Competition and Post-Transplant Outcomes in Cadaveric Liver Transplantation under the MELD Scoring System

Harry J. Paarsch, PhD
Amazon.com, Inc.
HPaarsch@Gmail.com

Alberto M. Segre, PhD
Department of Computer Science
University of Iowa
Segre@vinci.cs.UIowa.edu

John P. Roberts, MD
Department of Surgery, Division of Transplant Surgery
University of California at San Francisco
RobertsJ@Surgery.UCSF.edu

Jeffrey B. Halldorson, MD
Department of Surgery, Division of Transplantation
University of Washington
jhalldor@U.Washington.edu

First Draft: 15 September 2010; this Draft: 5 July 2011.

Abstract

Previous researchers have modelled the decision to accept a donor organ for transplantation as a Markov decision problem, the solution to which is often a control-limit optimal policy: accept any organ whose match quality exceeds some health-dependent threshold; otherwise, wait for another. When competing transplant centers vie for the same organs, the decision rule changes relative to no competition; the relative size of competing centers affects the decision rules as well. Using center-specific graft and patient survival-rate data for cadaveric adult livers in the United States, we have found empirical evidence supporting these predictions.

Keywords: liver transplantation; competition; optimal stopping

JEL Classification: C14, I12, L1.

1 Motivation and Introduction

Among some policy analysts, it is almost an article of faith that increasing competition in a market will improve outcomes—efficiency. Of course, whether this will, in fact, obtain depends largely on the characteristics of the particular market. Consider, for example, healthcare. Because operating a major health center involves substantial fixed costs, the number of such institutions in a market is typically small, so issues of efficiency often revolve around volume; see, for example, Luft et al. [1987]; Shahian and Normand [2001] as well as Gaynor et al. [2005]. With relatively small numbers of participants, researchers are particularly concerned with market conduct as well as the effects of mergers; see, for example, Gaynor and Vogt [2003]. In situations with small numbers of decision-makers, particularly in the presence of private information, strategic behaviour can be especially important. A corollary of second-best economics is, then, that increases in competition need not necessarily improve efficiency. Nowhere is this, perhaps, more apparent than in organ transplantation.

In the United States, with the passing of the National Organ Transplant Act in 1984, Congress established the Organ Procurement and Transplantation Network (OPTN), a unified transplant network. Since 1986, OPTN has been been operated under federal contract by a non-profit organization—the United Network for Organ Sharing (UNOS).

While there are around $5,800$ hospitals in the United States, the number of organ transplantion centers is just in the hundreds. To be sure, depending on the type of organ to be transplanted, some variation in the number of transplantation centers exists. But, in the data we use below, there are just 143 liver transplant centers licensed by OPTN/UNOS, of which 121 are for adults and 22 are for children. OPTN/UNOS coordinates how these centers and others that transplant other types obtain organs suitable for transplantation.

Under the rules of the OPTN/UNOS, the United States is sub-divided into eleven regions; see figure 1 for a map outlining which region contains which states as well as the District of Columbia and the Commonwealth of Puerto Rico. Regions are further subdivided into Donation Service Areas; a Donation Service Area is a geographical service area designated by the federal government. Each Donation Service Area is assigned to an Organ Procurement Organization for purposes of recovering organs from all hospitals in that area. Donation Service Area designation and Organ Procurement Organization assignments are made every four years (since 2002) by the Centers for Medicare and Medicaid Services. In any particular Donation Service Area, some Organ Procurement Organizations serve just one transplant center, while others serve several transplant centers.

Within a Donation Service Area, the rules for allocating donor organs differ across the types of organs. The reasons for this surely reflect the benefits and costs of

Figure 1: UNOS Regions

transplanting an organ into the recipient who is next on the waiting list. Basically, the decision of the patient/surgeon concerning whether to accept a donor organ for transplantation is an optimal stopping problem.

On the benefit side, from the perspective of a potential recipient, not all organs are of equal utility: for a given potential recipient, some are better matches than others. That the blood type of the donor is important is uncontroversial. In addition, how long the donor organ has been preserved on ice is also important as *cold ischemia* time affects graft success. In the decision of whether to accept a donor organ today, the match quality of another organ available in the future is unknown. Optimal stopping problems involve a trade-off: in the vernacular of the late eighteenth century, does one accept a bird in hand or wait for the chance of getting the two in the bush?

The primary determinant of the cost of waiting is whether the life of the potential recipient is endangered without transplantation. For some organs, such as kidneys, a potential recipient can be kept alive by artificial means—for example, dialysis. For other organs, such as livers, this is impossible. Not surprisingly, the mechanisms for

allocating donated organs vary by the type of organ.

In the case of kidneys, for example, as the cost of waiting is relatively low, the primary determinant in getting a kidney is match quality. Because the demand for donor kidneys of suitable quality for transplantation greatly exceeds a relatively fixed supply, waiting lists exist. In the case of ties concerning matches of similar quality, time spent on the waiting list determines which potential recipient gets the next available donated kidney. Because dialysis is both painful and inconvenient, many potential recipients try to find living donors, often among close relatives, so they can eliminate the wait. However, even a close relative need not be a good match for some patients. In an attempt to exploit gains from trade and, thus, to increase the number of patients who can be freed from the chains of dialysis, Alvin E. Roth and co-authors (see, for example, Roth et al. [2004]; Roth et al. [2005a]; Roth et al. [2005b]; Roth et al. [2007] as well as Saidman et al. [2006]) have advocated pair-wise and even three-way trades. Such trades are feasible in the market for kidneys because living donors can live healthy and productive lives with just one kidney.

For hearts and lungs (and, for the most part, for livers as well), the supply of donor organs is from the deceased—cadaveric organs. Different allocation mechanisms exist to allocate these scarce resources. In particular, the allocation mechanisms that have been established for hearts, livers, and lungs use medical need as the primary determinant of priority on the waiting list for transplantation.

In this paper, we investigate cadaveric liver transplantation in adults. Liver transplantations in adults are sufficiently different from those in children that, in most locales, separate paediatric tranplantation centers exist. Also, paediatric transplantation centers are relatively rare—perhaps one to every six for adults. Liver transplantations involving living donors are also rare—historically, less than five percent.

Oguzhan Alagoz, Lisa M. Maillart, Andrew J. Schaefer, and Mark S. Roberts [Alagoz et al., 2004, 2007b] have discussed the differences between these sorts of transplantations and cadaveric-liver transplantations.

Below, we argue that, under the current allocation mechanism (known as the MELD scoring system), the presence of competing transplant centers in a Donation Service Area affects the patient/surgeon acceptance decision through a mechanism we refer to as *competitive impatience*: competition makes patient/surgeon decision-makers more likely to accept a donor organ than when no competition exists, which means (all other things being equal) the matches made under competition are predicted to be of weakly lower quality. Assuming match quality affects graft success in a weakly-positive way and holding all other factors constant, the survival function of the waiting time to graft failure after transplantation is then predicted to be weakly greater when no competitors exist than that under competition. When the competing transplant centers in a Donation Service Area perform different numbers of transplantations and have waiting lists of different lengths, such asymmetries are predicted to affect post-transplantation outcomes as well. Using center-specific actual as well as risk-adjusted average graft and patient survival-rate data concerning cadaveric-liver transplantations in adults in the United States, we have found that the predictions of the theory are borne out by the data. We have also found evidence that center-specific heterogeneity is important in explaining post-transplantation outcomes.

The remainder of this paper is divided into five sections. In the next section, we describe in some detail the institutional features of liver transplantation, while in section 3, we present a theoretical framework and then derive some of that framework's empirical implications. In section 4, we describe the data used in the empirical analysis of section 5, while in section 6, we summarize our research and present conclusions.

2 Institutional Details

In the United States, in the case of livers, medical need is the primary determinant of priority on the waiting list for transplantation. Specifically, the Model for End Stage Liver Disease (MELD) disease-severity score has been established as the sole determinant of waiting-list priority.[1] The MELD scoring system was adopted on 2 February 2002. A MELD score is calculated using easily-measurable laboratory values; the MELD score predicts, with some accuracy, a particular potential liver recipient's risk of dying without transplantation. The scale ranges between 6 and 40: a low score (for example, less than 10) predicts minimal risk of death without transplantation, while an high score (for example, above 30) predicts a life-expectancy of only a few months, or less, without transplantation.

Under the MELD scoring system, potential recipients have their names placed on a list at a transplant center. For example, on 31 October 2009, there were approximately $19,000$ potential recipients listed at some transplant center in the United States. Based on the severity of the end-stage liver disease, as measured by the MELD score, potential recipients then wait to see if they are eligible for the next-available donor liver in UNOS.

Competition varies considerably at the local level—the first stage in the allocation process. In Nebraska, for example, where only one Organ Procurement Organization exists, the Nebraska Medical Center (NEUN-TX1) is the only hospital currently performing liver transplantation in adults. NEUN-TX1 maintains a list of potential recipients. These patients await livers from those who will, by accident or stroke

[1]The major exception is in the case of Status 1 patients, who are given first priority. A Status 1A candidate has fulminant liver failure (a rapid, life-threatening loss of liver function) or has recently received a liver transplant that failed shortly afterward. Status 1B candidates are children who have chronic liver disease with severe and life-threatening complications. At any point in time in the United States, however, less than one percent of patients are classified as Status 1A or 1B.

or some other misfortune, die. In Missouri, on the other hand, which also has just one Organ Procurement Organization, two transplant centers exist for adults, both located in the city of Saint Louis: Barnes-Jewish Hospital (MOBH-TX1) and Saint Louis University Hospital (MOSL-TX1). Both MOBH-TX1 and MOSL-TX1 maintain separate lists of potential recipients awaiting liver transplantation.[2] We argue that the decision problem faced by a patient/surgeon in Saint Louis is quite different from the one in Omaha. Below, we explain why competition matters.

Populous states (like California, Illinois, New York, Ohio, Pennsylvania, and Texas) are sub-divided into smaller geographical areas, Donation Service Areas. For example, in California, in 2008, four Donation Service Areas existed—two in the northern part of the state and two in the southern part. The two northern California Donation Service Areas, for example, are the California Donor Network (CADN-OP1) and Golden State (CAGS-OP1). The University of California at Davis Medical Center transplant center (CASM-TX1) is part of CAGS-OP1, while California Pacific Medical Center (CAPM-TX1), the University of California at San Francisco Medical Center (CASF-TX1), and Stanford University Medical Center (CASU-TX1) vie ("compete") for organs in CADN-OP1.

