean value of $\hat{\rho}$ increases (at 1% significance level) with K, which is expected because disaggregating rank states essentially means using more detailed information about the waiting list. On the other hand, the mean value of $\hat{\rho}$ decreases with H (similarly, with L), which makes intuitive sense because the incremental benefit of revealing the perfect rank information to someone who already has a finer health state description

[4] For $H = 4$, we aggregate MELD scores as ≤ 10, 11–18, 19–24, ≥ 25; for $H = 9$, we aggregate MELD scores as ≤ 9, 10–13, 14–17, 18–21, 22–25, 26–29, 30–33, 34–37, ≥ 38; and for $H = 18$, we aggregate MELD scores into groups of two. For $L = 4$, we classify livers using donor's age as described in §6.2; for $L = 14$, we use the classification given in Alagoz et al. (2007b). For a given value of K, we aggregate rank states as described in §4.1 of Sandıkçı et al. (2008).

Table 4 Effect of State Aggregation on $\hat{\rho}$ (Values Reported as %)

		Min $\hat{\rho}$			Mean $\hat{\rho}$			Max $\hat{\rho}$		
H	L	$K=10$	$K=12$	$K=30$	$K=10$	$K=12$	$K=30$	$K=10$	$K=12$	$K=30$
4	4	0.05	0.04	0.04	2.37*	2.49*	2.61*	8.70	10.05	12.17
9	4	0.10	0.10	0.10	2.02*	2.23*	2.31*	8.24	8.58	10.59
18	4	0.07	0.07	0.07	1.75*	2.02*	2.12*	5.98	9.20	11.67
4	14	0.07	0.06	0.06	1.75*	1.91*	1.98*	6.54	6.47	7.00
9	14	0.10	0.09	0.10	1.68*	1.91*	1.98*	7.96	7.94	7.95
18	14	0.10	0.10	0.10	1.56*	1.85*	1.95*,**	8.13	8.11	9.90

*Denotes statistical significance at 1% confidence level.
**This value is different than 5% reported in Sandıkçı et al. (2008), because we use (i) $\hat{\rho}$ as opposed to ρ, and (ii) $\mathcal{R}\{k' \mid h', k, h\}$ as opposed to $\mathcal{R}\{k' \mid k\}$.

(large H) would not be as high as revealing the same information to someone who does not have much precision about her health states (small H).

7. Concluding Remarks

We develop a POMDP model with a hybrid state space, where one of the completely observable components (l) also acts as another signal to inform about the partially observable component (k), to optimize the accept/reject decisions of an ESLD patient. We provide structural analyses of this partially observable model, and derive results including conditions that guarantee the existence of intuitive threshold-type optimal policies with respect to the partially observed rank (k). Furthermore, we solve this model for 200 randomly sampled patients and compare the solutions with those from the models of Sandıkçı et al. (2008) and Alagoz et al. (2007b) for each patient to assess the impact of the imperfect waiting list information published by UNOS.

We find that the optimal accept/reject decisions of ESLD patients can change significantly depending on what they know about their waiting list positions. Our comparisons demonstrate that a patient can do significantly better by taking advantage of the (imperfect) information available to her through UNOS's website. In particular, using the partially observed waiting list model reduces the price of privacy, which averages around 2%, but can be as high as 12%, to nearly 0%. Furthermore, on average, it reduces a patient's rejection rate and the probability of death without a transplant, while also reducing the number of offers they would receive as patients wait (about three months) shorter to accept an offer.

This study comes with some limitations. First, despite the fact that we can guarantee that the grid-based solution we offer is the best solution that can be obtained for a given set of grid points, we cannot assess the quality of this solution compared with the true optimum. Nevertheless, the fact that our estimates of the price of privacy ($\delta(q)$) are all negligibly small for various grid resolutions (q) leads us to believe that our solutions are not far from the optimum. Second, our optimization model assumes a risk-neutral decision maker with the objective of maximizing her total expected discounted life days. Incorporating the risk preferences of patients and/or considering other objectives is left for future research. Third, our numerical study does not consider status 1 patients. However, it is unlikely that waiting list information will affect any status 1 patient much. Such patients are almost surely no lower than third in their regional list (recall that there are typically a dozen such patients nationwide), and they should know this without having any access to the waiting list. Finally, assuming strategic decision makers and considering the accept/reject problem in such settings would be an interesting future research direction. One possible way to approach this problem would be to embed the optimal accept/reject decisions, hence the mentioned optimization models, into a simulation model that simulates the transplant waiting list.

Acknowledgments

The authors thank department editor Assaf Zeevi, an anonymous associate editor, and three anonymous referees for providing constructive comments that helped improve this paper. Burhaneddin Sandıkçı thanks Donald D. Eisenstein for his extensive support during the preparation of this paper and Laurens G. Debo for his comments on an earlier version of this paper. The data and analyses reported in the 2008 Annual Report of the U.S. Organ Procurement and Transplantation Network and the Scientific Registry of Transplant Recipients have been supplied by UNOS and Arbor Research under contract with HHS/HRSA. The authors alone are responsible for reporting and interpreting these data; the views expressed herein are those of the authors and not necessarily those of the U.S. Government. This research was supported by the National Science Foundation [Grant CMMI-0546960]. Burhaneddin Sandıkçı also acknowledges the financial support from the University of Chicago Booth School of Business.

Appendix A. Analytical Results and Proofs

Considering the completely observable l component of the state space, we establish, in Proposition 1, the intuitive fact that it is always better to be offered a higher-quality liver. Similar structural results for the health component, h, can be derived as in Sandıkçı et al. (2008).

