Herding in Quality Assessment: Identification and Welfare Analysis*

Stephanie De Mel† Kaivan Munshi‡ Soenje Reiche§ Hamid Sabourian¶

October 16, 2019

Abstract

There are many economic environments in which an object is offered sequentially to prospective buyers. It is often observed that once the object for sale is turned down by one or more agents, those that follow do the same. One explanation that has been proposed for this phenomenon, which goes back to Banerjee (1992) and Bikhchandani et al. (1992) is that agents making choices further down the line (rationally) ignore their own assessment of the object's quality and herd behind their predecessors. We develop novel tests to detect information-based herding, together with a methodology to quantify its welfare consequences, which are applied to organ transplantation in the U.K. We find that herding is common and is an important contributor to the high rate at which organs are rejected (and subsequently discarded). However, herding does not raise discard rates much above the level that would be obtained if private assessments were made publicly available, with the (limited) information contained in predecessors' decisions substantially reducing false acceptances. The information loss highlighted by the canonical models of herding appears to be less serious in this setting, but this does not imply that information problems are absent; a counterfactual simulation of the estimated model indicates that an intervention that makes a recently developed risk index of organ quality available to all centers would substantially reduce discard rates, while simultaneously improving the selection of organs.

Keywords. Social learning. Herd behavior. Information Cascades. Organ transplant decisions.
JEL. J12. J16. D31. I3.

*We are grateful to Kourosh Saeb-Parsy and Patrick Trotter for their collaboration on this project. Research support from the Keynes Fund and Cambridge-INET is gratefully acknowledged. We are responsible for any errors that may remain.

†Institute for Fiscal Studies
‡Yale University
§Yale University
¶University of Cambridge

1 Introduction

There are many economic environments in which prospective buyers, acting sequentially, must choose whether or not to buy an object. Examples of such environments include venture capital and property development, where a startup or a piece of land is offered for sale, and the labor market, notably the draft in professional sports leagues and the academic job market. It is often observed in these environments that once the object for sale is turned down by one or more agents, those that follow do the same. One explanation for this correlation in decisions is that the object is (correctly) assessed to be of poor quality by all agents. An alternative explanation, which goes back to seminal contributions by Banerjee (1992) and Bikhchandani et al. (1992) is that agents who must make choices further down the line (rationally) ignore their own assessment of the object's quality and herd behind their predecessors. The statistical identification of information–based herding is a challenging problem. In this paper, we develop novel strategies to detect herding in settings where decisions are characterized by a sequence of rejections followed by either an acceptance or termination of trade, together with a methodology to quantify its welfare consequences.

We apply our methodology to organ transplantation in the United Kingdom. The organ transplant program in the U.K. is organized around a nationwide network of centers (hospitals). When a deceased donor organ becomes available, all patients on the National Transplant Registry are assigned a priority rank based on a predetermined allocation algorithm. Transplant centers are offered the organ in order of their patients' priority, until the organ either is accepted or, having deteriorated with time, is no longer viable and is discarded. We will see that the organization of the U.K. transplant program makes it an ideal test-case for our methodology. This is also a setting in which information-based herding may have large practical consequences. Currently, demand outstrips the availability of both livers and kidneys, the two organs that dominate transplantation activity in the U.K. and that constitute the focus of our analysis. National Health Service Blood and Transplantation (NHSBT) statistics indicate that, five years after being listed, approximately 20% of patients on the National Registry have either died or been removed from the waiting list as their condition has deteriorated below the minimum eligibility criterion for transplantation. Simultaneously, however, approximately 50% of livers and 30% of kidneys are discarded. Our methodology allows us to quantify the contribution of information-based herding to this relatively high discard rate.

To understand why transplant centers might rationally condition their decisions on those of the centers that preceded them, and why such herding behavior could generate inefficiencies, suppose that there are two types of organs: good (G) and bad (B). The optimal decision is to accept a good organ and reject a bad organ. Each transplant center makes an assessment of the quality of the organ that is made available to it. This assessment, which we characterize as an information signal, is not directly observed by other transplant centers. As in Bikhchandani et al. (1992) and Anderson and Holt (1997), we assume that signals are binary: good (g) and bad (b). Centers are not systematically misinformed; they are thus more likely to receive a g (b) signals when an organ is good (bad). For expositional convenience, we assume that transplant centers have a common negative prior about the quality of offered organs; thus, in the absence of any other information, each center's decision is to reject the organ. Additionally, we assume that, if a center receives a g signal, this dominates its negative prior and it will accept the offered organ.

It follows that the first center to be offered an organ will reject it on receipt of a b signal (which reinforces its prior), but will accept it on receipt of a g signal. The second center in line is only offered the organ if the first center rejects. It knows that the first center only rejects following a b signal. Thus, if the second center also receives a b signal, this reinforces the information contained in both the first center's b signal and the negative prior, and it will certainly reject the organ. If, however, the second center receives a g signal, it knows that its signal is not aligned with that of the first center. Under the assumption that all transplant centers receive signals of equal precision (i.e. that they are all equally competent in assessing the quality of an organ), the first and second centers' signals cancel each other out, and the second center also rejects (based on its negative prior). Next, consider the third center's decision: it knows that the second center rejects the organ regardless of it's signal, so the second center's decision gives the third center no additional information. Accordingly, the third center behaves as if it were second in line. Following the preceding argument, it also rejects the organ, regardless of the signal it receives. This process is repeated along the entire waiting list, regardless of the sequence of signals received by centers.

While it is individually rational for centers to ignore their signals in the manner described above, herding can result in the under-utilization of viable organs. To see why this is the case, suppose that the first center in line for an organ receives a b signal, but all the centers that follow receive g signals. This organ is rejected by all centers despite the fact that it is evidently a G organ. Banerjee (1992) and Bikhchandani et al. (1992) construct similar examples of this particular type of herding, which is referred to in the literature as an "information cascade." The distinguishing feature of an information cascade is that agents completely ignore their own signals when they herd behind their predecessors; consequently, agents that follow them in the decision-making process can learn nothing from their decisions, resulting in an information inefficiency.

