Patient Choice in Kidney Allocation:
A Sequential Stochastic Assignment Model

Xuanming Su
Walter A. Haas School of Business, University of California, Berkeley, California 94720, xuanming@haas.berkeley.edu

Stefanos A. Zenios
Graduate School of Business, Stanford University, 518 Memorial Way, Stanford, California 94305, stefzen@stanford.edu

This paper investigates the effect of patient choice on kidney allocation using the following sequential stochastic assignment model. There are $n$ transplant patients to be allocated $n$ kidneys that will arrive sequentially. Each patient and each kidney has its own type, kidney types are random and revealed upon arrival, and the reward from allocating a kidney to a particular patient depends on both their types. Patients may choose to accept or decline any kidney offer.

The objective is to determine a kidney allocation policy that maximizes total expected reward subject to the constraint that patients will only accept offers that maximize their own expected reward. A partition policy, in which the space of kidney types is divided into different domains (each corresponding to a different patient type) and in which each kidney is allocated to the patient type corresponding to its domain, is shown to be asymptotically optimal when patients must accept all kidney offers. To reflect patient choice, an incentive compatibility condition is derived to ensure that the offers made by the allocation policy are never declined. This condition is then used to derive a “second-best” partition policy. A numerical example, based on data from the US transplantation system, demonstrates that patient choice may introduce substantial inefficiencies, but the second-best policy recovers all the losses by minimizing the variability in the type of offers expected by each patient. Thus, policy makers should explicitly recognize the effect of patient choice when designing a kidney allocation system.

Subject classifications: kidney allocation: strategic behavior; incentive compatibility; assignment models: dynamic, stochastic.

Area of review: Policy Modeling and Public Sector OR.

History: Received February 2002; revisions received December 2002, October 2003, April 2004; accepted April 2004.

1. Introduction

In kidney transplantation, the conflict between patient choice and social welfare is not well recognized. While a social planner must develop and implement a kidney allocation system that enhances overall patient outcomes and promotes equity across different patient groups, transplant candidates wish to maximize their own welfare and may choose to refuse a kidney offer in anticipation of a better future offer. Existing studies either take the social planner’s perspective (see Zenios et al. 2000 or Votruba 2002) or the individual candidate’s perspective (see Ahn and Hornberger 1996) but not both. The goal of this paper is to develop a tractable modeling framework that integrates the two perspectives to assess the impact of patient choice on the kidney allocation system.

In 2000, 22,271 new candidates joined the kidney transplant waiting list but only 9,278 transplantations were performed. At the end of the year, 47,873 were on the waiting list, and its size has been growing steadily. The number of candidates dying while on the waiting list is also on the rise. Despite this acute shortage, nearly half of the available kidneys were refused by the candidates to whom these kidneys were first offered. Although clinically acceptable for transplantation, these refused kidneys do not represent an attractive option from the candidate’s viewpoint because of their quality (such as size and weight), the donor’s age, or the donor’s medical condition (Chertow 2002). During the search for alternative recipients following an offer refusal, kidneys accumulate cold ischemia time (i.e., the time during which the kidney is kept frozen), which leads to inferior transplant outcomes (Chertow et al. 1996). Kidneys are eventually discarded if not accepted for transplantation within 48 hours from the time they were retrieved from the donor. The total number of discarded kidneys in 2000 was 1,626, or roughly 15% of all those donated; see UNOS (2002).

While such wastage occurs, repeated attempts to increase the supply of donor kidneys remain unsuccessful. In this light, Gridelli and Remuzzi (2000) argue that the marginal kidneys (e.g., from older donors or from donors with health problems) discarded in the current system should be better utilized and several researchers have offered suggestions for their efficient placement. For example, kidneys from elderly donors as old as 89 years can be successfully transplanted to elderly recipients, with average kidney life exceeding the recipients’ remaining life expectancy (Andres et al. 2000). Similarly, kidneys from donors with Hepatitis C can be offered to candidates already infected with the virus (Chertow 2002). However, such allocations are infeasible if transplant candidates refuse kidneys from these “marginal” donors.
In this paper, we develop a modeling framework that captures the conflict between patient choice and social welfare. The sequential stochastic assignment model of Derman et al. (1972), hereafter referred to as the DLR model, provides an appealing starting point. The DLR model assumes that there is a pool of transplant candidates of different (known) types. A finite supply of kidneys arrives sequentially and is allocated to the transplant candidates. The kidney types are random and revealed upon kidney arrival. When a kidney is allocated to a candidate, the candidate receives a reward that depends on both the kidney and the patient types. The reward can be any clinically-valid measure of transplant success such as quality adjusted life expectancy, short-term or long-term graft survival, etc; the medical term “graft” refers to the transplanted kidney. A “social planner” will presumably allocate the kidneys to maximize total expected reward for all transplant candidates. By contrast, individual transplant candidates wish to maximize their own expected reward and may decline a kidney offer if they expect future offers to be more attractive.

While the DLR modeling framework is stylized and suppresses important dynamic features of the underlying system, it captures two forces underlying patient choice. First, by having a random stream of kidneys, the model captures the value of declining a kidney: because a future offer may be more appealing than the current one, declining that offer and waiting can be advantageous. Second, because the supply of kidneys in the model is finite, offer declinations can be “costly” because some patients who refuse an offer may never receive another one. Our premise is that these two conflicting forces shape patient choice and interfere with the social planner’s overall objective. Hence, the analysis of our stylized model will show how to balance these two forces and alleviate welfare losses caused by self-interested patient behavior.

The analysis proceeds in two steps. Starting with a scenario that assumes patients will accept all kidneys offered, we find the allocation policy that maximizes social welfare; this is referred to as the “first-best” policy (the term is borrowed from the economics literature) as it represents an unattainable ideal given that patients do have choice. The second step introduces patient choice and derives an incentive compatibility condition, which an allocation policy must satisfy in order that candidates accept all kidney offers. Imposing this condition, we modify the first-best policy to obtain the so-called “second-best” policy: this policy maximizes total expected reward while ensuring that organ offers are never declined. A comparison of the first-best and second-best policies shows that patient choice causes welfare losses if there is significant variability in kidney types expected by each candidate, if the imbalance between supply and demand for kidneys is not acute, or if patient mortality rates are low.

We then present a numerical study, which used data from the current kidney allocation environment, to provide realistic estimates for the magnitude of welfare losses caused by patient choice. We also compare the first-best and second-best policies with the first-come-first-transplanted (FCFT) policy, which is a rough representation of the current allocation system in the United States (see Zenios et al. 2000). The numerical results reveal that in the presence of patient choice, approximately 17% of the donated kidneys are discarded under the first-best policy; this percentage increases to almost 30% under the FCFT policy. On the other hand, the second-best policy ensures that patients accept all kidney offers.

The seminal work by Derman et al. (1972) analyzed the sequential allocation of a random stream of independent, identically distributed (i.i.d.) jobs to workers of differing capabilities. Albright and Derman (1972) analyzed the asymptotic behavior of the optimal assignment policy for the DLR model, and Albright (1974) considered its continuous-time analogue. Over the next 15 years, researchers pursued numerous extensions, introducing uncertainty in the number and distribution of these job arrivals (Albright 1977, Sakaguchi 1984) and allowing dependent (Kennedy 1986) and Markov-modulated job arrival processes (Nakai 1986a, b; Righter 1987, 1989). The basic model was applied to kidney allocation by David and Yechiali (1985), who considered the candidate’s (worker’s) problem of deciding whether or not to accept a kidney offer (job). They also considered the social planner’s problem in David and Yechiali (1990, 1995) and its continuous time analogue in David (1995). Our analysis extends these results in two directions. First, we impose no specific functional forms on assignment rewards. Second, we integrate the perspective of the social planner and that of the transplant candidate within a single framework.