Proposition 1. *If $r_T(h,l)$ is nonincreasing in l, then*

(a) *$v(h,l,\boldsymbol{\pi})$ is nonincreasing in l for any $h \in \Omega$ and $\boldsymbol{\pi} \in \Pi(\Psi)$;*

(b) *there exists a threshold liver $l^* \in \Phi$ such that, for any $h \in \Omega$ and $\boldsymbol{\pi} \in \Pi(\Psi)$, the optimal action is*

$$a^*(h,l,\boldsymbol{\pi}) = \begin{cases} T & \text{if } l \le l^*, \\ W & \text{if } l > l^*. \end{cases}$$

Proof of Proposition 1. The proof follows from $r_T(h,l)$ being nonincreasing in l and Equations (3) and (4). □

Considering the partially observable rank component k and the associated belief $\boldsymbol{\pi}$, we first establish, in Proposition 2, sufficient conditions under which Bayesian updating yields MLR-ordered beliefs. These results also help prove the main result in Theorem 1.

Proposition 2. (a) *If $\mathcal{L}\{l \mid k\}$ is nonincreasing in $k \in \Psi$ for all $l \neq L+1$, then $\boldsymbol{\pi}'(\cdot, l', \cdot) \preceq_{LR} \boldsymbol{\pi}'(\cdot, L+1, \cdot)$ for $l' \in \Phi$.*

(b) *If $\mathcal{L}$ is TP_2, then $\boldsymbol{\pi}'(\cdot, l', \cdot) \preceq_{LR} \boldsymbol{\pi}'(\cdot, l'', \cdot)$ for $l' \le l''$ in Φ.*

(c) *If $\mathcal{O}$ is TP_2, then $\boldsymbol{\pi}'(\cdot, \cdot, y') \preceq_{LR} \boldsymbol{\pi}'(\cdot, \cdot, y'')$ for $y' \le y''$ in Υ.*

(d) *If $\mathcal{O}$ is more informative than $\mathcal{L}$, then $\boldsymbol{\pi}'(\cdot, l'', y') \preceq_{LR} \boldsymbol{\pi}'(\cdot, l', y'')$ for $l' \le l''$ in Φ and $y' \le y''$ in Υ.*

(e) *If $\mathcal{H}^{h,h'}$ is TP_2 for all $h, h' \in \Omega$, then $\boldsymbol{\pi}'(\boldsymbol{\pi}^1, \cdot, \cdot) \preceq_{LR} \boldsymbol{\pi}'(\boldsymbol{\pi}^2, \cdot, \cdot)$ for $\boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2$ in $\Pi(\Psi)$.*

Proof of Proposition 2. (a) For any given $h, h' \in \Omega$, $y' \in \Upsilon$, $\boldsymbol{\pi} \in \Pi(\Psi)$, and $l' \in \Phi$,

$$\begin{aligned}
&\boldsymbol{\pi}'(\cdot, l', \cdot) \preceq_{LR} \boldsymbol{\pi}'(\cdot, L+1, \cdot) \\
&\quad \Leftrightarrow \quad \pi'_j(\cdot, L+1, \cdot) \cdot \pi'_i(\cdot, l', \cdot) \le \pi'_i(\cdot, L+1, \cdot) \cdot \pi'_j(\cdot, l', \cdot) \\
&\qquad\qquad \text{for } i \ge j \in \Psi \\
&\quad \Leftrightarrow \quad \frac{\mathcal{O}\{y' \mid j\}\mathcal{L}\{L+1 \mid j\}\sum_k \mathcal{H}\{j \mid h', k, h\}\pi_k}{\sum_{k'} \mathcal{O}\{y' \mid k'\}\mathcal{L}\{L+1 \mid k'\}\sum_k \mathcal{H}\{k' \mid h', k, h\}\pi_k} \\
&\qquad\qquad \cdot \frac{\mathcal{O}\{y' \mid i\}\mathcal{L}\{l' \mid i\}\sum_k \mathcal{H}\{i \mid h', k, h\}\pi_k}{\sum_{k'} \mathcal{O}\{y' \mid k'\}\mathcal{L}\{l' \mid k'\}\sum_k \mathcal{H}\{k' \mid h', k, h\}\pi_k} \\
&\qquad\qquad \le \frac{\mathcal{O}\{y' \mid i\}\mathcal{L}\{L+1 \mid i\}\sum_k \mathcal{H}\{i \mid h', k, h\}\pi_k}{\sum_{k'} \mathcal{O}\{y' \mid k'\}\mathcal{L}\{L+1 \mid k'\}\sum_k \mathcal{H}\{k' \mid h', k, h\}\pi_k} \\
&\qquad\qquad \cdot \frac{\mathcal{O}\{y' \mid j\}\mathcal{L}\{l' \mid j\}\sum_k \mathcal{H}\{j \mid h', k, h\}\pi_k}{\sum_{k'} \mathcal{O}\{y' \mid k'\}\mathcal{L}\{l' \mid k'\}\sum_k \mathcal{H}\{k' \mid h', k, h\}\pi_k} \\
&\quad \Leftrightarrow \quad \mathcal{L}\{L+1 \mid j\}\mathcal{L}\{l' \mid i\} \le \mathcal{L}\{L+1 \mid i\} \cdot \mathcal{L}\{l' \mid j\} \\
&\qquad\qquad \text{for } i \ge j \in \Psi,
\end{aligned}$$

which is always true because

$$\frac{\mathcal{L}\{L+1 \mid j\}}{\mathcal{L}\{L+1 \mid i\}} \le 1 \le \frac{\mathcal{L}\{l' \mid j\}}{\mathcal{L}\{l' \mid i\}} \quad \text{for } i \ge j \in \Psi.$$

(b)–(d) These proofs are similar to the proof of Proposition 2(a).