Over the past three decades, the literature on information-based herding has advanced on multiple fronts. Early theoretical contributions examined the robustness of information cascades to alternative signal and choice structures (see Gale (1996) for an overview). In parallel, empirical contributions in finance and development sought to establish the key implication of herding, which is that agents condition their decisions on their predecessors' decisions (Lakonishok et al., 1992; Grinblatt et al., 1995; Wermers, 1999; Foster and Rosenzweig, 1995; Munshi, 2004; Conley and Udry, 2010). More recently, the theoretical focus has shifted to model misspecification (Bohren, 2016; Frick et al., 2019) and learning on networks (Golub and Sadler, 2016). The empirical literature has developed sharper tests of social learning by exploiting experimentally induced variation in predecessors' decisions (Dupas, 2014) and has moved to estimating sophisticated structural models of herding (Zhang, 2010; Cipriani and Guarino, 2014). Although much progress has been made, there remains one important gap: there has been no test, in a real world setting, of the information inefficiency that lies at the heart of the early herding literature. The structural analyses cited above quantify this inefficiency. However, our analysis is the first we are aware of to develop reduced form tests of the information loss associated with herding. Only after these tests are complete do we proceed to the structural analysis, which quantifies the welfare consequences of this information loss in a setting where it is likely to be of practical importance. Apart from the credibility that comes with the reduced form tests, the additional value of our two-step analysis is that it allows us to independently cross-validate the model, as discussed below.

When choices are sequential, one implication of herding is that the same patient (and her associated center) are more likely to reject an organ when they are further down the line. Zhang (2010) exploits this source of variation to estimate a structural model of herding in organ transplant decisions in the U.S. However, other explanations for this observation are also available. For example, lower quality organs, which are less likely to be accepted at any position on average, travel further down the line even without herding. Furthermore, patient priority is determined, in part, by the organ-patient match. Patients lower in line will have a worse match on average and, moreover, the organ is more likely to have deteriorated with time. Zhang controls for these factors in her analysis, but such conditioning is always imperfect. We break now ground by developing tests of herding and its associated information loss that leverage variation in decisions at the *same* position. This variation arises because (i) centers differ in their ability to distinguish between good and bad organs, and (ii) the ordering of centers varies from one organ to the next depending on the priority of their patients.

An obvious implication of herding with center heterogeneity is that a center in second position will be more likely to reject the organ when it follows a higher-ability center in first position (who must necessarily have rejected). This is because the b signal received by the center in first position is more informative about the state of the world; i.e. that the organ is a bad (B) organ. Letting p_2 be the probability that center 2 rejects (conditional on being offered the organ) and q_1 be center 1's ability, this implies that α_1 is positive in the following equation:

$$p_2 = \alpha_0 + \alpha_1 q_1. \tag{1}$$

While an estimated $\alpha_1 > 0$ is consistent with herding, a special feature of our application, which involves sequential decision making with a single object being passed along, is that this result could be obtained even if centers independently follow their own signals with no regard for the decisions of their predecessors. This is because a higher-ability center is more likely to accept a good organ and less likely to accept a bad organ. Compared to a center of lower ability, it thus passes on a worse pool of organs when in first position.

Our first test of herding is thus based on an augmented specification of the preceding equation:

$$p_2 = \alpha_0 + \alpha_1 q_1 + \alpha_2 q_2 + \alpha_3 q_1 \cdot q_2, \tag{2}$$

where q_2 is center 2's ability and our focus is on the interaction term, $q_1 \cdot q_2$. In our model, both choices and signals are binary, allowing center heterogeneity to be incorporated relatively easily; higher ability centers are more likely to receive a $g(b)$ signal with a $G(B)$ organ. Given this structure, a center can either independently follow its own signal or it can herd behind its predecessors, in which case it will reject for sure (regardless of the signal it receives). When centers follow their own signals, center 2 is more responsive to the deterioration in the organ pool that results from center 1 being of relatively high ability when it too is of higher ability. That is, we expect α_3 to be positive. When centers ignore their own signals and herd behind their predecessors, however, this effect could be reversed, resulting in a negative value for α_3. This is because lower ability centers in second position are more likely to abandon their signals and reject with certainty, especially when following higher-ability centers.

Our second test of herding is based on the behavior of centers in third position. To implement this test,

we estimate the following equation, with the probability of rejecting an organ in third position, p_3, as the dependent variable:

$$p_3 = \tilde{\alpha}_0 + \tilde{\alpha}_1 q_1 + \tilde{\alpha}_2 q_2. \tag{3}$$

Consider the case in which centers follow their own signals. As above, a higher-ability center passes on a relatively worse pool of organs to following centres. Thus, $\tilde{\alpha}_1$, $\tilde{\alpha}_2$ are both positive. In addition, $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ are equal to each other when we restrict attention to organs for which q_1 equals q_2. However, once we introduce the possibility of herding (this can happen in second, but not first, position), $\tilde{\alpha}_2$ is strictly smaller than $\tilde{\alpha}_1$. This is because centers that herd (and reject with certainty, regardless of their signal) do not alter the quality of the organ pool and, thus, the rejection probability of center 3.

With binary decisions and signals, herding is synonymous with information cascades (which by definition involve an information loss). While this ties our analysis directly to the canonical papers that motivate this research, our tests of information inefficiency are robust to alternative signal structures. In particular, as discussed below, our tests of herding would go through even with a continuum of signals over an infinite support. Agents will always place some weight on their own signals; i.e. information cascades are ruled out in this case (Smith and Sørensen, 2000). With binary decisions and a continuum of signals, herding is nevertheless associated with information loss (Vives, 1996). Unlike Çelen and Kariv (2004) who use a clever lab experiment to distinguish between herding and information cascades, our objective in this analysis is to simply establish that there is an information inefficiency and our tests suffice for this purpose.