Apart from the sequential stochastic assignment model, matching models (see Roth and Sotomayor 1992) have been used to examine issues of patient choice in kidney allocation. In particular, a recent paper by Roth et al. (2004) studies patient incentives in the context of kidney exchanges; see Rapaport (1986) and Ross and Woodle (2000) for some background of kidney exchanges.

The remainder of the paper is organized as follows. Section 2 describes the model and identifies a policy that is asymptotically optimal in the first-best case. Section 3 discusses the structural properties of the optimal policy. Section 4 develops an incentive compatibility condition and derives the second-best policy. The implications of the incentive compatibility condition are discussed in §5. An extensive numerical study is presented in §6. In §7, we discuss extensions and limitations of the model, and concluding remarks appear in §8. The appendix presents all the proofs.

2. Problem Formulation and Asymptotic Analysis

2.1. Basic Model

We first develop a model for the situation where patients do not exercise choice. In this model, there are $n$ transplant
candidates to be assigned \( n \) kidneys that will arrive sequentially over a discrete time horizon of \( n \) periods. Associated with the kidney arriving in period \( t \), where \( t = 1, \ldots, n \), is a random variable \( X_t \) indicating its type. The random variables \( \{X_t: t = 1, \ldots, n\} \) are i.i.d. with probability \( P \) over the space of all possible kidney types \( \mathcal{X} \). Each transplant candidate has its own type \( i \in \{1, \ldots, m\} \), and \( p_i \) denotes the proportion of type-\( i \) candidates. All candidates’ types are known at the beginning of the horizon, but each kidney’s type is only revealed upon arrival. Each new kidney must either be assigned to a candidate or be discarded upon arrival; if assigned, we assume that the kidney will be accepted by the candidate. The reward from assigning kidney \( x \) (where \( x \in \mathcal{X} \)) to a candidate of type \( i \) is a continuous function \( R_i(x) \in [0, B] \), and the assigned candidate becomes unavailable for future assignment. The reward function reflects graft survival or quality-adjusted life expectancy following transplantation, and can be estimated empirically from historical transplant data (details are described in §3.2).

The problem is to determine an allocation policy that maximizes total expected reward. Let \( Y_t \) denote the current candidate pool at time \( t \) and \( \{i(t, X_t, Y_t)\} \) the candidate type assigned the kidney. We seek a dynamic assignment policy \( \Pi = \{(i(t, X_t, Y_t))_{t=1}^{n}\} \) that maximizes

\[
E\left\{\sum_{t=1}^{n} R(i(t, X_t, Y_t))(X_t)\right\}.
\]

A solution to (1) is called exactly optimal (for \( n \)) if it holds for some finite value of \( n \). A solution to (1) is called asymptotically optimal if it holds as \( n \to \infty \).

This formulation relies on several assumptions. First, there are no dynamics of patient arrivals and death-related departures. Essentially, the transplant waiting list is frozen at its current composition, and future patients will be considered for transplantation only after the current waiting list is exhausted. While this is a very simplified view of the problem, it provides an analytically tractable starting point. Su (2004) describes an extension of the model that incorporates the relevant dynamics and demonstrates that the main insights remain unaffected. Second, the supply and demand for kidneys is assumed to be balanced. (Allowing the arrival of “phantom” kidneys that generate zero rewards can assure this.) Third, the objective ignores possible inequities between different patient types. Because equity is in reality a critical objective of the kidney allocation system, we discuss the impact of this assumption in §§6 and 7.

### 2.2. Partition Policies

A partition policy is defined as follows: Let \( A = \{A_i\}_{i=1}^{m} \) be a partition of the kidney space \( \mathcal{X} \), with elements \( A_i \) referred to as domains. The partition policy \( A \) allocates each kidney according to its domain: The kidney is assigned to an arbitrarily chosen candidate of type \( i \) if its type belongs to domain \( A_i \); if no candidates of type \( i \) are available, the kidney is discarded. A partition policy \( A^* = \{A_i^*\}_{i=1}^{m} \) is called optimal if it is a solution of the optimization problem

\[
\max_{\{A_i\}_{i=1}^{m}} \sum_{t=1}^{n} E[R_i(X_t)1_{X_t \in A_i}]
\]

such that \( \Pr(X \in A_i) = p_i \) \( \forall i \),

where the probability distribution of \( X \) is \( P(\cdot) \). Constraint (3) states that the expected number of kidneys assigned to each candidate type is equal to the number of candidates of that type. Without this condition, the partition policy may end up discarding kidneys, which is obviously suboptimal because assignment rewards are nonnegative and patients accept all kidney offers.

The main result states that an optimal partition policy is asymptotically optimal as \( n \to \infty \). Although an optimal partition policy is not necessarily an exactly optimal policy for a finite \( n \), its performance converges to the performance of an exactly optimal policy as \( n \) approaches infinity. To be more precise, consider an arbitrary sequence of kidneys denoted by the \( n \)-vector \( x^{(n)} \), let \( R^{opt}(x^{(n)}) \) denote the total reward from the exact optimal allocation policy (obtained by solving an appropriate dynamic program), and let \( R^{part}(x^{(n)}) \) denote the total reward from this partition policy. Then,

**Theorem 1.** The partition policy obtained by solving (2)–(3) is asymptotically optimal. That is,

\[
\lim_{n \to \infty} \left| \frac{R^{opt}(x^{(n)})}{n} - \frac{R^{part}(x^{(n)})}{n} \right| = 0.
\]

An outline of the proof follows; the details appear in the appendix. The proof uses the perfect hindsight upper bound on the optimal expected reward obtained by considering a policy that observes the complete sequence \( x^{(n)} \) of kidney realizations before making any allocation decisions. The proof then demonstrates that for large \( n \), the partition policy and the perfect hindsight policy make the exact same allocations over most of the sample paths. To obtain the optimal allocation with perfect hindsight, one must solve an assignment problem at the start of the horizon. This assignment problem coincides with (2)–(3) when the empirical distribution of realized kidney types along the observed sample path coincides with the underlying distribution. As the number of kidneys becomes increasingly large, it follows from the large deviations principle that the realized empirical distribution will become arbitrarily close to the underlying distribution. Thus, for large \( n \), the assignments made by the perfect hindsight policy will be almost identical to those made by the partition policy, which is hence asymptotically optimal. For future reference, we shall call this the first-best policy.

Optimal policies for the special cases \( R_i(x) = c_i x \) and \( R_i(x) = 1_{x \leq i} \) have previously been obtained by Albright and Derman (1972) and David and Yechiali (1990), respectively; their solutions are consistent with the more general findings in Theorem 1 (see Su 2004).
The first-best policy enjoys several advantages: It reduces the sequential stochastic assignment problem to a set-partitioning problem, it is stationary, and it needs to be computed only once at the start of the planning horizon. Numerical experiments (see Su 2004) also show that, under several special cases for which exact optimal policies are known, more than 95% of optimality is attained when \( n = 50 \) and more than 99% of optimality is attained when \( n = 500 \).