When a donor liver is harvested in a Donation Service Area, the rules for its allocation under UNOS are clear and well-defined. In figure 2, we present a decision tree in which are depicted the potential paths. In that figure, we have inserted

[2]In fact, potential organ recipients are permitted to be on waiting lists at more than one transplant center, without penalty—a practice referred to as *multi-listing*. Multi-listing is believed by some to give wealthy patients an advantage in getting to the head of the queue, so to speak. The geographical distribution of donor organs is not coincident with the demand for organs. For example, because livers have limited viability after harvest (about twelve hours), the *tyranny of distance* rears its ugly head: a liver harvested in Honolulu may not be a viable option for a potential recipient in New York. For this reason, and others, some fraction of harvested livers (David H. Howard [Howard, 2002] has reported over ten percent, while Oguzhan Alagoz, Lisa M. Maillart, Andrew J. Schaefer, and Mark S. Roberts [Alagoz et al., 2007a] have reported six percent) is never used, by anyone.

numbers beneath each node:

1) the liver is offered to Status 1A candidates, either locally (within the Donation Service Area) or in the Region;

2) the liver is offered to Status 1B candidates locally, within the Donation Service Area;

3) the liver is offered locally, within the Donation Service Area, to non-Status 1 candidates having MELD scores greater than or equal to 15, in descending order of MELD score;

4) the liver is offered in the Region to non-Status 1 candidates having MELD scores greater than or equal to 15, in descending order of MELD score;

5) the liver is offered locally, within the Donation Service Area, to non-Status 1 candidates having MELD scores less than 15, in descending order of MELD score;

6) the liver is offered in the Region to non-Status 1 candidates having MELD scores less than 15, in descending order of MELD score;

7) the liver is offered nationally to Status 1A candidates;

8) the liver is offered nationally to Status 1B candidates;

9) the liver is offered nationally to all other candidates, in descending order of MELD score.

In general, the numbers of patients affected by decisions at nodes 1 and 2 are quite small; Alagoz et al. [2007a] have reported under twenty Status 1A candidates

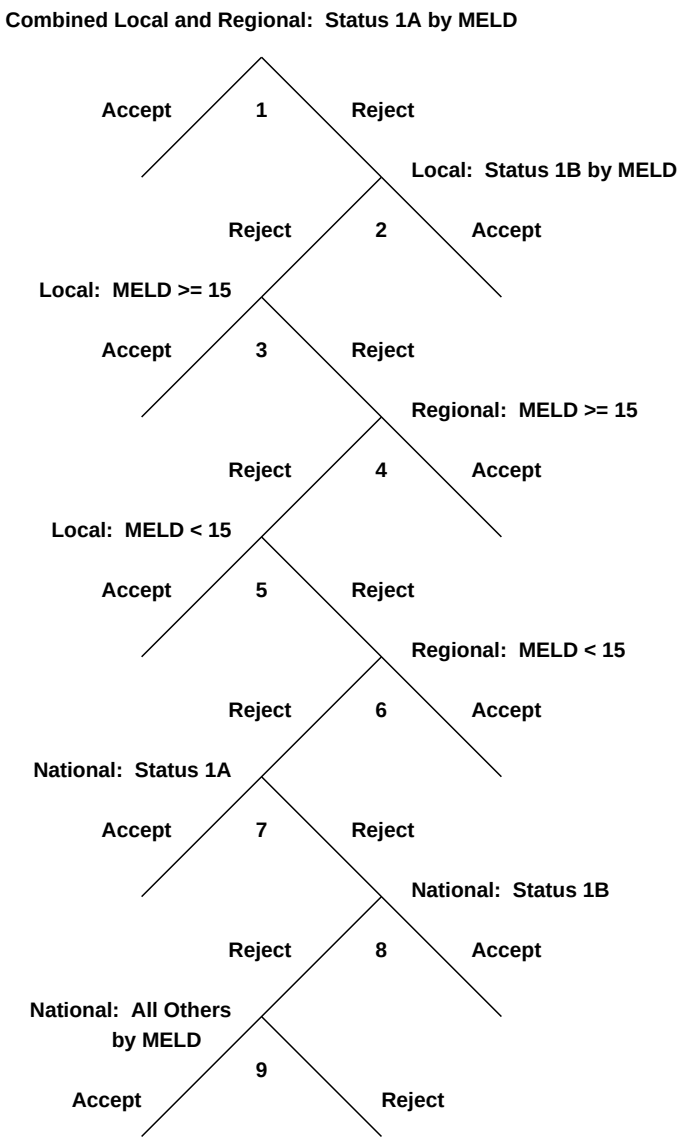

Figure 2: UNOS Rules for Allocating a Donor Liver by Status and Location

in the entire nation at a particular point in time. Moreover, unlike non-Status 1 patients, few Status 1A patients sojourn on the UNOS list very long: Status 1A patients have a life expectancy of fewer than seven days. On the other hand, in any given Region, the numbers of patients affected by decisions at nodes 3, 4, 5, and 6 can be relatively large, perhaps a thousand at any given point in time. But the number of transplant centers served in a Donation Service Area is often very small, sometimes just one or two. In addition, under the MELD scoring system, only patients with the highest scores are relevant to decisions being made. For example, in Region 6 (which contains Alaska, Hawaii, Idaho, Montana, Oregon, and Washington), there are three Donation Service Areas and four transplant centers: HISF-TX1 (Hawaii Medical Center–East), ORUO-TX1 (Oregon Health and Science University) and ORUO-VA1 (Portland Veterans Administration Medical Center, and WAUW-TX1 (University of Washington Medical Center). In short, the number of relevant decision-makers is small.

In situations with small numbers of decision-makers, strategic behaviour can arise. Elsewhere, Dennis P. Scanlon, Christopher S. Hollenbeak, Woolton Lee, Evan Loh, and Peter A. Uber [Scanlon et al., 2004] (who investigated heart transplantations) as well as Jason Snyder [Snyder, in press] (who investigated liver transplantations before the MELD scoring system was implemented) have reported evidence of this. While we do not solve for the equilibrium of a (potentially asymmetric) game of incomplete information, we appeal to game-theoretic notions when interpreting empirically the effects of competition on the post-transplantation outcomes for cadaveric livers in adults in the United States.

Specifically, we argue below that competitive impatience affects the patient/surgeon acceptance decision: competition makes patient/surgeon decision-makers more likely

to accept a donor organ than when no competition exists, which means (all other things being equal) the matches made under competition are predicted to be of weakly lower quality. Assuming match quality affects graft success in a weakly-positive way and holding all other factors constant, the survival function of the waiting time to graft failure after transplantation is then predicted to be weakly greater when no competitors exist than that under competition. When the competing transplant centers in a Donation Service Area perform different numbers of transplantations and have waiting lists of different lengths, such asymmetries are also predicted to affect post-transplantation outcomes as well.

In the next section, we present a theoretical framework and then derive some of that framework's empirical implications.

3 Theoretical Framework

Basically, the decision of whether to accept a donor organ for transplantation is an optimal stopping problem. From the perspective of a potential recipient, not all livers are of equal utility: for a given potential recipient, some livers are better matches than others. That the blood type as well as the age and sex of the donor are important is uncontroversial. In addition, how long the donor liver has been preserved on ice is also important as *cold ischemia* time affects graft success. Whether the donor or the recipient are obese, have diabetes, or have hepatitis C can affect graft success, too. In the decision of whether to accept a donor liver today, the match quality of another liver available in the future is unknown. Optimal stopping problems involve a trade-off: does one accept a bird in hand or wait for the chance of getting the two in the bush? Intuitively, when the chance of getting the two birds in the bush goes

down relative to before, then one is more likely to accept the bird in hand. We should like to make formal this intuition and then to investigate the empirical importance for post-transplant outcomes of competition in the allocation of cadaveric livers for transplantation.

Of course, without transplantation, a potential recipient's health deteriorates over time. In the early stages of end-stage liver disease, the disease typically does not progress as quickly as it does towards the end. Nevertheless, end-stage liver disease is a progressive one which has, in the absence of successful transplantation, a known endpoint—death. Thus, the relative value of transplantation increases as the disease progresses.

In order to put structure on this problem, we initially make the following two assumptions:

Assumption 1. *The match quality of a donor organ can be summarized by a scalar random variable Q having probability density and cumulative distribution functions $f_Q(q)$ and $F_Q(q)$, respectively. Without loss of generality, let $Q \in [0,1]$ where low values of Q are 'bad' matches (ones having low chances of success) and high values are 'good' matches (ones having high chances of success).*

Assumption 2. *The health of a potential recipient can be summarized by a scalar random variable H. Without loss of generality, let $H \in [0,1]$ where a value of zero for H is death, while a value of one is complete health. Given an health state h today, the distribution of tomorrow's health state is weakly worse than today's.*

Given these assumptions, intuition suggests that some strategy involving a reservation quality will be optimal: that is, accept any organ above some threshold quality. It would indeed, however, be surprising if that threshold quality were independent

of health status. In fact, researchers have demonstrated that, under plausible assumptions, a control-limit optimal policy involves thresholds that vary with health status.

The first theoretical model of the patient/surgeon decision of whether to accept an organ for transplantation was developed by Howard [2002], who investigated livers. In the spirit of research pioneered in economics by John Rust [Rust, 1987], Howard constructed a stochastic dynamic-programming model of the decision faced by a patient/surgeon, without considering the effects of competition. Howard put the following structure on the problem: first, he assumed that per-period utility after a successful transplantation is a constant B, while death has utility zero; second, he assumed that the probability of successful transplantation π is a function of patient health h and match quality q where $\pi_h(h,q) > 0$, $\pi_q(h,q) > 0$, as well as $\pi(0,q) = 0$ and $\pi(h,0) = 0$. In words, healthier patients have higher graft-success rates, while better quality matches yield higher graft-success rates. Transplantations in patients near death have very low chances of graft success, while low-quality matches have very low chances of success. Third, Howard assumed that health evolves according to a Markov process. Thus, the conditional probability density function of health in the next period H', only depends on health this period H: health today is a sufficient statistic for a potential recipient's entire health history. Notationally, we write this conditional probability density function as $f_{H'|H}(h'|h)$. Fourth, he assumed a per-period utility while alive of A and a constant discount factor $\rho \in (0,1)$.