(e) This proof follows from Lemma 4.1 in Sandıkçı (2008), and from Lemmas 1.2.2 and 1.3.1 in Lovejoy (1987). □

Proposition 2 identifies conditions that lead to better belief states. As we later prove in Theorem 1, the value function $v(h,l,\boldsymbol{\pi})$ is nonincreasing in MLR-ordered belief states. With this result in mind, we say that $\boldsymbol{\pi}^1 \in \Pi(\Psi)$ is a better belief state than $\boldsymbol{\pi}^2 \in \Pi(\Psi)$ if $\boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2$. Proposition 2(a) provides conditions for which receiving an offer of any kind is better than not receiving any offers. The monotonicity condition in Proposition 2(a) is obviously more natural and easier satisfied than the TP_2 condition in Proposition 2(b). However, if $\mathcal{L}$ is TP_2, then it is better to receive a higher-quality offer l' than receiving an inferior offer l'' or not receiving any offers. Similarly, if $\mathcal{O}$ is TP_2, then it is better to receive a better Internet signal y'. Interpreting Proposition 2(d) requires imposing a lexicographic ordering for the elements of the set $\Phi \times \Upsilon$, that is, for $l' \le l''$ in Φ and $y' \le y''$ in Υ, we say $(l', y') \le (l'', y') \le (l', y'') \le (l'', y'')$. Hence, if one source ($\mathcal{O}$) is more informative than another source ($\mathcal{L}$), then a better signal (y') from the more informative source mixed with a worse signal (l'') from the other source is preferred to a worse signal (y'') from the more informative source mixed with a better signal (l') from the other source. Finally, Proposition 2(e) shows the conditions under which the belief states preserve the MLR order upon conditioning on new information.

Proof of Theorem 1. We first need three technical results before proving the main result.

Lemma 1 (Lemma 3.1 in Ohnishi et al. 1994). *If P and Q are TP_2 matrices of compatible sizes, then $P \cdot Q$ is also TP_2.*

Given $\mathbb{X} = \{1, 2, \ldots, n\}$ and belief vectors $\boldsymbol{\pi}^1$ and $\boldsymbol{\pi}^2$ in $\Pi(\mathbb{X})$, $\boldsymbol{\pi}^2$ is *stochastically greater* than $\boldsymbol{\pi}^1$, denoted $\boldsymbol{\pi}^1 \preceq_S \boldsymbol{\pi}^2$, if $\sum_{i \ge k} \pi^1_i \le \sum_{i \ge k} \pi^2_i$ for $k \in \mathbb{X}$. It is well known that $\boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2$ implies $\boldsymbol{\pi}^1 \preceq_S \boldsymbol{\pi}^2$ (Shaked and Shantikumar 2007).

Lemma 2 (Lemma 4.7.2 in Puterman 1994). *For $\mathbb{X} = \{1, 2, \ldots, n\}$ and $\boldsymbol{\pi}^1, \boldsymbol{\pi}^2 \in \Pi(\mathbb{X})$, $\boldsymbol{\pi}^1 \preceq_S \boldsymbol{\pi}^2$ implies $\sum_{i \in \mathbb{X}} \pi^1_i f_i \ge \sum_{i \in \mathbb{X}} \pi^2_i f_i$ for every nonincreasing sequence $\{f_i\}_{i \in \mathbb{X}}$.*

The *Kronecker product* of an $m_1 \times n_1$ matrix $A = [a_{ij}]$ and a $m_2 \times n_2$ matrix $B = [b_{kl}]$ is the $m_1 m_2 \times n_1 n_2$ matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

When $m_1 = m_2$, the *Khatri-Rao product* (or, *the row-wise Kronecker product*) of A and B is the $m_1 \times n_1 n_2$ matrix

$$A \odot B = \begin{bmatrix} \mathbf{a}_1. \otimes \mathbf{b}_1. \\ \mathbf{a}_2. \otimes \mathbf{b}_2. \\ \vdots \\ \mathbf{a}_m. \otimes \mathbf{b}_m. \end{bmatrix},$$

where $\mathbf{a}_i.$ and $\mathbf{b}_i.$ are the ith row of A and B, respectively.

Lemma 3. *If A and B are TP_2 matrices with equal number of rows and A is more informative than B, then $A \odot B$ is TP_2.*

Proof of Lemma 3. Let $A = [a_{ij}]_{m \times n}$ and $B = [b_{ik}]_{m \times q}$. By definition, for $i = 1, \ldots, m$, the ith row of $C = A \odot B$, denoted $\mathbf{c}_i.$, is $\mathbf{c}_i. = \mathbf{a}_i. \otimes \mathbf{b}_i.$ and each entry in $\mathbf{c}_i.$, denoted $c_{i,\{j,k\}}$ for $j = 1, \ldots, n$ and $k = 1, \ldots, q$, is $c_{i,\{j,k\}} = a_{ij}b_{ik}$. Given $i < i'$, $j < j'$ and $k < k'$, we consider the following exhaustive list of cases.

Case i. For $t = j, j'$, $c_{i,\{t,k\}} \cdot c_{i',\{t,k'\}} - c_{i,\{t,k'\}} \cdot c_{i',\{t,k\}} = a_{it}b_{ik} \cdot a_{i't}b_{i'k'} - a_{it}b_{ik'} \cdot a_{i't}b_{i'k} = a_{it}a_{i't} \cdot [b_{ik}b_{i'k'} - b_{ik'}b_{i'k}] \ge 0$, which follows because B is TP_2 and all coefficients are nonnegative.

Case ii. For $t = k, k'$, $c_{i,\{j,t\}} \cdot c_{i',\{j',t\}} - c_{i,\{j',t\}} \cdot c_{i',\{j,t\}} = a_{ij}b_{it} \cdot a_{i'j'}b_{i't} - a_{ij'}b_{it} \cdot a_{i'j}b_{i't} = b_{it}b_{i't} \cdot [a_{ij}a_{i'j'} - a_{ij'}a_{i'j}] \geq 0$, which follows because A is TP_2 and all coefficients are nonnegative.