### 2.3. The Case with Discrete Kidney Types

Suppose that the space \( \mathcal{X} \) consists of \( k \) discrete kidney types with probability distribution \( (q_1, \ldots, q_k) \), and the reward from assigning a kidney of type \( j \in \{1, \ldots, k\} \) to a candidate of type \( i \in \{1, \ldots, m\} \) is \( r_{ij} \). The partition policy can then be represented by the set of numbers \( \{a_{ij}\}_{1 \leq i \leq m, 1 \leq j \leq k} \). We can interpret \( a_{ij} \) as the joint probability that a kidney will be of type \( j \) and will be assigned to a type-\( i \) patient. This implies that the fraction of type-\( j \) kidneys assigned to type-\( i \) candidates is given by \( a_{ij}/\sum_{m=1}^{m} a_{ij} \). Then, problem (2)–(3) becomes the assignment problem

\[
\max_{\{a_{ij}\}} \sum_{j=1}^{m} \sum_{i=1}^{k} a_{ij} r_{ij} \tag{5}
\]

such that

\[
\sum_{i=1}^{m} a_{ij} = q_j \quad \forall j, \tag{6}
\]

\[
\sum_{j=1}^{k} a_{ij} = p_i \quad \forall i. \tag{7}
\]

By Theorem 1, the solution to (5)–(7) yields the first-best policy.

### 3. Structural Properties of the Optimal Policy

This section develops structural properties of the first-best policy and discusses their practical implications. Subsection 3.1 introduces the so-called “increasing differences condition” that establishes a natural ordering of candidate and kidney types, and develops an explicit expression for the first-best policy. Subsection 3.2 explains how this condition arises in the context of kidney transplantation and provides an interpretation of the structure of the resulting policy.

#### 3.1. The Increasing Differences Property

**Definition 1.** A sequence of reward functions \( R_1, \ldots, R_m \) satisfies the increasing differences property if \( R_1 > \cdots > R_m \), where \( R_i > R_j \) denotes the condition that \( R_i(x) - R_j(x) \) is increasing in \( x \).

The main result is as follows:

**Proposition 1.** With increasing differences, the first-best policy is given by the partition \( A^* = \{A^*_i\}_{i=1}^{m} \), where \( A^*_i = \{a_{i-1}, a_i\} \), \( a_0 = -\infty \), \( a_m = \infty \), and the \( a_i \)'s are selected such that

\[
\Pr(X \leq a_i) = p_1 + \cdots + p_i. \tag{8}
\]

Intuitively, the increasing differences property implies that candidate types can be ordered via their corresponding reward functions. Then, the optimal partition is obtained by laying out “intervals” of probability \( p_i \) over \( \mathcal{X} \) in the above order, with the leftmost interval corresponding to type 1 candidates, the next one to type 2 candidates, and so on. Note that the optimal thresholds \( a_i \) that separate the domains \( A^*_i \) are obtained from the critical fractile solution (8) that balances (in expectation) supply and demand for each candidate type.

#### 3.2. Justifying Increasing Differences Using Proportional Hazards Models

Beyond its desirable analytical properties, the increasing differences assumption seems to be a natural one in the context of kidney transplantation. Specifically, it requires an ordering of candidate and donor types such that the incremental benefit from receiving a “better” kidney is larger for “better” candidates. Such an ordering can be made precise using the notion of relative risk from proportional hazards survival models (Cox and Oakes 1984), which are widely used to assess transplant outcomes.

The hazard function \( h(t) \) is the conditional probability that the graft will fail at time \( [t, t+\delta) \) given that it is still functioning at time \( t \). If \( z \) is a vector of covariates that include the relevant clinical attributes of the recipient and the transplanted kidney, the hazard function is given by \( h(t) = h_0(t) e^{\theta^T z} \), where \( \theta \) is the vector of regression coefficients, and \( h_0(t) \), the “baseline” hazard, can be estimated empirically using standard techniques (see Cox and Oakes 1984). Therefore, a transplant has a higher risk of failure if it involves covariates with a higher “relative risk” \( e^{\theta^T z} \). The reward function can be then be defined to be \( S(t) \), the probability that the kidney functions beyond a given time \( t \), which is given by \( S(t) = \exp[-\int_0^t h(u) du] \). Alternatively, it can be defined to be the life expectancy of the kidney given by \( \int_0^\infty S(t) dt \).

Simple algebra shows that the hazard-based reward functions described above satisfy the increasing differences property if the covariates \( z \) do not include an interaction between the clinical characteristics of the donor and the candidate. While such interactions are currently present because of the role of tissue matching, their effect seems to be diminishing over time (see Su et al. 2004b) due to improvements in immunosuppressive drugs. Therefore, increasing differences will become an appropriate approximation in the future. This implies that an effective kidney allocation policy can be derived by ranking the candidates and donors according to their contribution to relative risk.

Then, kidneys from low-risk donors are allocated to low-risk candidates, while kidneys from high-risk donors are allocated to high-risk candidates.
4. A Model for Patient Choice

We now assume that there is patient choice. Upon receiving a kidney offer, a patient faces two alternatives: (i) accept the kidney for transplantation, or (ii) decline the offer and rejoin the candidate pool for a future reassignment. Assuming that candidates are rational, they will accept an offer that maximizes their individual expected reward. In the presence of choice, the planner must specify not only the particular patient who will be offered a kidney, but also what to do when an offer is refused: who should be offered the kidney next and how should the patient who refuses the offer be treated.

To simplify the analysis, we make three critical assumptions. First, we assume that within each candidate type, kidneys are allocated according to the FCFT rule. That is, candidates of each type are arranged into separate waiting lists, and each arriving kidney is offered to the candidate at the top of the waiting list of the chosen type. Second, we assume that kidneys declined by the first candidate will not be reassigned to another candidate. (This is a restrictive assumption, made purely for the sake of analytical tractability; because multiple offers of the same kidney occur in practice, this assumption is relaxed in the numerical investigation of §6.) Third, following an offer refusal, the candidate is penalized by moving to the $K$th position of the waiting list. In particular, $K=1$ implies that candidates are not penalized for refusing kidney offers because they retain their position at the head of the line.

Four additional assumptions are made about the information and reward structures: (a) The candidates and the social planner are equally well informed about candidate and kidney types and the reward functions; (b) Candidates may discount future rewards using a discount factor $\delta \leq 1$, while the social planner is interested in long-run average rewards; (c) The reward for a type-$i$ patient receiving a kidney $x$ is $R_i(x)$, which is the same as the reward obtained by the social planner; and (d) Candidates who have not accepted an offer by the end of the planning horizon receive zero reward.

Assumption (a) implies that candidates do not have private information about their preferences and health condition. Because information asymmetry magnifies the candidate’s opportunity for strategic behavior, our assumption understates the effects of patient choice. A model with information asymmetry is examined in Su and Zenios (2004b). Assumption (b) implies that a kidney transplant today is perceived by the candidate to be better than a kidney transplant of exactly the same quality in the future. This assumption attempts to capture in an imperfect and imprecise way the possibility that a patient may die before she receives a kidney, and it has the same behavioral effect as that of patient death in a more complete model (see §7 and Su 2004 for further details). However, the social planner is less concerned about time trade-offs given that it is overseeing a very large population of transplant candidates with frequent kidney arrivals. Assumption (c) implies that candidates wish to maximize the outcome of their own transplant, while the social planner wishes to maximize aggregate outcomes for all candidates. Finally, assumption (d) reflects the limited supply of kidneys and the possibility of not receiving a viable kidney by the end of the planning horizon.

We now proceed in two steps: First, we derive an incentive compatibility condition for an arbitrary partition policy. This condition ensures that transplant candidates will not wish to refuse a kidney offer. In the second step, we formulate the problem of determining the optimal incentive compatible partition policy (second-best policy), which can be obtained by solving the incentive constrained analogue of the partitioning problem (2)–(3). Our main result is Theorem 2, which shows that the incentive compatibility condition can be expressed simply.