Under these assumptions, the expected value of the next-period pay-off to a patient in health state h after transplantation (TX) of an organ having match quality q this period is

$$\mathbb{E}[\Omega^{\mathrm{TX}}(h,q)] = \pi(h,q)B.$$

On the other hand, the expected pay-off from rejecting an organ and continuing to wait (W) is

$$\mathbb{E}[\Omega^{\mathrm{W}}(h,q)] = \int_0^1 \int_0^1 \Omega^{\mathrm{W}}(h',q') f_{H'|H}(h'|h) f_Q(q') \, \mathrm{d}h' \mathrm{d}q'$$

where $\Omega^{\mathrm{W}}(h,q)$ is defined by the fixed point of the following Bellman equation of optimality:

$$\Omega^{\mathrm{W}}(h,q) = A + \rho \max\left\{\mathbb{E}[\Omega^{\mathrm{TX}}(h,q)], \mathbb{E}[\Omega^{\mathrm{W}}(h,q)]\right\}.$$

Howard argued that the Accept and Reject regions of the state space (H,Q) should look something like the regions depicted in figure 3. Specifically, in the early stages of end-stage liver disease, when health h is relatively high, a potential recipient has an high threshold match quality, but as the disease progresses and the potential recipient's health worsens, the threshold quality declines. The exact position of the threshold line going from the southwest to the northeast will depend on a variety of factors—the discount factor, per-period utility, the distribution of match qualities, and the dynamics of the end-stage liver disease progression. In figure 3, we have imposed the intuitively-reasonable condition that some one in near-perfect health (H near one) would not decide to undergo liver transplantation, regardless of the quality of the match.

Howard lacked the appropriate data to implement his model, so he relied on certain comparative static properties of the conjectured solution to interpret aggregate data concerning liver transplantation before the MELD scoring system was introduced. Broadly speaking, he found that his model did reasonably well at explaining the relevant and salient features of the aggregate data.

Alagoz et al. [2007a] provided the necessary mathematical analysis of the structure

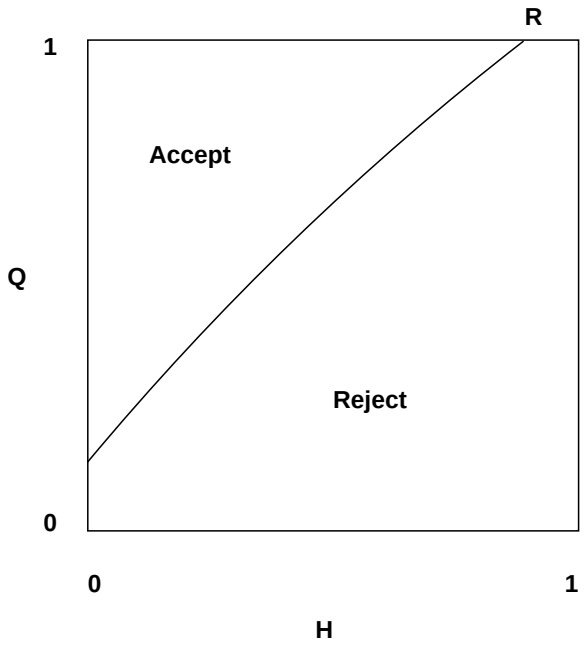

Figure 3: Accept and Reject Regions in Howard [2002] Model

of the problem in Howard [2002], demonstrating that, under plausible assumptions, the optimal solution to the Markov decision problem has a control-limit optimal policy: for a patient of a given health status, accept any organ whose match quality is above some threshold, which typically depends on the patient's health status; otherwise, wait for another organ to arrive. Using clinical data, Alagoz et al. also provided numerical solutions to the Markov decision problem as well as examples where the optimal policy exhibited properties other than those depicted in figure 3. For example, in figure 4, we depict an acceptance region in which the sickest potential recipients require better matches than some of the healthier ones, hence the U-shaped threshold locus in that figure. In fact, it is quite probable that for near-death patients even a liver of the highest match-quality would be of no utility: death would occur nonetheless. In any case, in their series of papers applied to organ transplantation, Alagoz et al. [2004, 2007a,b] demonstrated the practical value of this sort of modelling

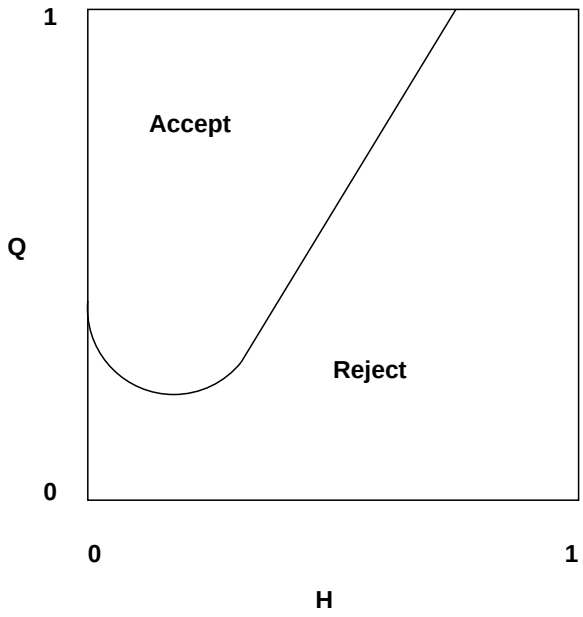

Figure 4: Alternative Accept and Reject Regions

approach.

Howard implicitly assumed that all surgeons are equally informed about the match quality of a donor organ and that each surgeon is only concerned about match quality for her/his patient. While Howard did not consider the effects of local competition in Howard [2002] or Howard [2001a,b], it is clear by statements in those papers that he was aware of potential effects. Alagoz et al. [2007a, page 25] explicitly ruled out strategic behaviour, stating

> we assume that the decisions made by the patient[/surgeon] do not affect the policies of the other patients.

3.1 Analysis of the Transplantation Behaviour

In this subsection, we deduce the implications of competition on behaviour. At the heart of our analysis is the question of whether local competition affects the decision

of a patient/surgeon to accept a donor liver for transplantation. Consider first a single transplant center having a waiting list having D potential recipients: in a particular period, if a donor liver arrives, then the potential recipient with the highest MELD score at the transplant center has the right of first refusal. When the donor liver is inspected, a match quality q is determined. At this point, the trade-off involves deciding whether a future draw from the urn of donors will produce a better expected match than the one currently in hand.

Now introduce a second transplant center in the Donation Service Area, and give this transplant center every second potential recipients from the first transplant center's ordered list. Under these conditions, the Donation Service Area has the same expected aggregate supply of donor organs as in the single transplant center case, so issues involving the relative supply of donor organs are unimportant. Note, too, that the demographics in both the single transplant center and the two transplant center cases are identical. Under UNOS, the allocation of livers between the two transplant centers is again done according to MELD score: the potential recipient in the Donation Service Area with the highest MELD score has the right of first refusal. When a patient/surgeon with the highest MELD score is making an acceptance decision, the trade-off involves not just deciding whether a future draw from the urn of donors can be expected to be a better match than the one currently in hand, but also whether the current highest MELD-score patient will have the right of first refusal in a future draw from the urn. End-stage liver disease is a progressive disease, but the progression is stochastic rather than deterministic: the potential recipient who currently has the highest MELD score could be easily eclipsed by another patient at the other transplant center in the Donation Service Area, or a new Status 1 patient in either the Donation Service Area or the Region. In short, holding the expected supply of

donor livers constant, waiting for another liver is more risky when competitors exist than when no competitors exist. Consequently, all things being equal, the minimum threshold quality for transplantation, at any health status, weakly decreases. In terms of figure 3, under the assumptions made, the curve rotates to the southeast about the point R. (It seems unlikely that those who did not wish to undergo transplantation would want to do so under competition, but it cannot be ruled out: there may be option value to having transplantation early rather than waiting.)

Competitors are like having an higher discount rate ρ in the problems investigated by Howard [2002] as well as Alagoz et al. [2004, 2007a,b], so we refer to this willingness to accept donor organs more readily under competition than when no competitors exist as *competitive impatience*. We depict the policy functions in figure 5: for points below a reservation quality, which has superscript *, the value of the policy is zero (do not accept the organ), while for points above a reservation quality, the value of the policy function is one (accept the organ). In that figure, the point S denotes the symmetric equilibrium, the intersection of the policy functions at point $q^*_{A,SC}$ and $q^*_{B,SC}$, while $q^*_{A,NC}$ and $q^*_{B,NC}$ denote the reservation qualities in the single transplant center case.