We implement the tests of herding with administrative data obtained from NHSBT. This data covers the universe of deceased-donor livers and kidneys offered between 2006 and 2015 in the United Kingdom. The data includes the sequence of centers that were offered each organ, as well as their decisions, which – with the possible exception of the final center in every sequence – must necessarily be rejections. The first step in estimating equations (2) and (3) is to construct a measure of center ability. Higher-ability centers are better than their low-ability counterparts at detecting both G and B organs. Thus, when the pool of organs is poor, higher-ability centers reject more often than do lower-ability centers. Conversely, when the pool of organs has high average quality, higher-ability centers accept more often. To determine which scenario is relevant for a given organ type – livers versus kidneys – we take advantage of equation (1), which indicates that centers in second position reject more often when following a higher-ability center. We find that, for livers, a center's ability increases in its first-position probability of rejection, while, for kidneys, the reverse is true. Our measure of ability is thus, for livers, precisely this probability and, for kidneys, one minus the corresponding probability. Using this measure, we detect herding with both tests. Based on the estimates of equation (1), the model additionally predicts that the interaction coefficient in equation (2), α_3, should be negative for livers and positive for kidneys. This prediction is supported by the data as well.

Having established that herding is present, the next step is to estimate its prevalence and to quantify its welfare consequences. Accordingly, we estimate the structural parameters of the model and then conduct counter-factual simulations. Heterogeneity in center ability is a distinguishing feature of our model; while this helps us identify herding, it also makes the structural estimation more challenging. Our model has two ability parameters – the probability of receiving a g signal with a G organ and the probability of receiving a b

signal with a B organ – that must be estimated separately for each center. Given the large number of centers, we estimate these ability parameters outside the model. All centers observe the same set of organ and donor characteristics. What distinguishes centers is the weight they place on different characteristics when assessing the quality of an organ. Indices of organ quality have recently been proposed in the organ transplant literature for both livers and kidneys. These indices are constructed from retrospective NHSBT data that covers organ transplant decisions and subsequent patient outcomes over many years; the set of characteristics included in these indices and the weight placed on each characteristic is thus (close to) optimal. Although individual centers base their decisions on many of the characteristics that are incorporated in the risk indices, their decisions are limited by their own experiences. In general, we expect quality assessment and the associated decisions taken by higher-ability centers to track more closely with the risk index. We use this intuition to construct center-specific measures of ability, to characterize the organ-center specific distribution of signals, and to compute the fraction of G organs in the population of organs.

The risk indices allow us to obtain internally consistent estimates of all the model's parameters, with one exception: the threshold belief that an organ is good, above which centers accept it. An increase in this threshold has no bearing on the decisions of centers that follow their own signal; they will accept if they receive a g signal and reject if they receive a b signal. It does, however, increase the fraction of centers that herd and ignore their own signal, rejecting with certainty. We estimate this parameter using the simulated method of moments (based on repeated draws of the information signals) by matching rejection rates predicted by the model to the data. Although the estimation is computationally straightforward, the model places strong restrictions on the parameter values that can be obtained. We verify that the estimated threshold satisfies a key assumption of the model, which is that centers follow their own signals in first position. We also verify that the ability parameters estimated from the risk indices correlate closely with the independently derived ability measure used in the reduced form tests. Finally, we verify that the model does a good job of matching the data; indeed, our model's goodness of fit is substantially better than that of the alternative "no-learning" model, in which centers ignore their predecessors' decisions and always follow their signals.

To measure the prevalence of information cascades, we compute the fraction of decisions in our data for which centers are predicted to have ignored their signals (given our estimate of the threshold belief). Based on this estimate, it is common for centers to ignore their signals: this occurs 66% of the time for livers and 54% of the time for kidneys, with an increase in these statistics at higher positions. However, a more important question is whether such herding has substantial welfare consequences. To answer this question, it is necessary to specify a benchmark, which we define as the counter-factual outcome in which information signals are pooled (common knowledge). In our model, centers know which of their predecessors followed their own signals and, consequently, must have received a b signal when they rejected. The missing information is associated with centers who herd, as the centers that follow them learn nothing from their decisions. With pooled information, by contrast, signals received by all preceding centers are utilized in the decision-making process. We construct the pooled information benchmark by drawing signals for those centers who are predicted by the model to herd (this is possible because the risk index provides us with more information than was historically available to individual centers). The signals we draw are also used to

predict decisions in the alternative no-learning model, in which centers always follow their own signals.

With herding, there are too many rejections relative to the pooled information benchmark. This is because g signals are ignored by centers who herd, and this has spillover effects on the centers that follow. Conversely, with no learning, there are too many acceptances because useful information contained in previous rejection decisions is ignored. Our estimates quantify these opposing effects; we find that over-rejection with the herding model is modest, while over-acceptance with the no-learning model is substantial. These findings motivate the final step of the analysis, in which we compare discard rates under the herding and no-learning models against the pooled information benchmark. The decision to discard an organ is taken by NHSBT and is outside our model. Our interest is in whether an organ that is discarded in the data (with herding) would have been accepted at an earlier position with the alternative models. As decisions are similar under herding and the pooled information benchmark, we do not expect discard rates to diverge substantially. As expected, discard rates with herding are roughly 10% higher than the pooled information benchmark. By contrast, discard rates under the assumption of no learning are 35% lower than the same benchmark, on account of the many false acceptances that arise when centers ignore the information that is contained in their predecessors' decisions. These results collectively indicate that herding contributes substantially to the high discard rate observed in the data, but that this increase in the discard rate is not necessarily inefficient. Centers often ignore their own signals, but their reliance on their predecessors actually protects them from accepting bad organs. This does not imply that information inefficiencies are absent in this setting. The analysis concludes by examining a counter-factual scenario in which the risk indices are made available to (and utilized by) all centers. We find that this easily implementable intervention would substantially reduce discard rates, while simultaneously improving the quality of accepted organs.

2 Institutional Setting

The shortage of suitable donor organs has always been the primary challenge faced by organ transplant programs. In response to this challenge, many countries, including the United Kingdom, have established national allocation schemes for the distribution of organs supplied by deceased donors. Organs obtained from deceased donors are classified according to the manner of death as either DBD (donation after brain death) or DCD (donation after cardiac death). Although DBD and DCD organs do not vary systematically with respect to *ex ante* quality and the same broad allocation protocols are utilized by the National Health Service Blood and Transplant (NHSBT) for both types of organs, DCD organs are useable for a shorter period of time before they must be discarded from the donor pool and set aside for research (Watson and Dark, 2012).