Case iii. Since $a_{ij}a_{i'j'} \geq a_{ij'}a_{i'j}$ (because A is TP_2), $b_{ik}b_{i'k'} \geq b_{ik'}b_{i'k}$ (because B is TP_2), and all coefficients are nonnegative, $c_{i,\{j,k\}} \cdot c_{i',\{j',k'\}} - c_{i,\{j',k'\}} \cdot c_{i',\{j,k\}} = a_{ij}b_{ik} \cdot a_{i'j'}b_{i'k'} - a_{ij'}b_{ik'} \cdot a_{i'j}b_{i'k} \geq 0$.

Case iv. Since A is more informative than B, $c_{i,\{j,k'\}} \cdot c_{i',\{j',k\}} - c_{i,\{j',k\}} \cdot c_{i',\{j,k'\}} = a_{ij}b_{ik'} \cdot a_{i'j'}b_{i'k} - a_{ij'}b_{ik} \cdot a_{i'j}b_{i'k'} \geq 0$.

Now that all possible cases have nonnegative determinants, the matrix C is TP_2. □

Similar to the proof of Proposition 1 in Lovejoy (1987), we prove Theorem 1 by induction on the steps of the value iteration algorithm. Let $v^i(h, l, \boldsymbol{\pi})$ be the value of state $(h, l, \boldsymbol{\pi})$ at the ith iteration of the algorithm and assume, without loss of generality, that $v^0(h, l, \boldsymbol{\pi}) = 0$ for all h, l, and $\boldsymbol{\pi}$. It is clear that $v^1(h, l, \boldsymbol{\pi})$ is constant and therefore nonincreasing in $\boldsymbol{\pi} \in \Pi(\Psi)$ for any $h \in \Omega$ and $l \in \Phi$.

For the induction hypothesis, given $h \in \Omega$ and $l \in \Phi$, assume $v^i(h, l, \boldsymbol{\pi})$ is MLR-nonincreasing in $\boldsymbol{\pi}$ for iterations $i = 2, \ldots, n$. If $a^{n+1}(h, l, \boldsymbol{\pi}^2) = T$, then, by Equation (4), $v^{n+1}(h, l, \boldsymbol{\pi}^2) = r_T(h, l) \leq v^{n+1}(h, l, \boldsymbol{\pi}^1)$ for $\boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2$.

To prove the intended result when $a^{n+1}(h, l, \boldsymbol{\pi}^2) = W$, let $\gamma(l, y \mid k) \equiv \mathcal{O}\{y \mid k\} \cdot \mathcal{L}\{l \mid k\}$ and $\boldsymbol{\Gamma}$ be the matrix with the kth row given by $\gamma(\cdot \mid k)$ so that $\boldsymbol{\Gamma} \equiv \mathcal{O} \odot \mathcal{L}$. Because $\mathcal{L}$ and $\mathcal{O}$ are TP_2 and $\mathcal{O}$ is more informative than $\mathcal{L}$, Lemma 3 implies that $\boldsymbol{\Gamma}$ is TP_2. Premultiplying $\boldsymbol{\Gamma}$ with the TP_2 matrix $\mathcal{H}^{h,h'}$ and using Lemma 1, we find another TP_2 matrix $\boldsymbol{\Theta}^{h,h'} \equiv \mathcal{H}^{h,h'} \cdot \boldsymbol{\Gamma}$. Note that the $(k, (l', y'))$th entry of the matrix $\boldsymbol{\Theta}^{h,h'}$ is $\sum_{k'} \mathcal{H}\{k' \mid h', k, h\}\mathcal{L}\{l' \mid k'\}\mathcal{O}\{y' \mid k'\}$. Now, consider a matrix $\boldsymbol{\alpha}$ formed by the MLR-ordered beliefs $\boldsymbol{\pi}^1$ and $\boldsymbol{\pi}^2$ (i.e., the ith row of $\boldsymbol{\alpha}$ is $\boldsymbol{\pi}^i$). Clearly, $\boldsymbol{\alpha}$ is TP_2. Therefore, $\boldsymbol{\eta}^{h,h'} \equiv \boldsymbol{\alpha} \cdot \boldsymbol{\Theta}$ is TP_2 by Lemma 1. As a result, rows of $\boldsymbol{\eta}^{h,h'}$ are in increasing stochastic order. That is, denoting the ith row of $\boldsymbol{\eta}^{h,h'}$ by $\boldsymbol{\eta}^{h,h'}(\boldsymbol{\pi}^i, \cdot)$,

$$\boldsymbol{\eta}^{h,h'}(\boldsymbol{\pi}^1, \cdot) \preceq_S \boldsymbol{\eta}^{h,h'}(\boldsymbol{\pi}^2, \cdot) \quad \text{for } \boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2. \tag{A1}$$

Also note that, for $l' \leq l''$ and $y' \leq y''$, Proposition 2(b) implies $\boldsymbol{\pi}'(\boldsymbol{\pi}, l', y') \preceq_{LR} \boldsymbol{\pi}'(\boldsymbol{\pi}, l'', y')$, because $\mathcal{L}$ is TP_2, and Proposition 2(d) implies $\boldsymbol{\pi}'(\boldsymbol{\pi}, l'', y') \preceq_{LR} \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y'')$, because $\mathcal{O}$ is more informative than $\mathcal{L}$. Employing induction on these results, we find $v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y')) \geq v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}, l'', y')) \geq v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y''))$. Furthermore, by Proposition 1(a), $v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y'')) \geq v^n(h', l'', \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y'')) \geq v^n(h', l'', \boldsymbol{\pi}'(\boldsymbol{\pi}, l'', y''))$, where the last inequality is because of the induction hypothesis and Proposition 2(b). Combining all these together, we find that $v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y'))$ is nonincreasing in (l', y'). Finally,