Of course, it is quite possible that competitors in a Donation Service Area are different in terms of *market presence* or *market power*. That is, some transplant centers will do more transplantations than others; some transplant centers have longer waiting lists than others. What does this mean? Well, obviously, this introduces an asymmetry into the transplantation game, which has interesting effects. Consider a Donation Service Area with a "large" transplant center (one which performs relatively many transplantations because it has a large number of its waiting list, in the extreme, say, $D-1$) which we shall refer to as A, and a small transplant center (one which

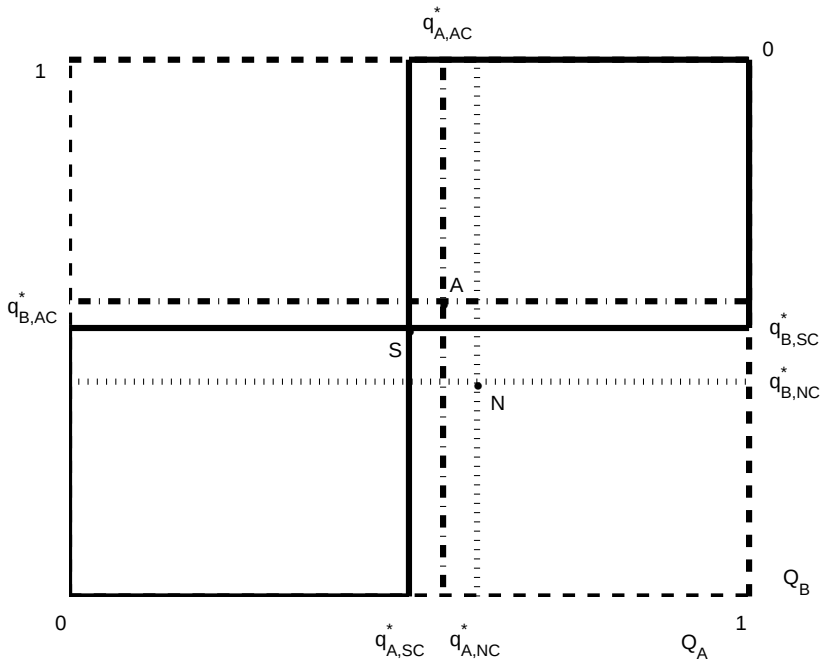

Figure 5: Reservation Quality under Different Amounts of Competition

performs relatively few transplantations because it has a fewer patients on its waiting list, in the extreme, say, one) which we shall refer to as B. All other things being equal, under the MELD scoring system, donor livers are allocated to patients at the top of the waiting list. Because transplant center A's list is larger in absolute size, a representative MELD ranking is more likely to have a sequence of patients from transplant center A at the top of the list than a sequence of patients from transplant center B; i.e.,

$$\mathrm{A}_1 \geq \mathrm{A}_2 \geq \mathrm{B}_1 \geq \mathrm{A}_3 \geq \mathrm{B}_2 \geq \cdots$$

is a more likely sequence than

$$\mathrm{B}_1 \geq \mathrm{B}_2 \geq \mathrm{A}_1 \geq \mathrm{A}_2 \geq \mathrm{A}_3 \cdots$$

Consequently, the reservation minimum threshold quality at transplant center A will be weakly higher than at transplant center B: all other things being equal, competitive impatience will be higher at smaller transplant centers than at larger ones. We depict this in figure 5 where point A denotes the asymmetric equilibrium. In this example, $q^*_{A,AC}$ is above $q^*_{A,SC}$, but below $q^*_{A,NC}$, while $q^*_{B,AC}$ is above both $q^*_{B,SC}$ and $q^*_{B,NC}$.

3.2 Empirical Implications of Theoretical Predictions

The existence of competing transplant centers in a Donation Service Area affects the patient/surgeon decision through the mechanism we have referred to above as competitive impatience: when competing transplant centers exist in a Donation Service Area, this competition makes patient/surgeon decision-makers more likely to accept a donor organ than when no competition exists. Thus, all other things being equal, the matches made under competition are of weakly lower quality than when no competitors exist. Assuming match quality affects graft success in a weakly-positive way and holding all other factors constant, one empirical implication of this is that the survival function of the waiting time to graft failure after transplantation is weakly greater when no competitors exist than that under competition, which means that the average waiting time to graft failure after transplantation is weakly greater when no competitors exist than that under competition.[3] Another empirical implication

[3]For a non-negative random variable T, the integral of its survival function $S_T(t)$, which equals $[1 - F_T(t)]$, where $F_T(t)$ is the cumulative distribution function, is the expected value of the random variable in question, provided the integral exists. Consider $S_1(t)$ and $S_2(t)$, the survival functions of two non-negative random variables T_1 and T_2. If

$$S_1(t) \geq S_2(t) \quad \forall\ t,$$

then

$$\mathbb{E}(T_1) = \int_0^\infty S_1(t)\ \mathrm{d}t \geq \int_0^\infty S_2(t)\ \mathrm{d}t = \mathbb{E}(T_2).$$

is this that, holding all other factors constant, the survival function of the waiting time to graft failure after transplantation is weakly less for small competitors than for large competitors, which means that the average waiting time to graft failure after transplantation is weakly shorter for small competitors than for large competitors. In section 5, we shall put specific structure on how we shall confront these predictions with data.

4 Center-Specific Graft and Patient Survival-Rate as well as Market-Specific Data

We obtained the publicly-available data used in this research from the Arbor Research Collaborative for Health.[4] Specifically, we used information from the SAS datasets that were used to produce the *Program Specific Reports*; these were provided to us by Andrew Barnes and Craig Lake of the Arbor Research Collaborative for Health. The construction of the data is described in detail in the "Guide to the Program-Specific Reports v 12.0, December 2009," which is published by the Arbor Research Collaborative for Health.

In any period, the four variables of interest to us are the actual, center-specific graft and patient survival-rates (`Actual_GSR` and `Actual_PSR`) and the center-specific, risk-adjusted average graft and patient survival-rates (`Predicted_GSR` and `Predicted_PSR`) which were constructed by the staff at Arbor at one month, one year, and three years after transplantation using the methods described in the "Guide to the Program-Specific Reports v 12.0, December 2009," pages 20–36. To provide some idea concerning how `Actual_GSR` and `Actual_PSR` as well as `Predicted_GSR` and `Predicted_PSR`

[4]Address: 315 West Huron Street, Suite 360, Ann Arbor, MI 48103.

were constructed, we begin by describing how the data are collected by UNOS.

Prior to the creation of each report, data are collected over the previous thirty-month observation period. For example, if the begin date of an observation period were 1 July 2002, then the end date of that observation period would be 31 December 2004. We depict in figure 6 this initial observation period as the interval between zero and thirty months. During this thirty-month period, data concerning all transplantations are gathered from each center in all eleven UNOS regions of the nation.

The most important data for our purposes are the graft and patient survival durations. Representative survival durations are depicted in figure 6 by the horizontal lines with "[" and "]" at the endpoints. These lines and endpoints represent completed spells—viz., grafts that have failed or patients who have died. Note that spells often go beyond the observation periods; that is, some grafts have not failed or some patients are still alive when data collection for that observation period ends, so these spells are censored for this particular observation period. In figure 6, these are depicted by two of the spells in the top right of the figure—those which began before month thirty, but ended after month thirty.

To describe the calculations that then take place, let us first introduce some additional notation. In any observation period indexed by $\ell = 1, 2, \ldots, L$, we index by $i = 1, 2, \ldots I_j^\ell$ the patients who received transplantation at the centers indexed by $j = 1, 2, \ldots, J^\ell$. We denote by t_{ij}^ℓ the survival duration of the graft (or the patient) for patient i at center j in observation period ℓ. Some of the t_{ij}^ℓs may be censored: to wit, the graft has not yet failed or the patient has not yet died when observation period ℓ ended, such as the two depicted in figure 6. We denote by $\boldsymbol{z}_{ij}^\ell$ a p-vector of covariates observed in period ℓ, concerning patient i at center j; these will be important when

we describe how `Predicted_GSR` and `Predicted_PSR` are calculated.

Using the product-limit estimator of the survival function proposed by Edwin L. Kaplan and Paul Meier [Kaplan and Meier, 1958], the `Actual_GSR` and `Actual_PSR` data are calculated for each center at one month, one year, and three years. Note that only center-specific data are used in each graft survival-rate and patient survival-rate calculation.

In order to calculate the `Predicted_GSR` and `Predicted_PSR` data, all of the data from all centers in the nation are used. Specifically, an empirical specification proposed by Sir David R. Cox [Cox, 1972], now commonly referred to as a *Cox proportional hazard rate* model, is used to control for the patient-specific covariates $\boldsymbol{z}$. In a Cox proportional hazard rate model, the p-vector $\boldsymbol{z}$ influences $\lambda(t|\boldsymbol{z})$, the conditional post-transplantation hazard rate of the waiting time to graft failure (or patient death) T, according to the following structure:

$$\lambda(t|\boldsymbol{z}) = \exp(\boldsymbol{z}\boldsymbol{\gamma})\lambda_0(t)$$

where $\boldsymbol{\gamma}$ is an unknown parameter vector conformable to $\boldsymbol{z}$ and $\lambda_0(t)$ is the baseline hazard rate of a patient having $\boldsymbol{z}$ equal to a p-vector of zeros $\mathbf{0}_p$. Both $\boldsymbol{\gamma}$ and $\lambda_0(t)$ must be estimated using $\left\{\{(t_{ij}^{\ell}, \boldsymbol{z}_{ij}^{\ell})\}_{i=1}^{I_j^{\ell}}\right\}_{j=1}^{J_\ell}$, data from all patients and all centers in the nation during the the ℓ^{th} observation period.

Now, the baseline survival function $S_0(t)$ is related to the baseline hazard rate $\lambda_0(t)$ according to

$$S_0(t) = \exp\left[-\int_0^t \lambda_0(y)\ \mathrm{d}y\right],$$

so the conditional survival function can be written as

$$S_{T|\boldsymbol{Z}}(t|\boldsymbol{z}) = \exp\left[-\exp(\boldsymbol{z\gamma})\int_0^t \lambda_0(y)\ \mathrm{d}y\right] = \exp\left[-\int_0^t \lambda_0(y)\ \mathrm{d}y\right]^{\exp(\boldsymbol{z\gamma})} = S_0(t)^{\exp(\boldsymbol{z\gamma})}.$$

Having estimated $\boldsymbol{\gamma}$ and $\lambda_0(t)$, one can then estimate the conditional survival function for each patient i at each center j in observation period ℓ at any duration t. We denote this as

$$\hat{S}_{T|\boldsymbol{Z}}(t|\boldsymbol{z}_{ij}^{\ell}) = \exp\left[-\int_0^t \hat{\lambda}_0^{\ell}(y)\ \mathrm{d}y\right]^{\exp(\boldsymbol{z}_{ij}^{\ell}\hat{\boldsymbol{\gamma}})} = \hat{S}_0(t)^{\exp(\boldsymbol{z}_{ij}^{\ell}\hat{\boldsymbol{\gamma}})}.$$

Thus, for center j in observation period ℓ at any duration t, the average adjusted survival rate for either graft or patient (depending on the dependent variable) is constructed by

$$\texttt{Predicted\_X}_j^{\ell} = \frac{1}{I_j^{\ell}}\sum_{i=1}^{I_j} \hat{S}_{T|\boldsymbol{Z}}(t|\boldsymbol{z}_{ij}^{\ell}) \quad \texttt{X} = \texttt{GSR},\ \texttt{PSR}.$$

The next report is created using information from an overlapping sample. Specifically, data from the first six months of the previous observation period are dropped, and data from the next adjacent six-month period are added. In figure 6, the next observation period would be between month six and month thirty-six, year three. Continuing with our initial example, the begin date would then be 1 January 2003, while the end date would be 30 June 2005.