**Theorem 2.** For an $n$-period allocation problem, consider a partition policy $A$ with $P(X \in A_i) \leq p_i$. Suppose that the following condition holds for $i = 1, \ldots, m$:

$$\inf_{x \in A_i} R_i(x) \geq \left( \frac{\delta P(X \in A_i)}{1 - \delta [1 - P(X \in A_i)]} \right)^{K} \cdot E[R_i(X) \mid X \in A_i]. \quad (9)$$

Then, the probability that at least one kidney is declined during the entire time horizon approaches zero as $n \to \infty$.

A partition policy satisfying the conditions of Theorem 2 is said to be asymptotically incentive compatible. The proof relies on a dynamic programming argument and is given in the appendix, but the idea behind (9) is intuitive. The left-hand side of (9) gives the lowest possible reward offered to a type-$i$ candidate, while the right-hand side is an upper bound on a candidate’s expected future reward if he or she refuses the current offer. If the worst-case reward from accepting an offer today is no less than the expected future reward from waiting, then it is optimal to accept all possible offers, and the partition policy will be incentive compatible.

We are now in a position to modify the formulation of (2)–(3) to obtain the second-best policy:

$$\max_{\{A_i\}_{i=1}^{m}} \sum_{i=1}^{m} E[R_i(X) \mathbb{1}_{\{X \in A_i\}}] \quad (10)$$

such that

$$\inf_{x \in A_i} R_i(x) \geq \left( \frac{\delta P(X \in A_i)}{1 - \delta [1 - P(X \in A_i)]} \right)^{K} \cdot E[R_i(X) \mid X \in A_i] \quad \forall i \quad (11)$$

and $\Pr(X \in A_i) \leq p_i \quad \forall i \quad (12)$

Compared to (2)–(3), which does not involve patient choice, this formulation includes the incentive compatibility constraint (9), and replaces the “supply balances demand” constraint (3) by the inequality constraint (11). When strict inequality holds, the assignment domains do not cover the entire kidney space $\mathcal{X}$ (so $A = \{A_i\}_{i=1}^{m}$ is no longer
considered a partition of the kidney space, but rather a collection of disjoint subsets). This implies that the social planner will have to discard some kidneys because it is not possible to design an incentive compatible policy in which supply is equal to demand. This may seem counterintuitive in view of the current kidney shortage.

One method to improve the utilization of donated kidneys is to assign the kidneys that would have been discarded (i.e., with types $x \notin \bigcup A_i$) to an arbitrary patient in position $L > K$, where $L$ is sufficiently large (so this patient is willing to accept it). While this does not influence the incentives of the first $K$ patients in line, it will influence the incentives of patients further down the line, and raises further questions: How far “down the line” should one go to find patients who will accept these kidneys? How does the reallocation of refused organs affect the incentives of patients below position $K$? What is the optimal way to ration the refused organs among the various patient types? We leave these extensions for future research. However, we do note that the benefits from reallocating refused organs are likely to be miniscule because the numerical results in §6 demonstrate that the second-best policy described above utilizes almost all organs.

The formulation presented in §2.3 can be used to obtain the second-best policy when kidney types are discrete, by solving

$$\max_{\{\alpha_i\}} \sum_{j=1}^{m} \sum_{i=1}^{k} a_{ij} r_{ij}$$

such that

$$\sum_{j=1}^{m} a_{ij} \leq q_j \quad \forall j,$$  

$$\sum_{j=1}^{k} a_{ij} \leq p_i \quad \forall i,$$  

$$a_{ij} \left( r_{ij} - \left( \frac{\delta \sum_{j=1}^{k} a_{ij}}{1 - \delta (1 - \sum_{j=1}^{k} a_{ij})} \right)^{k} \right) \geq 0 \quad \forall i, j.$$  

The inequality constraints (14) and (15) state that in expectation, some kidneys may be discarded and some candidates may not receive transplants. The incentive compatibility constraint (16) has a complementary slackness structure because the term in parentheses needs to be nonnegative only when $a_{ij}$ is strictly positive, i.e., when candidates of type $i$ may receive kidneys of type $j$. Although the optimization problem is no longer linear, the solution can be computed using algorithms developed for mathematical programs with equilibrium constraints (see Luo et al. 1996).

**Illustrative Example**

To illustrate the differences between the first-best and second-best policies, we provide a stylized numerical example. The example assumes that kidney types are uniformly distributed over the unit interval $[0, 1]$, candidates of all types have equal proportions, the discount factor $\delta = 1$, and supply balances demand. Three different cases are considered, as summarized in Figure 1.

For each case, assignment domains for the numerically computed second-best policy are shown below the horizontal axis, and the remaining region (a single interval) represents kidneys that have to be discarded. The second-best partition consists of intervals ranked according to the increasing differences ordering of the reward functions (see Definition 1). The first-best partition policies can be computed easily because of the increasing differences property: for the two cases with two patient types, the first-best partition is $A_1 = [0, 1/2]$ and $A_2 = (1/2, 1]$; for the case with three patient types, the partition is $A_1 = [0, 1/3], A_2 = [1/3, 2/3], A_3 = (2/3, 1]$.

5. **Implications of Patient Choice**

Next, we explore the implications of Theorem 2. The incentive compatibility condition (9) can be rewritten as

$$\inf_{x \in A_i} R_i(x) \geq \left( \frac{\delta P(X \in A_i)}{1 - \delta (1 - P(X \in A_j))} \right)^k,$$

which reveals four factors that dictate patient choice: (a) offer variability, (b) impatience, (c) size of potential
donor pool, and (d) penalties. As the variability in future offers diminishes (i.e., as the ratio \( \inf_{x \in A_i} R_i(x)/E[R_i(X) \mid X \in A_i] \) increases and becomes closer to 1), waiting becomes less attractive and patients are more likely to accept the existing offers. Similarly, as patients become increasingly impatient (i.e., \( \delta \) becomes smaller), the attractiveness of future offers diminishes, and patients become more likely to accept an existing offer. Third, as the potential donor pool becomes larger (i.e., as \( P(X \in A_i) \) increases), the likelihood of repeated offers increases and patients become more likely to decline an offer. Finally, as the penalty for refusing a kidney become more severe (i.e., as the drop in waiting list ranking \( K \) increases), patients become more reluctant to reject an offer.

The incentive compatibility condition changes depending on the system used to “penalize” patients who refuse an offer. For a system where candidates who have refused a kidney offer are moved to a randomly chosen position on the waiting list, Su (2004) shows that the incentive compatibility condition is

\[
\frac{P_i}{\delta P(X \in A_i)} \cdot \frac{\inf_{x \in A_i} R_i(x)}{E[R_i(X) \mid X \in A_i]} \geq 1. \tag{18}
\]

Here, incentive compatibility is dictated by candidate impatience, offer variability, and size of the potential donor pool as before. However, the size of the candidate pool \( P_i \) now plays the role of the penalty parameter \( K \) in (9); an increase in the relative size of the type-\( i \) waiting list discourages these candidates from refusing kidney offers, because the expected drop in waiting list position following offer refusals also increases. Note that explicit dependence on \( P_i \) is not present in (9). This suggests that, for the number of refused kidneys to decrease as waiting list sizes increase, penalties must be linked to waiting list sizes; for example, in our model, \( K \) should be increased for larger \( n \).

In the current allocation system without penalties, offer refusals are not expected to decline even as waiting list sizes continue to grow.

### 6. Numerical Study

We now investigate the effect of patient choice in the current kidney allocation environment by means of a numerical experiment based on recent transplantation data and evidence published in the literature. We compare several policies both with and without patient choice to demonstrate that ignoring patient choice in the design of the kidney allocation system diminishes its effectiveness.