$$\begin{aligned}
&v^{n+1}(h, l, \boldsymbol{\pi}^1) \\
&\quad \geq r_W(h) + \lambda \cdot \sum_k \sum_{(h', l', k')} \sum_{y'} \pi^1_k \cdot P\{h', l', k' \mid h, l, k\} \\
&\qquad \cdot \mathcal{O}\{y' \mid k'\} \cdot v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}^1, l', y')) \\
&\quad = r_W(h) + \lambda \cdot \sum_{h'} \mathcal{H}\{h' \mid h\} \Bigg[\sum_{k, k', l', y'} \pi^1_k \mathcal{H}\{k' \mid h', k, h\} \\
&\qquad \cdot \mathcal{L}\{l' \mid k'\}\mathcal{O}\{y' \mid k'\} v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}^1, l', y')) \Bigg] \\
&\quad = r_W(h) + \lambda \cdot \sum_{h'} \mathcal{H}\{h' \mid h\} \Bigg[\sum_{l', y'} \boldsymbol{\eta}^{h,h'}(\boldsymbol{\pi}^1, (l', y')) \\
&\qquad \cdot v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}^1, l', y')) \Bigg] \\
&\quad \geq r_W(h) + \lambda \cdot \sum_{h'} \mathcal{H}\{h' \mid h\} \Bigg[\sum_{l', y'} \boldsymbol{\eta}^{h,h'}(\boldsymbol{\pi}^2, (l', y')) \\
&\qquad \cdot v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}^1, l', y')) \Bigg] \quad \text{(A2)} \\
&\quad \geq r_W(h) + \lambda \cdot \sum_{h'} \mathcal{H}\{h' \mid h\} \Bigg[\sum_{l', y'} \boldsymbol{\eta}^{h,h'}(\boldsymbol{\pi}^2, (l', y')) \\
&\qquad \cdot v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}^2, l', y')) \Bigg] \quad \text{(A3)} \\
&\quad = v^{n+1}(h, l, \boldsymbol{\pi}^2),
\end{aligned}$$

where inequality (A2) follows from the fact that $\lambda \geq 0$ and $\mathcal{H}\{h' \mid h\} \geq 0$ for all h', and applying Lemma 2 to Equation (A1) along with the fact that $v^n(h', l', \boldsymbol{\pi}'(\boldsymbol{\pi}, l', y'))$ is nonincreasing in (l', y'). Inequality (A3) follows from applying Proposition 2(e) with $\boldsymbol{\pi}^1 \preceq_{LR} \boldsymbol{\pi}^2$ and the induction hypothesis. □

Proof of Theorem 2. The result holds trivially for $l = L + 1$. For $l \neq L + 1$, the result follows because the value of the transplant action $r_T(h, l)$ is constant in $\boldsymbol{\pi}$ and that of the wait action is nonincreasing across MLR-ordered belief states $\boldsymbol{\pi} \in \Pi(\Psi)$ by the result of Theorem 1. □

Proof of Theorem 3. We first need two technical results before proving the main result.

Lemma 4. *Any column (or row) of a TP_2 matrix is a band vector.*

Proof of Lemma 4. We prove the result for the columns by contradiction and note that the proof for rows follows similarly. Let $P = [p_{ij}]$ be a TP_2 matrix and assume otherwise. That is, assume that there exists indices $i^- < i < i^+$ and j such that $p_{i^-,j} > 0$ and $p_{i^+,j} > 0$, but $p_{ij} = 0$. But to preserve the TP_2 property of P, the following must hold:

(i) $p_{i,j^-} = 0$ for all $j^- < j$ (otherwise, the determinant of the submatrix $\begin{bmatrix} p_{i^-,j^-} & p_{i^-,j} \\ p_{i,j^-} & p_{i,j} \end{bmatrix}$ is negative).

(ii) $p_{i,j^+} = 0$ for all $j^+ > j$ (otherwise, the determinant of the submatrix $\begin{bmatrix} p_{i,j} & p_{i,j^+} \\ p_{i^+,j} & p_{i^+,j^+} \end{bmatrix}$ is negative).

These two results, together with $p_{ij} = 0$, imply an entire row of zeros for P. However, P cannot have a complete row of zeros. □

Lemma 5. *For any given matrix $P = [p_{ij}]$ and indices j_1 and j_2 such that $j_1 < j_2$, let $\mathcal{I}_1 = \{i\colon p_{i,j_1} > 0\}$ and $\mathcal{I}_2 = \{i\colon p_{i,j_2} > 0\}$. Also let $i_n^{\max} = \max\{i\colon i \in \mathcal{I}_n\}$ and $i_n^{\min} = \min\{i\colon i \in \mathcal{I}_n\}$ for $n = 1, 2$. If P is TP_2, then the following must hold*:

(a) *If $\mathcal{I}_1 \cap \mathcal{I}_2 \neq \varnothing$, then $p_{i,j} > 0$ for all $i \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $j_1 \leq j \leq j_2$.*

(b) *If $\mathcal{I}_1 \cap \mathcal{I}_2 = \varnothing$, then for each j such that $j_1 \leq j \leq j_2$, $p_{i,j} > 0$ for some i such that $i_1^{\min} \leq i \leq i_2^{\max}$.*

Proof of Lemma 5. (a) The proof follows directly from Lemma 4.

(b) Assume $\mathcal{I}_1 \cap \mathcal{I}_2 = \varnothing$. Clearly, $i_1^{\max} < i_2^{\min}$, because otherwise $p_{i_2,j_1} \cdot p_{i_1,j_2} - p_{i_1,j_1} \cdot p_{i_2,j_2} < 0$ for $j_1 < j_2$, $i_1 \in \mathcal{I}_1$, and $i_2 \in \mathcal{I}_2$, and therefore P cannot be TP_2.