As the MELD began on 1 February 2002, our first begin date is 1 July 2002. We depict this in figure 7. What to do about three-year survival rates? Well, these must begin thirty months plus two years after the MELD began, so for reporting date 1 January 2007 onward. The last report date we have is 1 January 2010.

For each transplant center (`center`) in each Donation Service Area (`dsa`), we

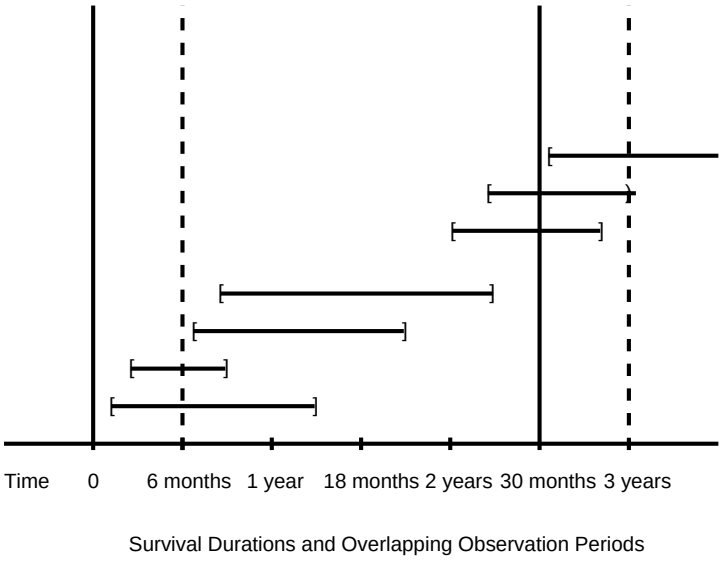

Figure 6: Timeline of Data Collection Periods

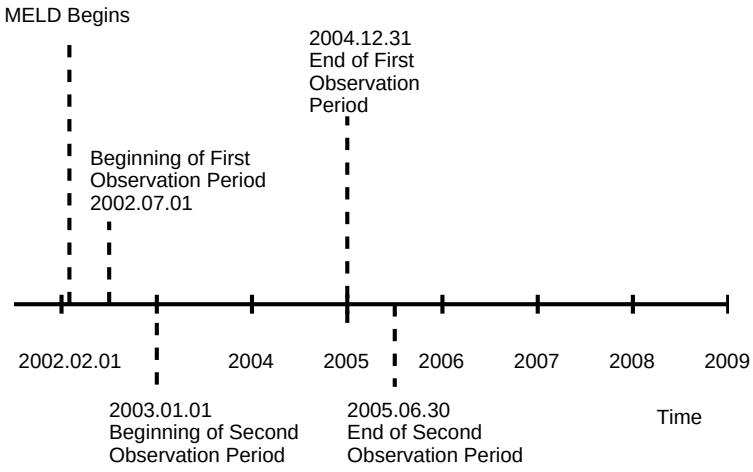

Figure 7: Post-MELD Data Gathering by UNOS

constructed a unique center/dsa-specific identification number (dcode), which was then used to construct some of the covariates described below. Covariates of interest to us are the number of transplantations performed at that transplant center during the observation period (ntxs) as well as the number of transplant centers competing in the Donation Service Area (ntxcs) and the total number of transplantations performed in the Donation Service Area (tot_txs) during the observation period. Using the ntxcs in the Donation Service Area as well as the proportion of transplantations performed by a transplant center in that Donation Service Area, we constructed the following covariate:

$$\texttt{share} = \log\left(\frac{\frac{\texttt{ntxs}}{\texttt{tot\_txs}}}{\frac{1}{\texttt{ntxcs}}}\right)$$

which is a summary statistic for the dispersion in market share. When the proportion of transplantations performed at a transplant center in a Donation Service Area is the same as the inverse of ntxcs, the covariate share equals zero, so the covariate has no effect. (This would be what the covariate would be under equal market shares for each transplant center in a Donation Service Area.) When the proportion of transplantations performed at a transplant center in a Donation Service Area is greater than the inverse of ntxcs, the covariate share is positive, while when the proportion of transplantations performed at a transplant center in a Donation Service Area is less than the inverse of ntxcs, the covariate share is negative.

With some entry and exit in this industry, constructing unequivocal measures of competition in any Donation Service Area during an observation period is difficult. For example, suppose that in 1 July 2002 there were two centers in a Donation Service Area, but on 30 November 2002 another entered, while on 14 November 2004 one of the first two centers exited the market, and on 21 February 2005 a third (fourth, depending on ones perspective) entered. What is the appropriate number of

centers to use? Provided at least one transplant were reported at each center during the observation period, we would report this as three for both observations because, during each observation period, three centers would have been active in that Donation Service Area.

In some states (for example, New York, Ohio, and Tennessee), sharing agreements apply statewide, so Donation Service Area boundaries are irrelevant. In our final empirical specification, we dealt with these by using the total number of centers in the state as a measure of competition.

In table 1, we present descriptive statistics for our sample which contained $2,322$ observations concerning 105 transplant centers in 53 Donation Service Areas in the eleven UNOS regions—870 each concerning one-month and one-year durations and 582 concerning three-year durations.[5] In the first three rows, we present summary statistics concerning the actual center-specific graft survival rates, while the next three concern the center-specific, risk-adjusted average graft survival rates that were constructed by the Arbor Research Collaborative for Health. In the next three rows, we present summary statistics concerning the actual center-specific patient survival rates, while the next three concern the center-specific, risk-adjusted average patient survival rates, again constructed by the Arbor Research Collaborative for Health. The next four rows concern the market-specific data—specifically, the number of transplantations performed at a transplant center, the total number of transplantations performed in a Donation Service Area, the number of transplant centers in a Donation Service Area, and the `share` variable defined above. In order to be flexible concerning the effects of the covariate `ntxcs`, we also constructed indicator variables for Donation Service Areas having just one transplant center (`one`), two transplant centers

[5]Fewer observations concerning three-year durations exist because fewer observation periods involving only post-MELD transplantations exist.

(`two`), three transplant centers (`three`), four transplant centers (`four`), and five or more transplant centers (`five+`). Thus, the final five rows concern dummy variables describing the number of competitors in the various Donation Service Areas.

Around one fifth of all Donation Service Areas in the sample have just one transplant center, while around another one quarter of the Donation Service Areas have five or more transplant centers.[6] About one third of the Donation Service Areas have just two transplant centers, which is also the median for this sample, while around ten percent have three or four Donation Service Areas each. In this sample, around 129 transplantations are performed at an average transplant center over the thirty month recording period used to calculate the center-specific actual graft survival rates and patient survival rates as well as risk-adjusted average graft survival rates and patient survival rates (around one per week), but there is considerable dispersion. For example, the sample minimum is one, while the sample maximum is 575, which translates into around four transplantations per week. The median `share` variable in the sample is zero, while the minimum is -4.1163 and the maximum is 1.3863.

5 Empirical Results

We investigated empirically the following question: does the presence of local competition affect the post-transplantation outcomes? The theoretical structure outlined above suggests that the control-limit quality Q^* will depend on patient-specific characteristics, which we collect in the vector $\boldsymbol{z}$, as well as center- and market-specific characteristics, which we collect in the vector $\boldsymbol{m}$. We hypothesize that the un-

[6] This does not mean that, at any point in time, one fifth of all Donation Service Areas have just one transplant center. Rather, that for one fifth of our data the dummy variable `one` is one. Note that our sample was gathered over a six-year period, during which some entry and exit have occurred in these Donation Service Areas: evidence the exit of CASM-TX1 mentioned in section 2.

Table 1: Sample Descriptive Statistics

Variable	Mean	St.Dev.	Median	Minimum	Maximum
Actual_GSR, one-month	0.9325	0.0471	0.9410	0.6360	1
Actual_GSR, one-year	0.8291	0.0685	0.8380	0.5000	1
Actual_GSR, three-year	0.7301	0.0869	0.7335	0.2500	1
Predicted_GSR, one-month	0.9370	0.0128	0.9380	0.8710	0.9670
Predicted_GSR, one-year	0.8324	0.0287	0.8350	0.7030	0.9030
Predicted_GSR, three-year	0.7319	0.0365	0.7340	0.6120	0.8640
Actual_PSR, one-month	0.9489	0.0461	0.9590	0.5000	1
Actual_PSR, one-year	0.8589	0.0657	0.8680	0.3330	1
Actual_PSR, three-year	0.7729	0.0804	0.7790	0.2500	1
Predicted_PSR, one-month	0.9573	0.0121	0.9590	0.8200	0.9810
Predicted_PSR, one-year	0.8687	0.0259	0.8720	0.6580	0.9290
Predicted_PSR, three-year	0.7809	0.0327	0.7850	0.4790	0.9160
ntxcs	2.9630	1.7143	2	1	6
ntxs	128.5310	98.6378	103	1	575
tot_txs	327.7196	204.2987	279	1	861
share	−0.1137	0.7149	0	−4.1163	1.3863
one	0.1960	0.3970	0	0	1
two	0.3695	0.4828	0	0	1
three	0.0930	0.2905	0	0	1
four	0.1064	0.3084	0	0	1
five+	0.2351	0.4242	0	0	1

observed control-limit Q^* affects post-transplantation graft survival duration in a weakly-positive way according to the following function:

$$T = \tau\left[Q^*(\boldsymbol{z}, \boldsymbol{m})\right] \quad \tau'(q) \geq 0.$$

Thus, the distribution of Q^*, conditional on $\boldsymbol{z}$ and $\boldsymbol{m}$, induces a distribution of T, conditional on $\boldsymbol{z}$ and $\boldsymbol{m}$.

We have chosen to confront the predictions of our theoretical framework using the data described in the previous section employing an empirical specification inspired by the research of Cox [1972] (which concerned hazard rates), but applied to survival functions. We cannot, of course, implement Cox's approach because we do not have micro-level data, only center-specific averages. We should note, however, that the center-specific, risk-adjusted average graft survival-rate and patient survival-rate data are generated using results from a Cox proportional hazard-rate model.