Our analysis focuses on livers and kidneys, for which donors have been matched to recipients in the United Kingdom through a national allocation scheme since the later 1990s. These two types of organ continue to dominate transplantation activity: NHSBT statistics indicate that over 80% of livers and kidneys obtained from DBD donors in 2014-2015 were transplanted, while the corresponding statistics for pancreases, hearts, and lungs were less than 35%. For DBD livers and kidneys, a Transplant Benefit Score (TBS), which puts weight on both the patient's need for a transplant and the patient's organ-specific quality of life after the transplant, is used to rank all patients listed on the National Registry when a given organ becomes available.

The TBS is calculated using a fixed set of donor and recipient characteristics. Transplantation delays are substantially more costly for DCD organs and, hence, the proximity between donor and recipient is also a factor in drawing up the priority list for them.

When an organ becomes available, it is offered to patients in order of their priority. Each patient's hospital (transplant center) has 45 minutes to accept or decline the offer, based on the information that is made available; this process continues until the organ has been accepted or until too much time has elapsed for it to remain useable. For livers, DBD (DCD) organs should be transplanted within 12 (6) hours, while the corresponding cutoffs for kidneys are 18 (12) hours. This is a narrow time window, leaving room for just a few centers to make decisions before an organ is discarded by NHSBT. We will see below that queue lengths for organs rarely exceed eight centers.

The allocation of deceased-donor organs in the United Kingdom differs in important respects from the allocation procedure in the United States. Zhang (2010), using data on kidney donations in Texas, documents that on average an organ is accepted by the 34^{th} patient in line, who has already turned down 15 offers. Such long queue lengths are possible because organs are only discarded after 48 hours. Under these circumstances, the condition of the organ becomes a major consideration in decision-making, particularly at higher positions. The mismatch between organ and recipient also becomes relevant (in Zhang's data, kidneys are accepted as late as the 77^{th} position). Given this mismatch, patients consider (and reject) many organs before finally accepting and, hence, dynamic considerations enter the decision rule. Both Zhang (2010) and Agarwal et al. (2019), who also study kidney allocation in the United States, model the acceptance decision as an optimal stopping problem. The institutional environment in the United Kingdom, where organs are almost always accepted by patients towards the very top of the national priority list and where mismatch and the associated strategic inter-temporal considerations are thus less relevant, allows us to focus in our model on a new aspect of decision-making, which is the ability of centers to correctly assess the quality of the organs that they are offered.

3 A Model of Organ Transplantation

3.1 Organs, Centers and Signals

There is a pool of organs of unknown quality. Organs can be either of good (G) or bad (B) quality. The outcome of an organ transplant is denoted by H if the organ is good, and by L if it is bad, with $H > 0 > L$. Payoffs H and L are realized independently of the center (hospital) undertaking the transplant and the identity of the patient who receives the organ. This implies that centers are equally competent in implementing the transplant procedure (as validated below) and that variation in the organ-recipient match and in recipient characteristics are small enough to be ignored when modeling transplant decisions (consistent with the discussion above). We normalize the outcome of not transplanting an organ to 0. The prior on organ quality is denoted by π; this is the probability that a random organ from the pool is a G organ. We define the *cut-off* belief $\tilde{\pi}$ as the belief at which every hospital is indifferent between accepting or rejecting an organ; i.e. $\tilde{\pi}H + (1-\tilde{\pi})L = 0$, or $\tilde{\pi} = \frac{-L}{H-L}$.

Centers individually assess organ quality before making a decision. This assessment is based on a set of organ (and donor) characteristics that is observed by everyone. What separates centers is the weight they place on each characteristic, with more able centers, who are better positioned to distinguish between G and B organs coming closer to the optimal weight. We characterize each center's independent assessment of an organ by a private information signal $s \in \{g, b\}$, where a g signal indicates that the organ is good, while b indicates that it is bad. The advantage of the binary signal is that it allows us to incorporate a continuum of abilities in the model. Center ability is captured by a parameter $q_j \in [\underline{q}, \overline{q}] \subset \mathbb{R}$ for each center j. A center's ability q_j determines the probability $\gamma_j \in (0,1)$, with which it correctly identifies a G organ, and the probability $\beta_j \in (0,1)$, with which it correctly identifies a B organ. In particular, for center j, $prob(g \mid G) = \gamma_j = \gamma(q_j)$ and $prob(b \mid B) = \beta_j = \beta(q_j)$, with strictly increasing functions $\gamma, \beta : [\underline{q}, \overline{q}] \to (0,1)$.

A center in first position with ability q_j updates its prior belief that an organ is good, π, upon receiving g and b signals, such that

$$\begin{aligned} \pi_j(G|g) &= \frac{\pi\gamma_j}{\pi\gamma_j + (1-\pi)(1-\beta_j)}, \\ \pi_j(G|b) &= \frac{\pi(1-\gamma_j)}{\pi(1-\gamma_j) + (1-\pi)\beta_j}. \end{aligned}$$

The prior shifts up (down) upon receipt of a $g(b)$ signal:

$$\pi_j(G|g) > \pi > \pi_j(G|b) \tag{4}$$

if the following condition is satisfied

Assumption 1: *For all j, $\beta_j \geq 1 - \gamma_j$.*

The preceding assumption implies that centers are not systematically misinformed; i.e. each center is more likely to receive a b signal with a bad organ than with a good organ. We further assume that centers always follow their own signals in first position (absent any other information) such that each center accepts the organ if it receives a g signal and declines the organ upon receipt of a b signal. This requires that:

Assumption 2: *For all j, $\pi_j(G|g) \geq \tilde{\pi} > \pi_j(G|b)$.*

3.2 Transplant Decisions

Organs are offered sequentially to centers on the basis of a predetermined algorithm. The priority list, both for a given organ and across organs, is based on recipients' characteristics and is independent of the centers to which they are attached. Patients on the waiting list are assigned to transplant centers in their region of residence. It is possible, given the variation in demographic characteristics across regions, that the average position in the sequence across organs will vary by center. This has no bearing on our analysis as long as centers cannot choose endogenously when to make their decision.

Center 1 receives a signal and, given Assumption 2, accepts after a g signal and declines after a b signal. If the organ is accepted, it is transplanted by center 1 and results in payoff H or L, depending on its quality.