Historically, transplant outcomes have been correlated with the compatibility between donor and recipient tissue types. The tissue type consists of six proteins located in six loci (A1, A2, B1, B2, DR1, DR2); the protein in each locus may be one of roughly 30 possible types. A “match” occurs if the patient and the donor have the same protein at a particular locus. Based on a proportional hazards model, Opelz et al. (1999) show that the effect of tissue matching on long-term survival is additive: Each additional match increases the five-year graft survival rate by the same number of percentage points independently of matches in any other loci. Earlier studies have found the effect of matches in the DR locus to be more significant (Chertow et al. 1996), but we will rely on the findings of the Opelz et al. (1999) study, which appears to be the most recent and comprehensive to date. Therefore, we define patient and donor types to be their tissue types, and the reward is the number of matching proteins (out of a maximum of six) in the assignment. However, this ignores the important effect of matching-independent factors such as age, gender, comorbidities. A more in-depth numerical analysis based on a more comprehensive set of patient and kidney characteristics is reported in Su et al. (2004a).

The goal of this numerical analysis is to investigate the outcome of allocating 40,000 sequentially arriving kidneys. To capture the imbalance between supply and demand, the initial candidate population is taken to be 40,000, 60,000, and 80,000 in three separate scenarios. In each scenario, a random set of candidates is first generated, followed by a random stream of 40,000 donors. Candidate and donor tissue types are generated from the distributions presented in Zenios (1996), which assumes independence of the six protein loci.

Three policies are tested in the simulations. (1) The first-best policy obtained from the solution to (2)–(3); the policy is obtained using the linear-programming formulation of §2.3 because tissue types are discrete. (2) The second-best policy computed by solving (10)–(12). (3) The FCFT policy, which resembles current allocation practices in the United States (see Zenios et al. 2000).

Two choice environments (with and without patient choice) are considered. Without choice, patients must accept all kidney offers. However, under patient choice, candidates will accept a kidney transplant offer only if it satisfies the incentive compatibility condition (9). Transplant candidates are assumed to discount future payoffs at a rate of 5% per year; the discount rate per period is thus \( 1 - \delta = (5 \times 10^{-4}) \) because approximately 10,000 cadaveric kidneys are procured each year. When a patient refuses a kidney offer, he or she is not penalized (i.e., \( K = 1 \)) and the kidney may be reassigned to another candidate according to the same allocation rule. The kidney is eventually discarded if it remains unassigned after 12 offers. This setup is designed to reflect current kidney allocation practices: Because kidneys have to be transplanted within 48 hours from the time of procurement, a time delay of four hours per offer (due to logistical reasons) implies that each kidney may not successively be offered to more than 12 candidates.

We have described two choice environments (with and without patient choice) and three policies (first-best, second-best, and FCFT). Because the second-best policy performs identically with and without choice, we have five distinct combinations to investigate. For each combination, 100 independent simulation runs are performed. For each
simulation, the following performance metrics are obtained: average number of tissue matches, and average one-year and average five-year graft survival rates after transplantation. Averages are taken over all 40,000 simulated kidneys, and over only the kidneys that were accepted in the simulation. If a simulated kidney is discarded, it counts as zero tissue matches and (more importantly) as an immediate kidney failure in the survival statistics; this is because discarded kidneys contribute to neither tissue matches nor graft survival. Survival rates are calculated using data from USRDS (2002), which incorporates the effect of cold-ischemia time. The number of discarded kidneys is also reported. Table 1 summarizes the results averaged over the 100 simulations. The simulation standard errors for all figures are of the order of $10^{-2}$ and are omitted. Figures averaged over the kidneys that were transplanted in the simulation are shown in brackets.

These results yield several observations:

(1) Patient choice significantly impacts kidney allocation system performance. The first-best policy performs poorly when patients exercise choice: In Scenario 1, for example, 6,849 out of the 40,000 kidneys (i.e., 17.1%) are discarded. These rejections imply a significant loss in functioning graft years: Compared to the corresponding outcomes for candidates who eventually receive transplants (results reported in brackets) have on average 4.56 tissue matches, which is an improvement from 4.17 matches when patients do not exercise choice. However, among this group of candidates, survival rates decrease: One-year survival decreases from 92.0% to 90.6% and five-year survival decreases from 68.1% to 63.7%. This decrease is caused by an accumulation of cold-ischemia time during the search for alternative recipients following offer refusals. Similar observations can be made for the FCFT policy.

(2) Allocating kidneys according to the FCFT policy exacerbates the effect of patient choice. In Scenario 1, 11,661 kidneys (out of 40,000) are discarded. Compared to the corresponding outcomes in the absence of choice, this kidney wastage leads to survival losses, despite an increase in the average number of tissue matches from 0.73 to 1.52. One-year survival decreases from 87.9% to 62.1% and five-year survival decreases from 62.2% to 43.9%. As in the first-best policy, an excess demand for kidneys does not neutralize the effect of patient choice.

(3) Even among candidates who eventually receive transplants, the introduction of patient choice generates a negative impact. Consider the first-best policy operating under patient choice in Scenario 1. Candidates who eventually receive transplants (results reported in brackets) have on average 4.56 tissue matches, which is an improvement from 4.17 matches when patients do not exercise choice. However, among this group of candidates, survival rates decrease: One-year survival decreases from 92.0% to 90.6% and five-year survival decreases from 68.1% to 63.7%. This decrease is caused by an accumulation of cold-ischemia time during the search for alternative recipients following offer refusals. Similar observations can be made for the FCFT policy.

(4) Redesigning the policy to explicitly capture patient choice leads to the second-best policy that performs almost as well as the frictionless first-best ideal. Here, each candidate is always assigned a kidney with the same number of matches, and this eliminates the incentive to decline kidney offers. Note that no explicit penalties were used; the policy maker has merely restricted the variability in the stream of kidneys offered to each candidate.

(5) The results highlight the negative externalities generated by patient choice. In all three scenarios and under both the first-best and FCFT policies, a wide discrepancy exists between outcomes for candidates who eventually receive kidneys and the transplant population as a whole. Patients who eventually accept a kidney enjoy reasonably good outcomes at the expense of a significant number of patients who are left without a kidney at the end of the planning

---

### Table 1. Summary of results from numerical experiments.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>First-best partition (no choice)</th>
<th>FCFT (no choice)</th>
<th>First-best partition (with choice)</th>
<th>FCFT (with choice)</th>
<th>Second-best partition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average number of matches</td>
<td></td>
<td>3.78 [4.56]</td>
<td>1.52 [2.14]</td>
<td>4.15</td>
</tr>
<tr>
<td></td>
<td>1-year graft surv. rate (%)</td>
<td>92.0</td>
<td>75.1 [90.6]</td>
<td>62.1 [87.5]</td>
<td>92.0</td>
</tr>
<tr>
<td></td>
<td>5-year graft surv. rate (%)</td>
<td>68.1</td>
<td>52.8 [63.7]</td>
<td>43.9 [61.9]</td>
<td>68.1</td>
</tr>
<tr>
<td></td>
<td>Number of discarded kidneys</td>
<td></td>
<td>6,849</td>
<td>11,661</td>
<td>0</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------</td>
<td>-----------------</td>
<td>-----------------------------------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Scenario 2: 60,000 candidates</td>
<td>Average number of matches</td>
<td>4.55</td>
<td>3.84 [4.59]</td>
<td>1.53 [2.14]</td>
<td>4.54</td>
</tr>
<tr>
<td></td>
<td>1-year graft surv. rate (%)</td>
<td>92.4</td>
<td>75.9 [90.7]</td>
<td>62.3 [87.5]</td>
<td>92.4</td>
</tr>
<tr>
<td></td>
<td>5-year graft surv. rate (%)</td>
<td>68.9</td>
<td>53.7 [64.2]</td>
<td>44.2 [62.1]</td>
<td>68.9</td>
</tr>
<tr>
<td></td>
<td>Number of discarded kidneys</td>
<td></td>
<td>6,543</td>
<td>11,550</td>
<td>0</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------</td>
<td>-----------------</td>
<td>-----------------------------------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>1-year graft surv. rate (%)</td>
<td>92.5</td>
<td>77.8 [90.8]</td>
<td>63.0 [87.7]</td>
<td>92.5</td>
</tr>
<tr>
<td></td>
<td>5-year graft surv. rate (%)</td>
<td>69.2</td>
<td>55.2 [64.5]</td>
<td>44.9 [62.2]</td>
<td>69.2</td>
</tr>
<tr>
<td></td>
<td>Number of discarded kidneys</td>
<td></td>
<td>5,743</td>
<td>11,027</td>
<td>0</td>
</tr>
</tbody>
</table>
horizon. By manipulating their option to wait, individual candidates can game the kidney allocation system if it ignores patient choice.