Within our empirical framework, the observed covariates, which are collected in the vectors $\boldsymbol{z}$ and $\boldsymbol{m}$, as well as a single dimension of unobserved heterogeneity $\hat{U}$ (or $\hat{V}$), are assumed to influence the `Actual_GSR` (or `Actual_PSR`) post-transplantation graft (or patient) survival-rate according to the following equations:

$$\texttt{Actual\_GSR}(t) = \texttt{Predicted\_GSR}(t) \times \exp(\boldsymbol{m}\boldsymbol{\delta}_G + \hat{U})$$

or

$$\texttt{Actual\_PSR}(t) = \texttt{Predicted\_PSR}(t) \times \exp(\boldsymbol{m}\boldsymbol{\delta}_P + \hat{V})$$

where $\boldsymbol{\delta}_G$ and $\boldsymbol{\delta}_P$ are unknown parameter vectors conformable to $\boldsymbol{m}$. Within this framework, the patient-specific covariates $\boldsymbol{z}$ are embedded in `Predicted_GSR` (and

Predicted_PSR) as a result of the conditioning performed using the Cox proportional hazard-rate model.

Now,

$$\log\left[\texttt{Actual\_GSR}(1)\right] - \log\left[\texttt{Predicted\_GSR}(1)\right] = \boldsymbol{m}\boldsymbol{\delta}_G + \hat{U}_1$$
$$\log\left[\texttt{Actual\_GSR}(12)\right] - \log\left[\texttt{Predicted\_GSR}(12)\right] = \boldsymbol{m}\boldsymbol{\delta}_G + \hat{U}_{12}$$
$$\log\left[\texttt{Actual\_GSR}(36)\right] - \log\left[\texttt{Predicted\_GSR}(36)\right] = \boldsymbol{m}\boldsymbol{\delta}_G + \hat{U}_{36}$$

which we write, for each transplant center indexed by j in each Donation Service Area indexed by k, as

$$Y_{jk,1} = \boldsymbol{m}_{jk}\boldsymbol{\delta}_G + U_{jk,1}$$
$$Y_{jk,12} = \boldsymbol{m}_{jk}\boldsymbol{\delta}_G + U_{jk,12}$$
$$Y_{jk,36} = \boldsymbol{m}_{jk}\boldsymbol{\delta}_G + U_{jk,36}$$

where $Y_{jk,t}$ equals $(\log[\texttt{Actual\_GSR}_{jk}(t)] - \log[\texttt{Predicted\_GSR}_{jk}(t)])$ and where U differs from $\hat{U}$ by measurement or reporting errors, which are probably small.[7] Introducing the index over all observations of duration t in all reporting periods $n =$

[7]To save space, we do not write-out the equations for Actual_PSR and Predicted_PSR as the reasoning is straightforward.

$1, 2, \ldots, N_t$, we can write this in matrix notation as

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_{12} \\ \boldsymbol{Y}_{36} \end{bmatrix} = \begin{bmatrix} \boldsymbol{M}_{N_1} \\ \boldsymbol{M}_{N_{12}} \\ \boldsymbol{M}_{N_{36}} \end{bmatrix} \boldsymbol{\delta}_G + \begin{bmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_{12} \\ \boldsymbol{U}_{36} \end{bmatrix} = \boldsymbol{X}\boldsymbol{\delta}_G + \boldsymbol{U} \tag{1}$$

where the matrix $\boldsymbol{M}_{N_t}$ collects the vectors of center- and market-specific covariates. The parameters of the measurement equation (1) can estimated a number of ways—e.g., by the method of least squares (LS), or by the method of least absolute deviations (LAD).

The overlapping nature of the samples introduces potential dependence among some of the error terms in the empirical specifications. This dependence could affect the standard errors of the estimates and, hence, influence inference. Thus, we used estimators of the least-squares standard errors that are robust to arbitrary forms heteroskedasticity as well as some forms of autocorrelation. Such so-called "HAC estimators" were first proposed by Whitney K. Newey and Kenneth D. West [Newey and West, 1987] and then developed further by Donald K. Andrews [Andrews, 1991] as well as Andrews and J. Christoper Monahan [Andrews and Monahan, 1992]; this work built on the research of Friedhelm Eicker [Eicker, 1963] and Halbert L. White [White, 1980]. The computation of these standard errors is discussed in detail by Achim Zeileis [Zeileis, 2004].

In table 2, we present the LS and LAD estimates of the parameters of interest.[8] Beneath each estimate, in parentheses, is reported the robust standard error. In the row denoted "`SSR`," is reported the sum of squared residuals. The LAD estimates,

[8]Here, we have also included as covariates one-month, one-year, and three-year dummy variables (which are denoted `d01`, `d12`, and `d36`) to ensure that no residual duration dependence exists in the data; the results are qualitatively similar when these covariates are excluded.

Table 2: Estimates for Equation (1): Dependent Variable Y, Graft Survival Rate

Covariate	LS	LAD
d01	0.00430	−0.00001
	(0.00606)	(0.00314)
d12	0.00337	0.00112
	(0.00628)	(0.00314)
d36	0.00136	0.00125
	(0.00560)	(0.00331)
ntxs	−0.00002	0.00001
	(0.00002)	(0.00001)
ntxcs	−0.00195	0.00020
	(0.00119)	(0.00060)
share	0.02058	0.00366
	(0.00644)	(0.00178)
SSR	14.94720	15.45540

which correspond to the regression equivalent of the median, are robust to contamination.

In general, what can one conclude from the inclusion of the market covariates into equation (1)? First, contrary to what has been reported previously in the literature (see, for example, the research by Erick B. Edwards, John P. Roberts, Maureen A. McBride, James A. Schulak, and Lawrence G. Hunsicker [Edwards et al., 1999]), we have found that the number of transplantations performed at a transplant center has no important effect on average graft survival rates; the estimated coefficient on ntxs, while negative, is statistically insignificant—a p-value of 0.38. Second, when estimated by the method of LS, the number of transplant centers in the Donation Service Area, ntxcs, is calculated to have a negative effect on the post-transplantation graft survival rate, as the theory above predicted: each additional transplant center in a Donation Service Area is predicted to reduce the average survival-rate by about one-quarter of one percent, but the p-value is only 0.102. On the other hand, when estimated by the method of LAD, the number of transplant centers in the

Donation Service Area, `ntxcs`, is predicted to have no significant effect on the post-transplantation graft survival rates; its p-value is 0.74. Third, as predicted above, an above "average" market `share` translates into an increase in the post-transplantation graft survival rate. For example, consider a Donation Service Area in which two transplant centers compete, where one transplant center performs two-thirds of the transplantations, while the other performs one third: the "larger" transplant center has an average graft survival rate that is around 2.85 percent greater than the "smaller" one; the LAD estimate is about half that, but both have p-values less than 0.04. That Edwards et al. [1999] did not include measures of competition may explain why they found a positive relationship between post-transplant outcomes and the number of transplants performed: the distribution of transplantations among transplant centers was ignored.

In table 3, we report next the LS and LAD estimates for the dependent variable Y where we have also included center-specific dummy variables. In order to implement the empirical specification, we have had to alter it slightly, without affecting its empirical content. Specifically, we substituted a constant for the `d01` dummy variable. Also, in order to relax the linearity constraint concerning the effect of the number of competing transplant centers in a Donation Service Area, as noted in section 3, we introduced dummy variables to represent varying levels of competition; the omitted dummy variable is the one-transplant center category, `one`. While there are 105 transplant centers in our sample, only 104 dummy variables were included in the empirical specifications; the omitted center was WAUW-TX1, a transplant center with no other competitors in its Donation Service Area.

What do the estimates in table 3 tell us? In short, the number of transplantations performed at a transplant center is predicted to decrease the average graft survival

Table 3: LS and LAD Estimates of Equation (1) Including Center-Specific Dummy Variables, Graft Survival Rate

Variable	LS	LAD
Constant	0.07431	0.04643
	(0.02005)	(0.01276)
d12	−0.00092	0.00320
	(0.00254)	(0.00181)
d36	−0.00442	0.00750
	(0.00425)	(0.00205)
ntxs	−0.00022	−0.00015
	(0.00007)	(0.00003)
two	−0.03109	−0.02301
	(0.01660)	(0.01517)
three	−0.07028	−0.04742
	(0.02743)	(0.01977)
four	−0.08418	−0.05815
	(0.03387)	(0.02115)
five+	−0.09331	−0.06634
	(0.03791)	(0.02234)
share	0.04754	0.02100
	(0.01918)	(0.00370)
SSR	9.55863	10.92940

rate for each additional transplant; its p-value is 0.002. Additional competition is predicted to decrease the average graft survival rate by between 2.3 and 9.3 percent, depending on the number of transplant centers in the Donation Service Area and the estimation method. The share of transplantations performed by a transplant center in a Donation Service Area is predicted to increase the average graft survival rate; its p-value is 0.001 under least squares. For example, again consider a Donation Service Area in which two transplant centers compete, where one transplant center performs two-thirds of the transplants, while the other performs one third: the "larger" transplant center has an average graft survival rate that is around 6.6 percent greater than the "smaller" one: bigger is better in this empirical specification.

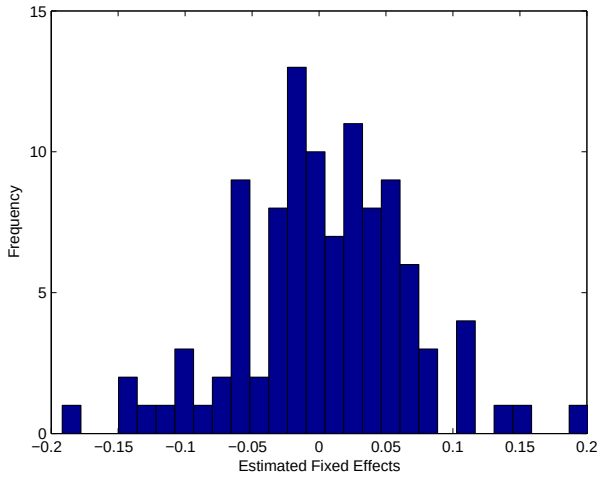

Figure 8: Histogram of LS Center-Specific Estimated Effects

We do not think it either fruitful or illuminating to report all 104 estimated coefficients on the dummy variables. However, in figure 8, we depict the histogram of estimated coefficients, while in figure 9, we depict the kernel-smoothed estimated density of the percentage differences across transplant centers in our sample. The striking feature of this latter figure is the extreme variability in center-specific predicted outcomes. The range is between -20 percent and $+20$ percent. That is, having controlled for patient-specific characteristics, different transplant centers can differ by forty percent in the post-transplantation outcomes as measured by the average graft survival rate. (To make sure that this extreme range is not an artifact of the method of estimation, we also investigated the center-specific estimates generated by LAD; they were qualitatively the similar, but the range is slight smaller, only thirty percent, and are depicted in figures 10 and 11.)