If it is declined, an administrator from NHSBT decides either to offer the organ to the next center or to set it aside for research. The decision to discard an organ is based on its condition or useability, which is determined by the delay in retrieving the organ and the subsequent delay prior to transplantation. NHSBT administrators monitor the condition of the organ during the offering process (this information is not available to centers), discarding it as soon as it is considered to be unsuitable for transplantation. Because the decision to discard an organ is orthogonal to its quality, this will not affect the next center's prior on organ quality. We will, nevertheless, allow for the possibility that organ condition declines systematically by position (despite NHSBT's efforts to avoid this outcome) in the empirical analysis.

Centers positioned further along in the sequence learn from the (rejection) decisions of their predecessors. Each center knows the identity of its predecessors and the order in which they made their decisions. If this were not the case, then all the tests reported below would fail to be supported by the data. In the discussion that follows, we identify a center by its position in this sequence, such that the center at position j has ability q_j. We use an iterative process to describe centers' equilibrium beliefs and strategies moving down the line. The equilibrium concept that characterizes the learning process is Perfect Bayesian Nash Equilibrium.

If center 2 is offered an organ, it knows that center 1 must have received a b signal, given Assumption 2. Its prior belief (before it receives its own signal) is therefore

$$\pi_2 = \pi_1(b) = \frac{\pi(1-\gamma_1)}{\pi(1-\gamma_1)+(1-\pi)\beta_1}. \tag{5}$$

Its posterior belief when it receives a g signal is

$$\pi_2(g) = \frac{\pi_2\gamma_2}{\pi_2\gamma_2+(1-\pi_2)(1-\beta_2)},$$

and its posterior belief when it receives a b signal is

$$\pi_2(b) = \frac{\pi_2(1-\gamma_2)}{\pi_2(1-\gamma_2)+\pi_2\beta_2}.$$

Center 2 always rejects the organ if it receives a b signal, because its prior belief, π_2 (which is lower than $\tilde{\pi}$ from Assumption 2), is downgraded even further following a b signal. Center 2 could reject the organ even if it receives a g signal – which implies that it is herding – if this updating does not raise its posterior above $\tilde{\pi}$. To summarize, center 2's optimal decision is to accept the organ if it received a g signal and $\pi_2(g) \geq \tilde{\pi}$, and to decline otherwise.

Next, center 3 knows center 2's decision-making process and its prior belief, π_2, but does not necessarily know center 2's signal. If center 2 herds, its decision provides no information about its signal to center 3, and the latter's prior belief is therefore equal to π_2. If, on the other hand, center 2 uses its signal to make its decision ($\pi_2(g) \geq \tilde{\pi}$), center 3 infers from center 2's rejection that it must have received a b signal, and

therefore has a prior belief equal to $\pi_2(b)$:

$$\pi_3 = \begin{cases} \pi_2, & \text{if } \pi_2(g) < \tilde{\pi} \\ \pi_2(b) & \text{otherwise.} \end{cases}$$

The preceding discussion can be easily generalized. In the same way as center 2, center $n > 3$, given its prior belief (π_n), forms its posterior belief (either $\pi_n(g)$ or $\pi_n(b)$), and then chooses optimally either to accept or to decline the organ. Center $n+1$'s prior then depends (as just described for center 3) on whether its predecessor, center n, herds or not.

3.3 Center Heterogeneity

As discussed, the novelty of our tests of herding is that they are based on variation in center decisions at the same position. This variation arises because centers differ in their ability to distinguish between good and bad organs and because the order of centers varies from one organ to the next. The first step in deriving our tests is thus to use the model to construct a measure of center ability.

Our measure of a center's ability is based on its probability of rejection when in first position. Given Assumption 2, which specifies that all centers follow their signals in first position, the probability that center 1 (with ability q_1) rejects an organ is equal to the probability that it receives a b signal:

$$p_1(q_1) = \pi(1-\gamma_1) + (1-\pi)\beta_1. \tag{6}$$

$$\frac{dp_1(q_1)}{dq_1} = -\pi\gamma'(q_1) + (1-\pi)\beta'(q_1). \tag{7}$$

It is evident from the preceding expression that $p_1(q_1)$ could be increasing or decreasing in q_1 because more able centers are better at detecting both good and bad organs; i.e. $\gamma'(q_1)$ and $\beta'(q_1)$ are both positive. In an inferior organ pool, with many bad organs, the $\beta'(q_1)$ term dominates and $p_1(q_1)$ is increasing in q_1. In a superior organ pool, with many good organs, the $\gamma'(q_1)$ term dominates and $p_1(q_1)$ is decreasing in q_1. We allow for both possibilities, with the restriction that the probability of rejection, for a given organ pool, is either monotonically increasing or decreasing in ability for all centers; i.e. for any $q_1 \in [\underline{q}, \overline{q}]$.

To determine whether the probability of rejection in first position is increasing or decreasing in center ability, we examine the decisions of centers in second position. To begin with, assume that center 2 follows its own signal. This would be the case if it ignores its predecessors' decisions or if it does not herd; that is, its posterior belief upon receiving a g signal exceeds $\tilde{\pi}$. In this case, the probability that center 2 rejects the organ is the probability that it receives a b signal, conditional on center 1 also having received a b signal. Applying Bayes' Rule:

$$p_2(q_1, q_2) = \frac{\pi(1-\gamma_1)(1-\gamma_2) + (1-\pi)\beta_1\beta_2}{\pi(1-\gamma_1) + (1-\pi)\beta_1}. \tag{8}$$

We can now compute, using Assumption 2, the manner in which center 2's rejection probability varies with center 1's ability:

$$\frac{\partial p_2(q_1, q_2)}{\partial q_1} = \frac{\pi(1-\pi)(\gamma_1'\beta_1 + \beta_1'(1-\gamma_1))(\beta_2 - (1-\gamma_2))}{(\pi(1-\gamma_1) + (1-\pi)\beta_1)^2} \geq 0. \tag{9}$$

Furthermore, an increase in center 1's ability makes it more likely that center 2 herds and rejects with certainty. This is because a higher-ability predecessor's rejection has a bigger impact on center 2's prior belief, thereby increasing the likelihood that its posterior belief will remain below $\tilde{\pi}$ even when it receives a g signal. In general, center 2 is more likely to reject when center 1 has high ability, regardless of whether centers learn from their predecessors or not.