7. Extensions and Limitations

Su (2004) present a more complex model that includes candidate arrivals as well as death-related departures. (The online technical supplement describes this model. It is available at http://or.pubs.informs.org/Pages/collect.html.) However, the key observations remain unchanged: Partition policies are asymptotically optimal, and an incentive compatibility condition similar to (9) can be imposed to discourage patients from refusing kidney offers. The factors determining patient choice include supply and demand (i.e., the difference between candidate arrival rates and kidney arrival rates) and death rates, but these were already captured by our analysis above: The relative difference between the sizes of the donor pool $P(X \in A_i)$ and the candidate pool $p_i$ reflects the supply-demand imbalance, and the discount factor $\delta$ indirectly captures the death rate. Therefore, our insights also apply under continuous time dynamics with patient arrivals (following general renewal processes) and death.

For our results to be applicable, patient “types” have to be defined. The numerical example of §6 suggests that tissue types can be used to determine patient types. However, a wide array of other variables, such as age, comorbidities, history of dialysis therapy, and number of previous transplants also influence transplant outcomes. A straightforward approach is to discretize these clinical variables and classify patients accordingly, but this would result in an enormous number of patient types. Another possibility is to use statistical techniques to partition the space of all patients into clusters with similar risk profiles, but this approach raises another question: How many clusters (i.e., patient types) should we use? A large number of patient types would make the system too complex to implement, while a small number of patient types would create large within-group variabilities in transplant rewards, which will create incentives for offer refusals. Selection of a precise implementation methodology is left as a topic for future research.

Our model assumes that candidates within each type receive kidneys in a FCFT manner. While FCFT is often regarded as the universal gold standard for fairness, it ironically generates perverse incentives: Patients on the top of the waiting list are more inclined than patients farther down to refuse an offer. Su and Zenios (2004a) analyze the effect of the priority discipline on patient choice and establish that the FCFT rule aggravates the efficiency loss. Two mechanisms can alleviate the problem: (i) assigning absolute priority to a fixed proportion of new patients, or (ii) having future arrivals join the waiting list in front of patients who have refused offers previously, but behind all other patients. Another limitation of FCFT is that it assigns kidneys to the candidates who have been waiting for the longest time. Because the health condition of waiting candidates progressively worsens, FCFT effectively assigns kidneys to the sickest patients. Our analysis does not capture the evolution of patient health and therefore may overstate the benefits of partition policies under FCFT.

Partition policies are designed to eliminate offer refusals by anticipating them. Our model assumes that candidates of the same type have identical preferences, but in reality, patients with almost identical clinical characteristics may have widely different preferences that cannot be captured in any kidney allocation system. Therefore, there may be “unforeseen refusals” driven by patients’ subjective assessments of different kidney offers. Su et al. (2004a) examine a proposal in which candidates specify in advance the range of donor kidneys acceptable to them. Under heterogeneous preferences, this system eliminates offer refusals and leads to increased life expectancy across all demographic groups.

Finally, our analysis has focused on optimizing clinical efficiency, although equity is equally important. The term “equity” in this context captures the desire to “minimize discrepancies across patients,” and has two components: (i) process equity, achieved when allocation procedures do not discriminate between patient types either directly or indirectly, and (ii) outcomes equity, achieved when all patients enjoy equal outcomes. The partition policies considered in our analysis are not equitable because different types of patients are allocated different types of kidneys and thus experience different outcomes. The FCFT policy, in which all patients wait for kidneys in the same queue, is process equitable but not outcomes equitable due to the inherent heterogeneity present in the transplant population. The search for an allocation system that is both equitable and efficient must rely on a precise definition of equity, chosen among many possible ones. For example, Rawls’s (1999) study of fairness and justice suggests the max-min criterion: Maximize the lowest expected reward among all patients. Ultimately, public opinion will determine the most crucial equity metrics that should be used in conjunction with the less-controversial efficiency objective, to determine the “optimal” kidney allocation system. No matter which equity metric is chosen, our analysis suggests that ignoring patient choice can cause a significant waste of irreplaceable kidneys.

8. Concluding Remarks

This analysis has identified four factors that contribute to a patient’s desire to refuse kidneys allocated by a centrally designed allocation system: variability in type of kidneys offered, supply-demand imbalance, patient mortality/discounting, and penalties for kidney refusal. While some of these factors are beyond the control of policy makers, variability in kidney type can be controlled. Our numerical results suggest that in the absence of adjustments for patient choice, repeated candidate refusals cause as many as
30% of all available kidneys to be discarded. However, an adjustment that minimizes the variability in kidney offers made to each candidate resolves the problem and achieves an efficient distribution of kidneys.

The problem of patient choice in kidney allocation is multifaceted and complex. This is because a patient’s desire to exercise choice interacts with three broad factors: heterogeneity (clinical characteristics and survival prospects), implementation (allocation rules and priority points), and information (unobserved preferences and health conditions). These factors shape the strategic interactions that take place between transplant candidates and the kidney allocation system. The development of an analytical model that simultaneously captures all these factors appears to be an intractable task. Therefore, we adopt a divide-and-conquer approach. In this paper, we focus on patient heterogeneity and consider a stylized model that abstracts away incentive problems caused by either information asymmetry or by queueing priority disciplines. In Su and Zenios (2004a), we examine the role of the priority discipline, while in Su and Zenios (2004b), we develop a mechanism design model that incorporates information asymmetry. We hope that this research program will lead to a better understanding of the interplay between patient choice and various policy options.

While penalties could be used (at least in theory) to control the externalities created by patient choice, we have demonstrated that explicit penalties are not necessary to induce an efficient allocation of kidneys. However, one may argue that the modification proposed by our second-best mechanism constitutes a penalty of sorts. By restricting the allocation decision of the perfect hindsight policy. The reader may object that the modification proposed by our second-best policy are denoted

\[ R_{\text{part}}(x^{(\ast)}) \]

and allocated optimally given this perfect information), it implies the asymptotic optimality of the partition policy \( A^\ast \).

Before proceeding, we need a few more definitions. Given a sequence of kidney realizations \( x^{(i)} \), let \( R_{\text{part}}(x^{(i)}) \) denote the reward from following the partition policy \( A^\ast \), and let \( l_{\text{part}}(x^{(i)}, t) \) denote the candidate type that kidney \( x_i \) is allocated to (or zero if the kidney is discarded) for \( t = 1, \ldots, n \). For an index set \( I \subseteq \{1, \ldots, n\} \), let

\[ R_{\text{part}}(x^{(i)}) \mid I \]

denote the total rewards from allocations made during time periods \( t \in I \). Analogous quantities for the perfect hindsight policy are denoted \( R_{\text{ph}}(x^{(i)}), l_{\text{ph}}(x^{(i)}, t), \) and \( R_{\text{ph}}(x^{(i)}) \mid I \). We can now proceed with the proof.