What is causing this extreme variability? Because patient-specific characteristics have been used to construct the risk-adjusted average graft survival rates, the estimates of coefficients corresponding to the center-specific dummy variables would

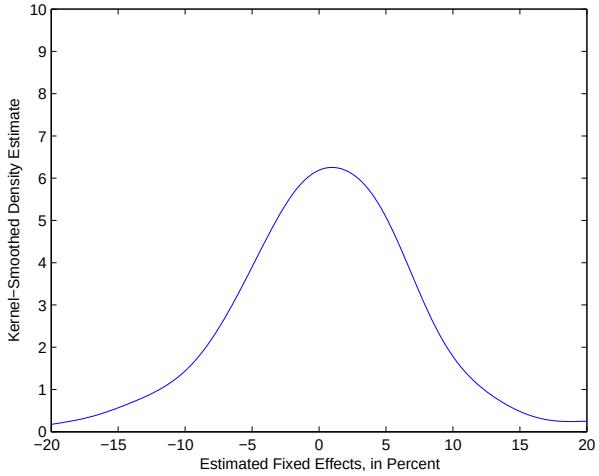

Figure 9: Kernel-Smoothed, LS Center-Specific Estimated Effects

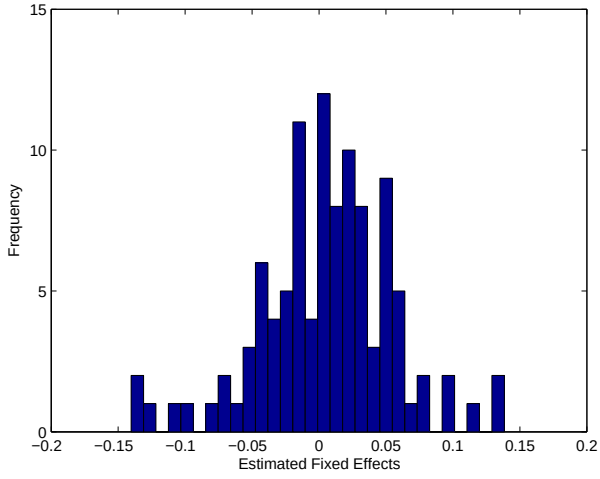

Figure 10: Histogram of LAD Center-Specific Estimated Effects

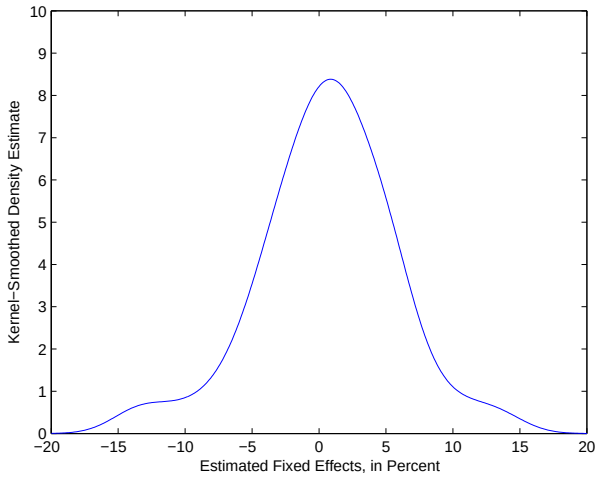

Figure 11: Kernel-Smoothed, LAD Center-Specific Estimated Effects

appear to be capturing center-specific differences in graft success outcomes. Why are these differences obtaining? Is it because of differences in the culture of surgery at different transplant centers (viz., increased risks are taken at some transplant centers, and not at others) or because the quality of surgeons differs across transplant centers (viz., some transplant centers attract the very best, while others can only attract those who just barely passed the Board examinations)? Unfortunately, the average data to which we have had access do not permit us to address such important questions.

We next proceeded to analyse the effects of competition on patient survival rates, the `Actual_PSR` and `Predicted_PSR` data. In table 4, we present the LS and LAD estimates of the parameters of interest.

In general, what can one conclude from the inclusion of the market covariates? First, again, contrary to what has been reported previously in the literature, we have found that the number of transplantations performed at a transplant center has no important effect on average patient survival rates; the LS estimated coefficient on

Table 4: Estimates for Equation (1): Dependent Variable Y, Patient Survival Rate

Covariate	LS	LAD
d01	−0.00088	−0.00653
	(0.00596)	(0.00283)
d12	−0.00181	−0.00552
	(0.00612)	(0.00283)
d36	−0.00647	−0.00590
	(0.00517)	(0.00297)
ntxs	0.00003	0.00007
	(0.00002)	(0.00001)
ntxcs	−0.00316	−0.00140
	(0.00116)	(0.00053)
share	0.02268	0.00400
	(0.00651)	(0.00159)
SSR	13.20820	13.71610

ntxs, while positive, is small 0.00003, one-third of one basis point, and statistically insignificant. Second, the number of transplant centers in the Donation Service Area, ntxcs, is estimated to have a significantly negative effect on the post-transplantation patient survival rate, as the theory above predicted: each additional transplant center in a Donation Service Area is predicted to reduce the average patient survival rate by between one-tenth and three-tenths of one percent, with a p-value of less than 0.009. Third, as predicted above, an above "average" market share translates into an increase in the post-transplantation patient survival rate.

In table 5, we report next the LS and LAD estimates when we also included center-specific dummy variables. Again, in order to implement the empirical specification, we have had to alter it slightly, without affecting its empirical content. Specifically, we substituted a constant for the d01 dummy variable. Also, in order to relax the linearity constraint concerning the effect of the number of competing transplant centers in a Donation Service Area, we introduced dummy variables to represent varying levels of competition; the omitted dummy variable is the one-transplant center category, one.

Only 104 dummy variables were included in the empirical specifications; the omitted center was again WAUW-TX1, a TXC with no other competitors in its Donation Service Area.

What do the LS estimates in table 5 tell us? In short, the number of transplantations performed at a transplant center is predicted to decrease the average patient survival rate for each additional transplant; its p-value is 0.081. Additional competition is predicted to decrease the average patient survival rate by between 3.4 and 10.5 percent, depending on the number of transplant centers in the Donation Service Area. The `share` of transplantations performed by a transplant center in a Donation Service Area is predicted to increase the average patient survival rate; its p-value is 0.002. In the example in which two transplant centers compete, where one transplant center performs two-thirds of the transplants, while the other performs one third, the "larger" transplant center has an average patient survival rate that is around 5.0 percent greater than the "smaller" one: bigger is better in this empirical specification, too. The LAD estimates are qualitatively similar, but quantitatively smaller in magnitude.

We do not report all 104 estimated coefficients on the dummy variables, nor do we present graphs as we did for the graft survival-rate data, but the results are similar to the graft survival-rate ones.

Of course, the reader may rightly be concerned that using of the observations across different durations could be inducing the empirical results we observe. Thus, we estimated the graft survival rate and patient survival rate equations across separate durations, including as covariates the competition dummy and center-specific dummy variables as well as the `share` variable, thus performing the estimation on the same durations—one month, one year, and three years. We report in tables 6 and 7 the LS

Table 5: LS and LAD Estimates of Equation (1) Including Center-Specific Dummy Variables, Patient Survival Rate

Variable	LS	LAD
Constant	0.03956	0.02232
	(0.01999)	(0.01203)
d12	−0.00269	0.00296
	(0.00240)	(0.00167)
d36	−0.00807	0.00525
	(0.00421)	(0.00190)
ntxs	−0.00013	−0.00009
	(0.00007)	(0.00004)
two	−0.03360	−0.03018
	(0.01482)	(0.01401)
three	−0.07645	−0.05877
	(0.02500)	(0.01821)
four	−0.08993	−0.06541
	(0.03046)	(0.00194)
five+	−0.10555	−0.07707
	(0.03415)	(0.02049)
share	0.04184	0.01997
	(0.01748)	(0.00307)
SSR	8.91731	10.59230

Table 6: LS Estimates of Competition and Share Effects: Graft Survival-Rate and Patient Survival-Rate Equations, Including Center-Specific Dummy Variables, By Duration

Variable	GSR(1)	PSR(1)	GSR(12)	PSR(12)	GSR(36)	PSR(36)
`ntxs`	−0.00018	−0.00009	−0.00032	−0.00010	−0.00027	−0.00015
	(0.00005)	(0.00007)	(0.00008)	(0.00013)	(0.00018)	(0.00020)
`two`	−0.01864	−0.01752	−0.05684	−0.05003	−0.09653	−0.05653
	(0.01219)	(0.01211)	(0.01799)	(0.02323)	(0.04687)	(0.02920)
`three`	−0.05526	−0.05128	−0.11293	−0.10745	−0.16393	−0.12885
	(0.01955)	(0.02180)	(0.03168)	(0.04147)	(0.05351)	(0.05615)
`four`	−0.07220	−0.06385	−0.13198	−0.11991	−0.18075	−0.15210
	(0.02542)	(0.02886)	(0.03969)	(0.05517)	(0.08134)	(0.07835)
`five+`	−0.09278	−0.08834	−0.13945	−0.12436	−0.18682	−0.17375
	(0.02942)	(0.03368)	(0.04708)	(0.06495)	(0.09000)	(0.08745)
`share`	0.04931	0.03644	0.08560	0.05673	0.04998	0.05665
	(0.01713)	(0.02093)	(0.02633)	(0.03983)	(0.05319)	(0.05219)

and LAD estimates of the coefficients concerning the competion dummy variables as well as the `share` variable.