Consider an index j such that $j_1 < j < j_2$. If we assume the contrary that $p_{ij} = 0$ for all i such that $i_1^{\min} \le i \le i_2^{\max}$, then either of the following two must be true: (i) $p_{ij} > 0$ for some $i < i_1^{\min}$, or (ii) $p_{ij} > 0$ for some $i > i_2^{\max}$. If (i) is true, then the determinant of $\begin{bmatrix} p_{i,j_1} & p_{i,j} \\ p_{i_1,j_1} & p_{i_1,j} \end{bmatrix}$ for $i_1 \in \mathcal{I}_1$ is negative and therefore P cannot be TP_2. If, on the other hand, (ii) is true, then the determinant of $\begin{bmatrix} p_{i_2,j} & p_{i_2,j_2} \\ p_{i,j} & p_{i,j_2} \end{bmatrix}$ for $i_2 \in \mathcal{I}_2$ is negative and therefore P cannot be TP_2. □

Let $\boldsymbol{\pi}$ be a band belief vector. Then, for some given indices $i^- \le i^+$,

$$\pi_i = \begin{cases} \text{positive} & \text{if } i^- \le i \le i^+, \\ \text{zero} & \text{otherwise.} \end{cases} \tag{A4}$$

Let κ denote the denominator of the Bayesian updating formula (2). We can rewrite the jth component of the updated belief vector $\boldsymbol{\pi}'(\boldsymbol{\pi}, l', y')$, for some given h, h', y' and l', as

$$\begin{aligned} \pi'_j &= \kappa^{-1} \cdot \mathcal{O}\{y' \mid j\} \cdot \mathcal{L}\{l' \mid j\} \cdot \sum_i \mathcal{H}\{j \mid h', i, h\} \cdot \pi_i \\ &= \kappa^{-1} \cdot \mathcal{O}\{y' \mid j\} \cdot \mathcal{L}\{l' \mid j\} \\ &\quad \cdot \sum_{i=i^-}^{i^+} \mathcal{H}\{j \mid h', i, h\} \cdot \pi_i \quad \forall\, j \in \Psi, \end{aligned} \tag{A5}$$

where κ is the same positive constant for all $j \in \Psi$. To prove the intended result, we need to show that if $\pi'_{j^-} > 0$ and $\pi'_{j^+} > 0$ for some $j^- < j^+$, then $\pi'_j > 0$ for all j such that $j^- \le j \le j^+$. Assume that $\pi'_{j^-} > 0$ and $\pi'_{j^+} > 0$ for some indices $j^- < j^+$. We analyze the three components of Equation (A5):

(1) For $j = j^-, j^+$, $\pi'_j > 0$ implies $\mathcal{O}\{y' \mid j\} > 0$. Because $\mathcal{O}$ is TP_2, Lemma 4 implies that $\mathcal{O}\{y' \mid j\} > 0$ for all j such that $j^- \le j \le j^+$.

(2) For $j = j^-, j^+$, $\pi'_j > 0$ implies $\mathcal{L}\{l' \mid j\} > 0$. Because $\mathcal{L}$ is monotone in j, we find that $\mathcal{L}\{l' \mid j\} > 0$ for all j such that $j^- \le j \le j^+$.

(3) For $j = j^-, j^+$, $\pi'_j > 0$ implies $\sum_{i=i^-}^{i^+} \mathcal{H}\{j \mid h', i, h\} \cdot \pi_i > 0$. This fact along with Equation (A4) implies that $\mathcal{H}\{j^- \mid h', i_1, h\} > 0$ and $\mathcal{H}\{j^+ \mid h', i_2, h\} > 0$ for at least one $i_1, i_2 \in \{i^-, \ldots, i^+\}$. Let $\mathcal{I}_1 = \{i\colon \mathcal{H}\{j^- \mid h', i, h\} > 0\}$ and $\mathcal{I}_2 = \{i\colon \mathcal{H}\{j^+ \mid h', i, h\} > 0\}$. Lemma 5 implies that, for each j such that $j^- \le j \le j^+$, $\mathcal{H}\{j \mid h', i, h\}$ must be strictly positive for some i such that $i_1^{\min} \le i \le i_2^{\max}$, where $i_1^{\min} = \min\{i\colon i \in \mathcal{I}_1\}$ and $i_2^{\max} = \max\{i\colon i \in \mathcal{I}_2\}$. Therefore, $\sum_{i=i^-}^{i^+} \mathcal{H}\{j \mid h', i, h\} \cdot \pi_i > 0$ for each j such that $j^- \le j \le j^+$.