**Typical Sequences.**

**Lemma 1.** Let \( x^{(i)} = (x_1, \ldots, x_n) \) be a typical kidney sequence, so that it consists of exactly \( n_i \) kidneys with types \( x \in A^\ast_i \) for each \( i \). Then,

(i) \( l_{\text{ph}}(x^{(i)}, t) = l_{\text{part}}(x^{(i)}, t) \) for \( t = 1, \ldots, n \).

(ii) \( R_{\text{ph}}(x^{(i)}) = R_{\text{part}}(x^{(i)}) \).

**Proof.** First, note that no kidneys are discarded under both the perfect hindsight policy (because rewards are nonnegative) and the partition policy (because \( x^{(i)} \) is typical). Let \( J \subseteq \{1, \ldots, n\} \) be the set of time indices over which these two policies make different allocation decisions.

Suppose that (i) does not hold, implying that the partition policy characterized by \( \{l_{\text{part}}(x^{(i)}, t)\}_{t=1}^n \) is suboptimal for \( x^{(i)} \) given perfect hindsight. Then, \( J \) must be nonempty and

\[ \sum_{t \in J} R_{\text{ph}}(x^{(i)}, t) > \sum_{t \in J} R_{\text{part}}(x^{(i)}, t) \]  \[ (19) \]

Now, let us modify the partition policy \( A^\ast \) over small neighborhoods of kidney types \( x_i \), for every \( t \in J \), by mimicking the allocation decisions of the perfect hindsight policy. The continuity of reward functions, together with (19), implies that the objective function (2) is improved, thus contradicting the optimality of \( A^\ast \). Therefore, (i) must be true and (ii) follows as a corollary. \( \square \)

**Approximately Typical Sequences.** For the next result, recall that \( B \) is an upper bound on all possible rewards.

**Appendix. Main Proofs**

**Proof of Theorem 1.**

**Preliminaries.** Consider the allocation problem for a fixed value of \( n \). Under the partition policy \( A^\ast \), where \( A^\ast = \{ A^\ast_i \} \) denotes the solution to (2)–(3), a kidney of type \( x \) is allocated to a type-\( i \) candidate if and only if \( x \in A^\ast_i \). Given a particular sequence of kidney realizations denoted by \( x^{(i)} = (x_1, \ldots, x_n) \), the vector-valued map \( \phi \) summarizes these allocations by giving the proportion of kidneys allocated to each candidate type. That is, \( \phi_i(x^{(i)}) = \sum_{t=1}^{n} 1_{\{x_t \in A^\ast_i\}}/n \) for \( i = 1, \ldots, m \). Note that constraint (3) implies that the expected value of \( \phi_i(x^{(i)}) \) is \( p_i \). For this reason, we say that the kidney sequence \( x^{(i)} \) is typical if \( \phi_i(x^{(i)}) = p_i \) for every \( i \). Similarly, we say that the sequence \( x^{(i)} \) is approximately typical if \( \phi_i(x^{(i)}) \) is “close” to \( p_i \) in a sense to be made precise later. Finally, all other kidney sequences are called nontypical.

Through a sequence of lemmas that consider each of these three types of kidney realizations, we will show that the difference in reward between the partition policy \( A^\ast \) and the perfect hindsight policy converges to zero as \( n \to \infty \). Lemma 1 shows that this difference is zero when the kidney sequence is typical. Next, for approximately typical sequences, Lemmas 2 and 3 establish an asymptotically negligible bound for this difference. Finally, Lemma 4 demonstrates that the probability of the nontypical region vanishes as \( n \) becomes large. Because the perfect hindsight policy yields an upper bound on optimal rewards (because kidney types are revealed at the beginning of the horizon and allocated optimally given this perfect information), it implies the asymptotic optimality of the partition policy \( A^\ast \).
Lemma 2. For every kidney sequence \(\mathbf{x}^{(n)} = (x_1, \ldots, x_n)\),
\[
R_{ph}(\mathbf{x}^{(n)}) - R_{part}(\mathbf{x}^{(n)}) \leq 2Bn \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i|.
\] (20)

Proof. Step 1. Fix the kidney sequence \(\mathbf{x}^{(n)}\). Let \(I\) denote the periods during which kidneys are discarded under the partition policy and let \(I\) denote its complement. Then, construct a kidney sequence \(\mathbf{y}^{(n)}\) by letting \(y_i = x_i\) for all \(t \in I\) and choosing \(y_i\) for all \(t \in J\) arbitrarily so that the new sequence \(\mathbf{y}^{(n)}\) is typical. Lemma 1 implies that
\[
R_{ph}(\mathbf{y}^{(n)}) = R_{ph}(\mathbf{x}^{(n)} | I) + R_{ph}(\mathbf{x}^{(n)} | J).
\] (21)

Equation (22) holds because the sequences \(\mathbf{x}^{(n)}\) and \(\mathbf{y}^{(n)}\) are identical on the index set \(I\), and Equation (23) holds because kidneys arriving in the complement of \(I\) are discarded by the partition policy. Finally, (24) is due to the upper bound \(B\).

Step 2.
\[
R_{ph}(\mathbf{x}^{(n)}) = R_{ph}(\mathbf{x}^{(n)} | I) + R_{ph}(\mathbf{x}^{(n)} | J).
\] (25)
\[
\leq R_{ph}(\mathbf{x}^{(n)} | I) + |J|B.
\] (26)
\[
= R_{ph}(\mathbf{y}^{(n)} | I) + |J|B.
\] (27)
\[
\leq R_{ph}(\mathbf{y}^{(n)}) + |J|B.
\] (28)
\[
\leq R_{part}(\mathbf{x}^{(n)}) + |J|B.
\] (29)

We use the upper bound \(B\) in (26). Inequality (29) follows from Step 1. Finally, because \(|J| \leq n \sum_{i = 1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i|\), we have the desired result.

In the next two lemmas, let \(\varepsilon > 0\) denote an arbitrary small constant.

Lemma 3. Consider the set of approximately typical kidney sequences defined by \(A^{(n)} = \{\mathbf{x}^{(n)}: \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i|^2 \leq n^{-1}\}\). Then, for each \(\mathbf{x}^{(n)} \in A^{(n)}\),
\[
R_{ph}(\mathbf{x}^{(n)}) - R_{part}(\mathbf{x}^{(n)}) \leq 2mn^{(e+1)/2}B.
\] (30)

Proof. By Lemma 2, it suffices to show that
\[
n \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i| \leq mn^{(e+1)/2}
\] for \(\mathbf{x}^{(n)} \in A^{(n)}\):
\[
n \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i| \leq nm \cdot \max_{i} |\phi_i(\mathbf{x}^{(n)}) - p_i|.
\] (31)
\[
= nm \left( \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i|^2 \right)^{1/2}
\] (32)
\[
\leq nm \left( \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i|^2 \right)^{1/2}
\] (33)
\[
\leq nm \cdot n^{(e+1)/2} = mn^{(e+1)/2},
\] (34)

where (34) follows from the definition of set \(A^{(n)}\).