Consider the results in table 6. The general pattern is that the effects of competition on graft survival rates across durations are stronger than for patient survival rates, but these effects increase as the duration increase. For example, consider the columns GSR(1) and PSR(1): each of the estimates in the GSR(1) column is smaller (more negative) than those in the PSR(1) column. Across the GSR columns, however, the effect becomes more pronounced at longer durations: the estimates in the GSR(36) column for competition are all much smaller (more negative) than those for the GSR(1) column. The `share` variable fluctuates across the columns, but never changes sign. However, it is imprecisely estimated (has large standard errors) for the longest duration, three years. In figures 12 and 13, we present the LS percentage reductions in the graft survival rates and patient survival rates, by duration, for different numbers of transplant centers in the Donation Service Area. This pattern is

Table 7: LAD Estimates of Competition and Share Effects: Graft Survival-Rate and Patient Survival-Rate Equations, Including Center-Specific Dummy Variables, By Duration

Variable	GSR(1)	PSR(1)	GSR(12)	PSR(12)	GSR(36)	PSR(36)
`ntxs`	−0.00010	−0.00006	−0.00026	−0.00018	−0.00014	−0.00000
	(0.00003)	(0.00003)	(0.00005)	(0.00006)	(0.00009)	(0.00009)
`two`	−0.00204	−0.00490	−0.06570	−0.03967	−0.04027	−0.03726
	(0.01142)	(0.00970)	(0.01979)	(0.01795)	(0.02933)	(0.02810)
`three`	−0.02891	−0.02582	−0.11639	−0.11305	−0.07396	−0.08174
	(0.01576)	(0.01333)	(0.02731)	(0.02466)	(0.03692)	(0.03541)
`four`	−0.03295	−0.03200	−0.12854	−0.12129	−0.07129	−0.09477
	(0.01715)	(0.01442)	(0.02973)	(0.02667)	(0.03901)	(0.03779)
`five+`	−0.04504	−0.04027	−0.11989	−0.10983	−0.10438	−0.13255
	(0.01828)	(0.01533)	(0.03169)	(0.02835)	(0.04000)	(0.03867)
`share`	0.01850	0.01277	0.06034	0.05746	0.00286	0.02418
	(0.00360)	(0.00259)	(0.00624)	(0.00479)	(0.00914)	(0.00869)

qualitatively similar for the LAD estimates presented in table 7.

6 Summary and Conclusions

Previous researchers have modelled the patient/surgeon decision of whether to accept a donor organ for transplantation as a Markov decision problem, the solution to which (under plausible assumptions) is a control-limit optimal policy: for a patient of a given health status, accept any organ whose quality is above some threshold, which typically depends on the patient's health status; otherwise, wait for another organ to arrive. The presence of competing transplant centers in a Donation Service Area is predicted to affect the patient/surgeon acceptance decision through a mechanism we have referred to as competitive impatience: competition makes patient/surgeon decision-makers more likely to accept a donor organ than when no competition exists, which means (all other things being equal) the matches made under competition

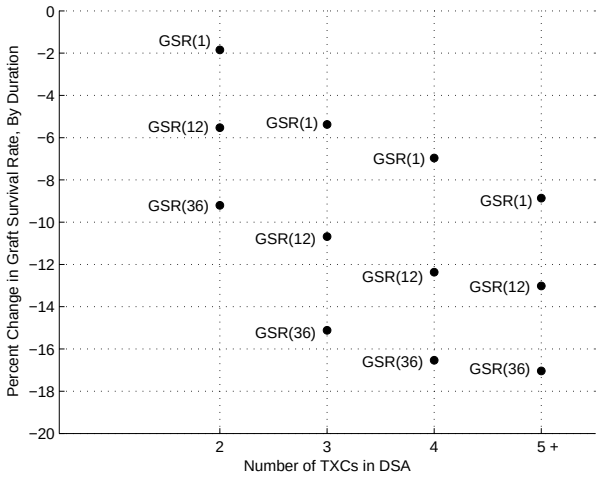

Figure 12: LS Percentage Change in Graft Survival Rates, by Duration and Number of transplant centers in Donation Service Area

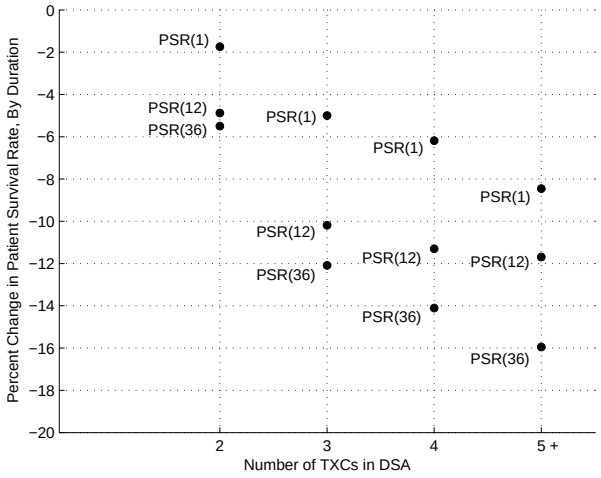

Figure 13: LS Percentage Change in Patient Survival Rates, by Duration and Number of transplant centers in Donation Service Area

are predicted to be of weakly lower quality. Assuming match quality affects graft success in a weakly-positive way and holding all other factors constant, the survival function of the waiting time to graft failure after transplantation is then predicted to be weakly greater when no competitors exist than that under competition. When the competing transplant centers in a Donation Service Area perform different numbers of transplantations and have waiting lists of different lengths, such asymmetries are also predicted to affect post-transplantation outcomes. Using center-specific actual as well as risk-adjusted average graft survival-rate and patient survival-rate data concerning cadaveric-liver transplantations in adults in the United States, we have found that the predictions of the theory are borne out by the data. We have also found evidence that center-specific heterogeneity is very important in affecting post-transplantation outcomes. Our research also suggests that, in order to test the hypothesis of competitive impatience, disaggregated, center-specific, match-list data are required.

Acknowledgements

The authors would like to thank Jay Bhattacharya, Srihari Govindan, Han Hong, Kenneth L. Judd, Clinton J. Levitt, John Rust, Che-Lin Su, Stephen J. Turnbull, and Robert B. Wilson for enlightening discussions as well as useful comments and helpful suggestions. We are grateful to Andrew Barnes and Craig Lake of the Arbor Research Collaboration for Health, who provided us with a collection of SAS datasets from which we distilled the relevant data and who answered a number of questions when we created our final dataset. Paarsch gratefully acknowledges that some of the research for this paper was completed while he was a visiting research scholar at the Center for Economic Institutions in the Institute of Economic Research at

Hitotsubashi University in Kunitachi, Japan as well as a visiting research scholar at the Collegio Carlo Alberto in Moncalieri, Italy.

References

Oguzhan Alagoz, Lisa M. Maillart, Andrew J. Schaefer, and Mark S. Roberts. The optimal timing of living-donor liver transplantation. *Management Science*, 50: 1420–1430, 2004.

Oguzhan Alagoz, Lisa M. Maillart, Andrew J. Schaefer, and Mark S. Roberts. Determining the acceptance of cadaveric livers using an implicit model of the waiting list. *Operations Research*, 55:24–36, 2007a.

Oguzhan Alagoz, Lisa M. Maillart, Andrew J. Schaefer, and Mark S. Roberts. Choosing among living-donor and cadaveric livers. *Management Science*, 53:1702–1715, 2007b.

Donald W.K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59:817–858, 1991.

Donald W.K. Andrews and J. Christopher Monahan. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60:953–966, 1992.

David R. Cox. Regression models and life tables. *Journal of the Royal Statistical Society, Series B*, 34:187–220, 1972.

Erick B. Edwards, John P. Roberts, Maureen A. McBride, James A. Schulak, and Lawrence G. Hunsicker. The effect of the volume of procedures at transplantation

centers on mortality after liver transplantation. *New England Journal of Medicine*, 341(27):2049–20453, 1999.

Friedhelm Eicker. Asymptotic normality and consistency of the least squares estimator for families of linear regressions. *Annals of Mathematical Statistics*, 34:447–456, 1963.

Martin S. Gaynor and William B. Vogt. Competition among hospitals. *RAND Journal of Economics*, 34:764–785, 2003.

Martin S. Gaynor, Harald Seider, and William B. Vogt. The volume-outcome effect, scale economies, and learning-by-doing. *American Economic Review Papers & Proceedings*, 95:243–247, 2005.

David H. Howard. Dynamic analysis of liver allocation policies. *Medical Decision Making*, 21:257–266, 2001a.

David H. Howard. Hope versus efficiency in organ allocation. *Transplantation*, 72: 1169–1173, 2001b.

David H. Howard. Why do transplant surgeons turn down organs? A model of the accept/reject decision. *Journal of Health Economics*, 21:957–969, 2002.

Edwin L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481, 1958.

Harold S. Luft, Sandra S. Hunt, and Susan C. Maerki. The volume-outcome relationship: Practice-makes-perfect or selective-referral patterns. *Health Services Research*, 22:157–182, 1987.

Whitney K. Newey and Kenneth D. West. A simple positive semi-definite heteroskedasticity and autocorrelation consitent covariance matrix. *Econometrica*, 55: 703–708, 1987.

Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119:457–488, 2004.

Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125:151–188, 2005a.

Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver. A kidney exchange clearinghouse in New England. *American Economic Review Papers & Proceedings*, 95:368–371, 2005b.

Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver. Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. *American Economic Review*, 97:828–851, 2007.

John Rust. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, 55:999–1033, 1987.

Susan L. Saidman, Alvin E. Roth, Tayfun Sönmez, Utku Ünver, and Francis L. Delmonico. Increasing the opportunity of live kidney donation by matching for two- and three-way exchanges. *Transplantation*, 81:773–782, 2006.

Dennis P. Scanlon, Christopher S. Hollenbeak, Woolton Lee, Evan Loh, and Peter A. Uber. Does competition for transplantable hearts encourage 'gaming' of the waiting list? *Health Affairs*, 23:191–198, 2004.

David M. Shahian and Sharon L. Normand. The volume outcome relationship: From Luft to leapfrog. *Annals of Thoracic Surgery*, 75:1048–1058, 2001.

Jason Snyder. Gaming the liver transplant market. *Journal of Law, Economics, and Organization*, doi:10.1093/jleo/ewq003, in press.

Halbert L. White, Jr. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica*, 48:817–838, 1980.

Achim Zeileis. Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, 11:1–17, 2004.