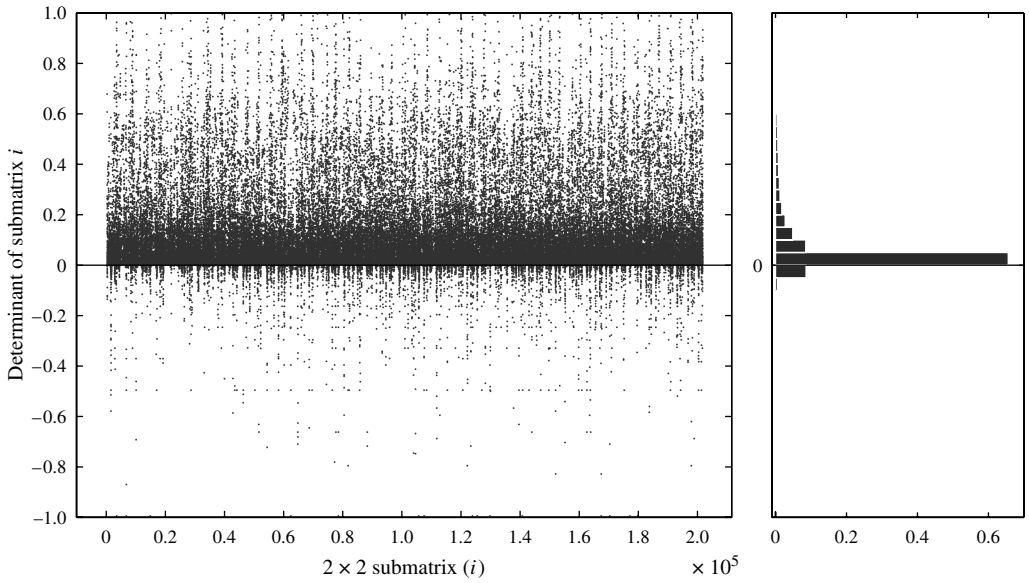
Combining these three results in Equation (A5), we conclude that if $\boldsymbol{\pi}$ is a band belief vector, then $\pi'_{j^-} > 0$ and $\pi'_{j^+} > 0$ for some $j^- < j^+$ imply $\pi'_j > 0$ for all $j^- \le j \le j^+$. □

Appendix B. Assessing the Robustness of Various Assumptions

We investigate the degree to which the conditions required for the structural results are satisfied by our parameter estimates. The details of parameter estimation are provided in §§6.1 and 6.2. Note that we estimate a new set of posttransplant rewards ($r_T(h, l)$ for all h and l) for each patient in our cohort and one set of $\mathcal{H}$, $\mathcal{L}$, and $\mathcal{O}$ matrices for each OPO.

The monotonicity condition on $r_T(h, l)$ in Proposition 1 is satisfied by all the reward estimates, as well as the monotonicity condition on $\mathcal{L}$ in Theorem 3 and Proposition 2(a). Furthermore, all of the $\mathcal{L}$ and $\mathcal{O}$ estimates yield TP_2 matrices as required for Theorems 1–3. However, $\mathcal{H}^{h,h'}$ matrices are not always TP_2, although violations are not significant. Across all $\mathcal{H}^{h,h'}$ estimates, on average, 98.7% of all 2×2 submatrices have a zero determinant, 1.2% have a positive determinant, and the remaining 0.1% have a negative determinant (indicating a violation of the TP_2 condition). Figure B.1 plots a scatter chart and the histogram of the nonzero determinants of all $\mathcal{H}^{h,h'}$ submatrices for a typical $\mathcal{H}$ matrix. The conditional fraction of negative determinants

Figure B.1 Scatter Plot and Histogram of the Nonzero Minors of Size 2×2 of the $\mathcal{H}^{h,h'}$ Matrices



amount to only 10% (the bottom half of the histogram) and those violations average into a negligible −0.02. Last, the condition on $\mathcal{O}$ being more informative than $\mathcal{L}$ is clearly more restrictive. Across all $\mathcal{O}$ and $\mathcal{L}$ estimates, the average fraction of time $\sigma^{\mathcal{O}} < \sigma^{\mathcal{L}}$ (indicating a violation) is about 60%, but the magnitude of these violations is again not significant (the average violation is about 0.011).

To test for the robustness of the proved results (i.e., the threshold-type policy proven in Theorem 2), we investigate the impact of the mentioned violations. For this purpose, we restrict our attention to MLR-ordered belief vectors. The points in a grid that is constructed as in §5, in general, may not be completely ordered in the MLR sense. There is, however, a subset of such points that can be completely ordered. One such set is $\mathcal{G}_2 = \{\boldsymbol{\pi} \in \mathcal{G}\colon \sum_i \mathbf{1}_{>0}(\pi_i) \le 2\}$, where $\mathcal{G}$ is the grid obtained as in §5 for any grid resolution $q \ge 2$ and $\mathbf{1}_{>0}(\cdot)$ is the indicator function indicating a positive value. In our numerical study, we find that the optimal actions associated with the points in $\mathcal{G}_2$ exhibit the threshold-type structure for all 200 instances solved under any grid resolution. These results lead us to believe that the indicated violations on the conditions required for proving Theorem 2 are not significant and such small violations have minimal to no impact on the end result.

References

Ahn JH, Hornberger JC (1996) Involving patients in the cadaveric kidney transplant allocation process: A decision-theoretic perspective. *Management Sci.* 42(5):629–641.

Alagoz O, Maillart LM, Schaefer AJ, Roberts MS (2007a) Choosing among living-donor and cadaveric livers. *Management Sci.* 53(11):1702–1715.

Alagoz O, Maillart LM, Schaefer AJ, Roberts MS (2007b) Determining the acceptance of cadaveric livers using an implicit model of the waiting list. *Oper. Res.* 55(1):24–36.

Alagoz O, Bryce CL, Schaefer AJ, Chang CH, Angus DC, Roberts MS (2005) Incorporating biological natural history in simulation models: Empiric estimates of the progression of end-stage liver disease. *Medical Decision Making* 25(6):620–632.

Ayer T, Alagoz O, Stout NK (2012) A POMDP approach to personalize mammography screening decisions. *Oper. Res.* 60(5): 1019–1034.

Cassandra AR (1997) A survey of POMDP applications. American Association of Artificial Intelligence (AAAI) Fall Symposium. Technical report, AAAI, Palo Alto, CA.

Cassandra AR (1998) Exact and approximate algorithms for partially observable Markov decision processes. Ph.D. thesis, Brown University, Providence, RI.

David I, Yechiali U (1995) One-attribute sequential assignment match processes in discrete time. *Oper. Res.* 43(5):879–884.

Dew MA, Switzer GE, Goycoolea JM, Allen A, DiMartini A, Kormos RL, Griffith BP (1997) Does transplantation produce quality of life benefits?: A quantitative analysis of the literature. *Transplantation* 64(9):1261–1273.