If we observe that centers in second position are more (less) likely to reject when they follow centers with a higher probability of rejection when in first position, then equation (9) implies that this probability is positively (negatively) associated with center ability. We will see later that the sign of this relationship is positive for livers and negative for kidneys. We therefore use the probability of rejection in first position to measure center ability for liver transplants and one minus this probability to measure center ability for kidney transplants in our tests of herding.

3.4 A Test of Herding (based on decisions in second position)

The preceding discussion indicates that center 2 is more likely to decline an organ when it follows a higher-ability center, with and without herding. To test for herding we thus need to put more structure on the relationship between center 2's rejection decision and center 1's ability, q_1. In particular, we examine the way in which this relationship varies with center 2's ability, q_2, by estimating equation (2), $p_2 = \alpha_0 + \alpha_1 q_1 + \alpha_2 q_2 + \alpha_3 q_1 \cdot q_2$, and then focusing on the interaction term.

First, assume that center 2 does not herd, such that its rejection probability is described by equation (8). The cross-partial with respect to q_1 and q_2, which is essentially the coefficient on the interaction term, is then

$$\frac{\partial^2 p_2(q_1, q_2)}{\partial q_1 \partial q_2} = \frac{\pi(1-\pi)(\gamma_1'\beta_1 + \beta_1'(1-\gamma_1))(\beta_2' + \gamma_2'))}{(\pi(1-\gamma_1) + (1-\pi)\beta_1)^2} > 0, \tag{10}$$

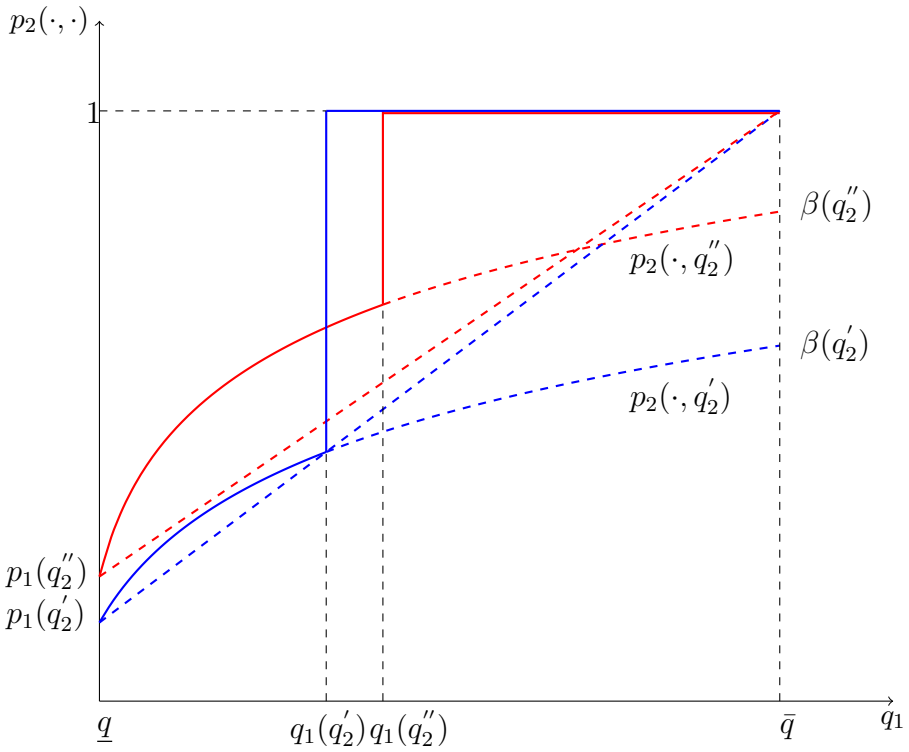
That is, the effect of an increase in center 1's ability on center 2's rejection probability is larger when center 2 has higher ability. This result is obtained because (i) an increase in q_1 reduces the quality of the organ pool passed on to center 2, π_2, and (ii) center 2's decision, p_2, is more sensitive to π_2 when it has higher ability. In the extreme case, if center 2 is completely uninformed, then it will choose at random and, hence, its decision is unaffected by the change in the quality of the organ pool.

Although equation (10) shows that the cross partial is positive in the absence of herding, this result does not necessarily hold when center 2 herds. In particular, there are now two effects: the *quality selection effect* and the *herding effect*. The former, which we described above, implies that a low-ability center reacts less to an increase in its predecessor's ability than does a high-ability center, because the former is less sensitive to the quality of its organ pool. The latter effect works in the opposite direction: as the rejection by a high-ability center 1 represents worse news about underlying organ quality than does the rejection of a low-ability center, it is more likely that a weak center 2 herds and also rejects. Which effect dominates depends on the underlying organ pool and on the ability of the relevant centers.

This is best seen in the following two figures. In drawing these figures, we have assumed that the lowest-ability center is completely uninformed and takes a random decision, and that the highest-ability center is

perfectly informed, although the results do not rely on those assumptions.[1] The figures show the rejection probabilities of two centers at position 2 as a function of center 1's ability q_1. The centers at position 2 differ in their own ability, and we assume that $q_2' < q_2''$. The curves labelled $p_2(\cdot, q_2')$ and $p_2(\cdot, q_2'')$ are the centers' respective rejection probabilities *without herding*, given by the expression in equation (8). If center 2 follows center 1 with $q_1 = \underline{q}$ it faces an organ pool with quality π and, consequently, draws signals as if it were in first position. This implies that $p_2(\underline{q}, q_2) = p_1(q_2)$, which pins down the intercept of each curve. Note that Figure 1 assumes that $p_1(q_2') < p_1(q_2'')$, which we later see applies to livers, while Figure 2 assumes that the inequality is reversed, which is relevant for kidneys. This represents the only difference between the two figures.