Nontypical Sequences. Before we present the lemma for nontypical sequences, it is necessary to introduce Sanov’s Theorem; see Cover and Thomas (1991) for a concise treatment of this result. Note that the set of approximately typical sequences used in Lemma 3 can be equivalently defined as \(A^{(n)} = \{\mathbf{x}^{(n)}: \phi_i(\mathbf{x}^{(n)}) \in E^{(n)}\}\), where
\[
E^{(n)} = \left\{ \mathbf{z} \in \mathbb{R}_+^m: \sum_{i=1}^{m} z_i = 1, \sum_{i=1}^{m} |z_i - p_i|^2 \leq n^{-1} \right\},
\] (35)

and can be interpreted as the set of discrete probability distributions \(\mathbf{z}\) that are “close” to the distribution \(\mathbf{p}\). We will also need the concept of relative entropy defined as
\[
H(\mathbf{q} | \mathbf{p}) = m \sum_{i=1}^{m} q_i \ln \left( \frac{q_i}{p_i} \right).
\] (36)

where \(\mathbf{q}\) is also a discrete probability distribution on the set \{1, \ldots, m\}. Then, Sanov’s theorem, applied to our context, states the following.

Sanov’s Theorem. Let \(\overline{A^{(n)}}\) denote the complement of set \(A^{(n)}\). Then,
\[
Pr(\overline{A^{(n)}}) \leq (n + 1)^m \exp \left( -n \inf_{\mathbf{q} \in E^{(n)}} H(\mathbf{q} | \mathbf{p}) \right).
\]

This result will be used in the next lemma.

Lemma 4. Consider the set of nontypical sequences \(\overline{A^{(n)}} = \{\mathbf{x}^{(n)}: \sum_{i=1}^{m} |\phi_i(\mathbf{x}^{(n)}) - p_i|^2 > n^{-1}\}\). Then, for a sufficiently large \(n\),
\[
Pr(\overline{A^{(n)}}) \leq (n + 1)^m \exp(-n^e/6).
\] (37)

Proof. Let \(\mathbf{q}^* = \arg \inf_{\mathbf{q} \in E^{(n)}} H(\mathbf{q} | \mathbf{p})\). Next, choose \(n\) large enough so that \(q_i^*\) is close to \(p_i\) for \(i = 1, \ldots, m\). This ensures that \(\{(p_i - q_i^*)/q_i^*\}\) is small and, in particular, less than one. Then, recalling the definition of relative entropy in (36), we have
\[
H(\mathbf{q}^* | \mathbf{p}) = \sum_{i=1}^{m} q_i^* \ln \frac{q_i^*}{p_i} = -\sum_{i=1}^{m} q_i^* \ln \frac{p_i - q_i^*}{q_i^*},
\] (38)
\[
\geq -\sum_{i=1}^{m} (p_i - q_i^*)
\]
\[
+ \sum_{i=1}^{m} q_i^* \left[ \frac{1}{2} \frac{(p_i - q_i^*)^2}{q_i^*} - \frac{1}{3} \left( \frac{p_i - q_i^*}{q_i^*} \right)^3 \right]
\] (40)
\[
\geq \frac{1}{6} \sum_{i=1}^{m} (p_i - q_i^*)^2 \geq \frac{1}{6} n^e - 1.
\] (42)

The inequality in (40) follows from an elementary Taylor expansion. The inequality in (41) holds because \(\mathbf{p}\) and \(\mathbf{q}^*\)
both sum to one, and because
\[
\left( \frac{p_i - q_i^*}{q_i^*} \right)^2 \geq \left( \frac{p_i - q_i^*}{q_i^*} \right)^3 \quad \text{because} \quad \left| \frac{p_i - q_i^*}{q_i^*} \right| < 1
\]
for every \( i \). Next, we have (42) because \( q_i^* < 1 \) for each \( i \), and because \( q^* \notin \text{int} E^{(n)} \), so \( (q_i^* - p_i)^2 \geq n^{e-1} \) from the definition of \( E^{(n)} \) in (35). To conclude, observe that
\[
\Pr(A^{(n)}) \leq (n+1)^m \exp\{-nH(q^* | p)\} \leq (n+1)^m \exp\{-n \cdot n^{-e}/6\} \quad \text{(43)}
\]
where (43) follows from Sanov's Theorem and (44) follows from (42) above. □

We are now ready to prove Theorem 1. Lemmas 3 and 4 imply that
\[
\frac{|R_{ph}(x^{(n)}) - R_{part}(x^{(n)})|}{n} \rightarrow 0
\]
in probability. Because rewards are bounded, it follows by the bounded convergence theorem that
\[
\lim_{n \rightarrow \infty} E\left|\frac{R_{ph}(x^{(n)}) - R_{part}(x^{(n)})}{n}\right| = 0. \quad \text{(46)}
\]

**PROOF OF THEOREM 2.** Denote \( \pi = P(X \in A_i) \). Let \( V^*_i(r, Y_n) \) be the expected discounted reward for a type-\( i \) candidate who is offered a kidney that would generate reward \( r \), when there are still \( Y_n \) type-\( i \) candidates unassigned and \( n \) time periods remaining. If this candidate refuses this offer, let \( \rho_i \) denote the probability that his next offer will arrive \( t \) periods later for \( t = 1, \ldots, n \). Then, Bellman's optimality equation can be written as follows:
\[
V^*_i(r, Y_n) = \max \left\{ r, \sum_{t=1}^{n} \rho_t \delta^t E[V^*_i(r, Y_{n-t})] \right\}. \quad \text{(47)}
\]

We will now establish that if (9) holds, then \( V^*_i(r, X_n) = r \) satisfies the dynamic programming recursion (47). We begin by observing that the second term in this recursion satisfies
\[
\sum_{t=1}^{n} \rho_i \delta^t E[V^*_i(r, X_{n-t})] \leq \sum_{t=1}^{n} \rho_t \delta^t E[V^*_i(r, X_{n-t})] \quad \text{(48)}
\]
and
\[
= \sum_{t=1}^{n} \rho_t \delta^t E[R_i(X) | X \in A_t] \quad \text{(49)}
\]

The first inequality in (48) follows from the fact that \( V^*_i(r, Y) \leq V^*_i(r, Y) \) for all \( k \leq n \) because extending the time horizon can only increase rewards. Then, (49) follows from using the trial solution \( V^*_i(r, X_n) = r \). Next, observe that
\[
\rho_i = \left( \frac{t-1}{K-1} \right) \pi^K (1 - \pi)^{t-K} \delta^t E[R_i(X) | X \in A_t] \quad \text{(50)}
\]
for every \( t = 1, \ldots, n \), because the number of periods that would elapse before the candidate receives another offer has a negative binomial distribution with parameters \( K \) and \( \pi \). The last inequality in (52) follows from the incentive compatibility condition (9). Therefore, \( V^*_i(r, X_n) = r \) satisfies the dynamic programming recursion (47), and the candidate's optimal action is to accept all offers.
For subsequent time periods, this same argument holds over the set of sample paths
\[ S_n^{(k)} = \{ Y_{1 \leq k < n}, Y_k \geq k p_n(1 - \epsilon) \}, \]

where \( \epsilon > 0 \) is arbitrary and small. An application of the large deviations principle (see Dembo and Zeitouni 1998) to the sample paths of the random process \( \{ Y_{1 \leq k < n} \} \) shows that as long as \( \Pr(X \in A) \leq p_n \), then \( \Pr\{S_n^{(k)}\} \to 1 \) as \( n \to \infty \). This establishes asymptotic incentive compatibility and concludes the proof. □

### Acknowledgments

The authors have received helpful comments from participants in a research seminar organized by the Operations, Information and Technology (OIT) group in the Graduate School of Business at Stanford University. Kathy Davis provided editorial assistance. Financial support from the National Science Foundation grant SBER-9982446 is gratefully acknowledged. Comments and suggestions from three anonymous referees and an associate editor have substantially improved this paper.

### References


Su, X., S. A. Zenios. 2004b. Mechanism design for kidney allocation. Worker paper, Graduate School of Business, Stanford University, Stanford, CA.