Harper AM, Taranto SE, Edwards EB, Daily OP (2000) Organ transplantation policies: An update on a successful simulation project: The UNOS liver allocation model. Joines JA, Barton RR, Kang K, Fishwick PA, eds. *WSC '00: Proc. 32nd Conf. Winter Simulation* (Society for Computer Simulation International, San Diego), 1955–1962.

Hauskrecht M (1997) Incremental methods for computing bounds in partially observable Markov decision processes. Kuipers B, Webber BL, eds. *Proc. 14th Natl. Conf. Artificial Intelligence (AAAI-97)* (AAAI Press/MIT Press, Danver, MA), 734–739.

Hauskrecht M, Fraser H (2000) Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artificial Intelligence in Medicine* 18(3):221–244.

Hu C, Lovejoy WS, Shafer SL (1996) Comparison of some suboptimal control policies in medical drug therapy. *Oper. Res.* 44(5):696–709.

Karlin S (1968) *Total Positivity*, Vol. I. (Stanford University Press, Stanford, CA).

Law AM (2007) *Simulation Modeling and Analysis*, 4th ed. (McGraw-Hill, New York).

Lovejoy WS (1987) Some monotonicity results for partially observed Markov decision processes. *Oper. Res.* 35(5): 736–743.

Lovejoy WS (1991a) Computationally feasible bounds for partially observed Markov decision processes. *Oper. Res.* 39(1): 162–175.

Lovejoy WS (1991b) A survey of algorithmic methods for partially observed Markov decision processes. *Ann. Oper. Res.* 28(1):47–66.

Maillart LM, Ivy JS, Ransom S, Diehl K (2008) Assessing dynamic breast cancer screening policies. *Oper. Res.* 56(6):1411–1427.

Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, TerBorg PCJ (2000) A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology* 31(4):864–871.

Ohnishi M, Morioka T, Ibaraki T (1994) Optimal minimal-repair and replacement problem of discrete-time Markovian deterioration system under incomplete state information. *Comput. Indust. Engrg.* 27(1–4):409–412.

Poupart P (2005) Exploiting structure to efficiently solve large scale partially observable Markov decision processes. Ph.D. thesis, University of Toronto, Toronto, ON.

Puterman ML (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, New York).

Roberts MS, Angus DC, Bryce CL, Valenta Z, Weissfeld L (2004) Survival after liver transplantation in the United States: A disease-specific analysis of the UNOS database. *Liver Transplantation* 10(7):886–897.

Sandıkçı B (2008) Estimating the price of privacy in liver transplantation. Ph.D. thesis, University of Pittsburgh, Pittsburgh.

Sandıkçı B, Maillart LM, Schaefer AJ, Alagoz O, Roberts MS (2008) Estimating the patient's price of privacy in liver transplantation. *Oper. Res.* 56(6):1393–1410.

Shaked M, Shantikumar JG (2007) *Stochastic Orders* (Springer, New York).

Shechter SM, Bryce CL, Alagoz O, Kreke JE, Stahl JE, Schaefer AJ, Angus DC, Roberts MS (2005) A clinically based discrete-event simulation of end-stage liver disease and the organ allocation process. *Medical Decision Making* 25(2):199–209.

Smallwood RD, Sondik EJ (1973) The optimal control of partially observable Markov processes over a finite horizon. *Oper. Res.* 21(5):1071–1088.

Spaan MTJ, Vlassis N (2005) Perseus: Randomized point-based value iteration for POMDPs. *J. Artificial Intelligence Res.* 24(2):195–220.

Striebel CT (1965) Sufficient statistics in the optimal control of stochastic systems. *J. Math. Anal. Appl.* 12(3):576–592.

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

Su X, Zenios SA (2005) Patient choice in kidney allocation: A sequential stochastic assignment model. *Oper. Res.* 53(3): 443–455.

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

United Network for Organ Sharing (UNOS) (2010a) Policy 3.6: Allocation of livers. Retrieved January 1, 2010, http://www.unos.org.

United Network for Organ Sharing (UNOS) (2010b) View data reports. Retrieved January 1, 2010, http://www.unos.org/data.

U.S. Department of Health and Human Services (HHS) (2008) Annual report of the U.S. Organ Procurement and Transplantation Network and the Scientific Registry of Transplant Recipients: Transplant data 1998–2007. Retrieved January 1, 2010, http://srtr.transplant.hrsa.gov/annual_reports/2010/ar_archives.htm.

Yokota F, Thompson KM (2004) Value of information literature: A review of applications in health risk management. *Medical Decision Making* 24(3):287–298.

Zenios SA (1999) Modeling the transplant waiting list: A queuing model with reneging. *Queuing Systems* 31(3–4):239–251.

Zenios SA (2004) Models for kidney allocation. Brandeau ML, Sainfort F, Pierskalla WP, eds. *Operations Research and Health Care: A Handbook of Methods and Applications* (Kluwer Academic Publishers, Boston), 537–554.

Zenios SA, Chertow GM, Wein LM (2000) Dynamic allocation of kidneys to candidates on the transplant waiting list. *Oper. Res.* 48(4):549–569.

Zenios SA, Wein LM, Chertow GM (1999) Evidence-based organ allocation. *Amer. J. Medicine* 107(1):52–61.

Zhang J, Denton BT, Balasubramanian H, Shah ND, Inman BA (2012) Optimization of prostate biopsy referral decisions. *Manuafacturing Service Oper. Management* 14(4):529–547.

Zhou R, Hansen EA (2001) An improved grid-based approximation algorithm for POMDPs. *Proc. 17th Internat. Joint Conf. Artificial Intelligence* (*IJAI-01*) (Morgan Kaufmann, San Francisco), 707–714.