Figure 1: $\frac{dp_1(q)}{dq} > 0$

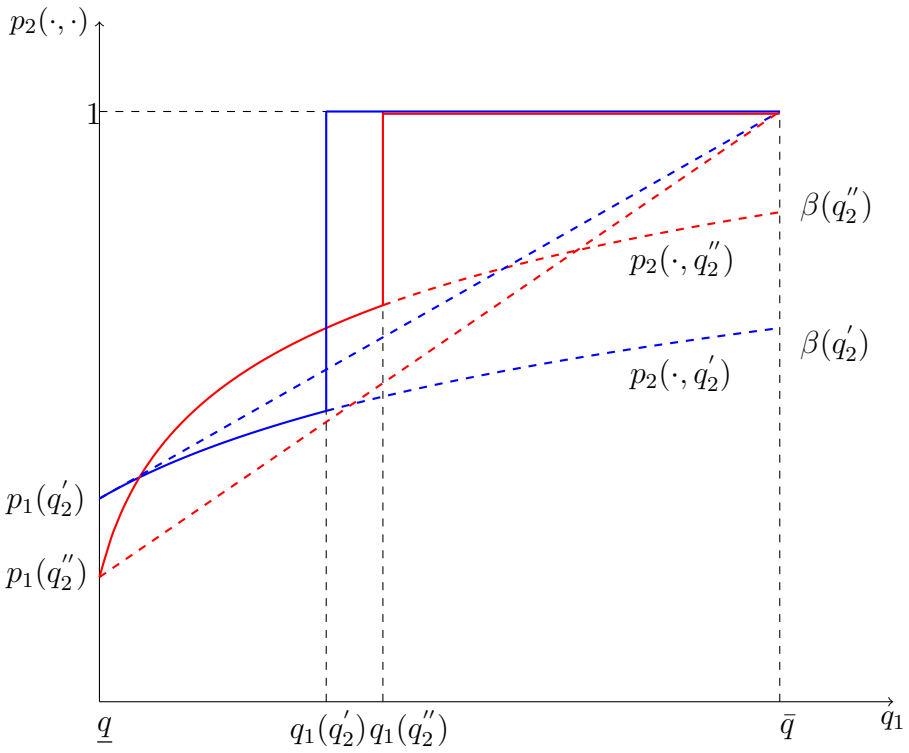


If center 2 follows center 1 with $q_1 = \bar{q}$, the organ is a B organ for certain, so center 2 draws signals from a B organ. This explains why the curves reach a height of $\beta(q_2)$, where $\beta(q_2') < \beta(q_2'')$. *With herding* the rejection probability of each center 2 jumps to 1 at a threshold q_1. This happens at $q_1(q_2)$, where $q_1(q_2') < q_1(q_2'')$, because a lower-ability center positioned at 2 starts herding sooner than does a high-ability one. Thus, for $q_1 < q_1(q_2)$, center 2 uses its signal and rejects according to the expression in equation (8) while, for $q_1 \geq q_1(q_2)$, center 2 herds and always rejects.

We now use the figures to derive the effect of the interaction term $q_1 \cdot q_2$, on center 2's rejection probability with herding. Expression (10) allows us to obtain a measure of this effect for each pair (q_1, q_2) in the absence of herding. With herding, we cannot use the same method, because center 2's rejection probability contains a jump when center 2 starts to herd and, hence, the derivative is not well-defined at each point. Instead,

[1] Formally, $1 - \gamma(\underline{q}) = \beta(\underline{q})$ and $\gamma(\bar{q}) = \beta(\bar{q}) = 1$.

Figure 2: $\frac{dp_1(q)}{dq} < 0$



we compute the"average effect" of an increase in q_1; this is the slope of the line starting at $(\underline{q}, p_1(q_2))$ and going to $(\bar{q}, p_2(\bar{q}, q_2))$. We then examine how this slope varies with q_2 (q_2' versus q_2''). Consistent with the cross-partial expression in (10), the average slope with respect to q_1 is increasing in q_2 in the absence of herding in both figures:

$$\frac{p_2(\bar{q}, q_2') - p_1(q_2')}{\bar{q} - \underline{q}} < \frac{p_2(\bar{q}, q_2'') - p_1(q_2'')}{\bar{q} - \underline{q}}$$

When we incorporate the effect of herding, however, we see that the slope with respect to q_1 in Figure 1 is decreasing in q_2:

$$\frac{1 - p_1(q_2')}{\bar{q} - \underline{q}} > \frac{1 - p_1(q_2'')}{\bar{q} - \underline{q}}.$$

By contrast, the slope with respect to q_1 in Figure 2 is increasing in q_2. Thus, when center ability is *increasing* in the rejection probability in first position,($\frac{dp_1(q)}{dq} > 0$, as observed for livers), we predict that the interaction effect is reversed with herding: lower-ability centers at position 2 react more to an increase in center 1's ability because the herding effect dominates. When center ability is *decreasing* in the rejection probability in first position, ($\frac{dp_1(q)}{dq} < 0$, as observed for kidneys), we predict that the interaction effect goes in the same direction as it does in the absence of herding; the quality selection effect then dominates, such that higher-ability centers at position 2 react more to an increase in center 1's ability. To summarize:

Test 1 *Without herding, the average cross-partial effect of an increase in both center* 1*'s and center* 2*'s ability on center 2's rejection probability is strictly positive. With herding, the average cross-partial effect is strictly positive if center ability is decreasing in rejection probability at first position and negative if center*

ability is increasing in rejection probability at first position.

3.5 A Test of Herding (based on decisions in third position)

Our second test of herding is based on decisions at position 3, in particular, on the relationship between these decisions and center abilities at position 1 (q_1) and position 2 (q_2), as expressed in equation (3): $p_3 = \tilde{\alpha}_0 + \tilde{\alpha}_1 q_1 + \tilde{\alpha}_2 q_2$. In deriving this test we assume that center 3 does not herd. Centers that herd at third position always reject, and their decision is thus unaffected by marginal changes in q_1 and q_2.

To develop our second test of herding we investigate the effect of a marginal increase in q_1 and q_2 on center 3's rejection probability, p_3. We first consider the case without herding, in which p_3, is the probability that center 3 receives a b signal, conditional on both center 1 and center 2 also having received b signals. Applying Bayes' Rule:

$$p_3(q_1, q_2, q_3) = \frac{\pi(1-\gamma_1)(1-\gamma_2)(1-\gamma_3) + (1-\pi)\beta_1\beta_2\beta_3}{\pi(1-\gamma_1)(1-\gamma_2) + (1-\pi)\beta_1\beta_2}. \tag{11}$$

It is easy to see that an increase in either center 1's or center 2's ability decreases the quality of the organ pool passed on to center 3, which increases the latter's rejection probability. Formally,

$$\frac{\partial p_3(q_1, q_2, q_3)}{\partial q_1} = \Theta\beta_2(1-\gamma_2)(\gamma_1'\beta_1 + \beta_1'(1-\gamma_1)), \tag{12}$$

$$\frac{\partial p_3(q_1, q_2, q_3)}{\partial q_2} = \Theta\beta_1(1-\gamma_1)(\gamma_2'\beta_2 + \beta_2'(1-\gamma_2)), \tag{13}$$

with $\Theta = \frac{\pi(1-\pi)(\beta_3-(1-\gamma_3))}{[\pi(1-\gamma_1)(1-\gamma_2)+(1-\pi)\beta_1\beta_2]^2}$. Both expressions clearly are strictly positive, but their exact magnitudes depend on the abilities of both centers. For $q_1 \neq q_2$, either effect could be larger than the other; for $q_1 = q_2$, however, expressions (12) and (13) are identical, such that the two effects are the same.

Now consider a situation in which center 2 herds. A marginal increase in its ability then has no impact on the organ pool passed on to center 3. By contrast, center 1 always uses its signal, which implies that an increase in its ability always worsens center 3's organ pool. It follows that a marginal increase in center 1's ability will have a bigger impact on center 3's rejection probability than a corresponding increase in center 2's ability.

To formally derive this result, denote center 3's rejection probability when center 2 herds by p_3^h. This is essentially the probability that center 3 receives a b signal, conditional only on center 1 having received a b signal:

$$p_3^h(q_1, q_2, q_3) = \frac{\pi(1-\gamma_1)(1-\gamma_3) + (1-\pi)\beta_1\beta_3}{\pi(1-\gamma_1) + (1-\pi)\beta_1}. \tag{14}$$

In this case, the effect of an increase in the ability of center 1, though different to the case without herding (see equation (12)), is still positive, while the effect of an increase in the ability of center 2 is zero:

$$\frac{\partial p_3^h(q_1,q_2,q_3)}{\partial q_1} = \Pi(\gamma_1'\beta_1 + \beta_1'(1-\gamma_1)), \tag{15}$$

$$\frac{\partial p_3^h(q_1,q_2,q_3)}{\partial q_2} = 0, \tag{16}$$

where $\Pi = \frac{\pi(1-\pi)(\beta_3-(1-\gamma_3))}{[\pi(1-\gamma_1)+(1-\pi)\beta_1]^2}$. Summarizing the preceding discussion:

Test 2 *Suppose that centers at position* 1 *and* 2 *have identical abilities. Then, without herding, the effect of an increase in center* 1*'s ability on center 3's rejection probability equals the effect of an increase in center 2's ability. With herding, the effect of an increase in center* 1*'s ability is larger than the effect of an increase in center* 2*'s ability.*

3.6 Herding with a Continuum of Organ Qualities and Signals

We have assumed thus far that there are two types of organs (G and B) and two types of signals (g and b). Suppose, instead, that there is a continuum of organ qualities, $\theta \sim N(\mu, \sigma_\theta)$, and a continuum of associated signals, such that for an organ of type $\tilde{\theta}$, center j receives signals that are distributed normally with mean $\tilde{\theta}$ and standard deviation σ_j. Centers are not systematically misinformed, but more able centers receive more precise signals; i.e. $\sigma_j < \sigma_{j'}$ if $q_j > q_{j'}$.

We normalize such that the optimal decision is to accept if organ quality θ is positive. The common prior, μ, is assumed to be negative and, hence, centers will only accept if they receive positive signals that shift their belief about an organ's quality above zero. In general, there is a threshold signal value (which is positive) above which centers accept. For centers in first position, this threshold is decreasing in ability because higher ability centers put more weight on their own (precise) signal, relative to the negative prior. This does not imply that higher ability centers accept more often; in general, they receive higher (lower) signals when the organ is of higher (lower) quality and thus they accept more (less) often when the organ pool is of superior (inferior) quality. Regardless of the quality of the organ pool, however, they always pass on a worse pool of organs (having disproportionately accepted the higher quality organs and rejected the lower quality organs).

The preceding argument implies that center 2 is more likely to reject if center 1 has higher ability, even if it makes its decision independently, simply because the organ pool will be of lower quality. Higher ability centers are more responsive to changes in the organ pool and, hence, the cross partial effect will be positive, mirroring the result that we obtained with binary organ types and signals. When centers herd, a given center will continue to be more responsive to a higher ability predecessor's rejection decision; as noted, the threshold signal value above which a higher ability center accepts in first position is lower and, hence, both the expected value and the variance of its signal (conditional on rejection) will be lower. However, the cross partial could now be negative. This is because lower ability centers in second position have less confidence in their own signals and thus put relatively more weight on their predecessors' signals, which they (imperfectly) recover from their rejection decisions.

Next, consider center 3's decision when it follows two centers with the same ability. Center 2 has a lower

prior than center 1 because it observes a previous rejection (by a center of the same ability). It thus requires a stronger positive signal to shift its belief to the point where it will accept the organ. This implies that the signal threshold is higher for center 2 than for center 1. Following the same argument as above, the expected value and variance of center 1's signal is lower than center 2's signal and, hence, center 3 will place more weight on its decision to reject. Both tests of herding that we have developed in the context of a binary signal structure thus extend to the polar opposite case with a continuum of signals and unbounded support.

4 Testing the Model

4.1 The Data

The data that we use to test the model consists of the sequence of decisions taken by centers for each deceased-donor organ (liver or kidney) that was offered for transplantation in the 2006-2015 period.[2] Each center that is offered an organ can either accept or reject it. If an organ is rejected, it is offered to the next center in line, unless NHSBT assesses that the condition of the organ has deteriorated to the point that it is no longer useable, in which case it is discarded, i.e. set aside for research. There are thus two possible end-points for an organ: it is accepted or it is discarded. Prior to either end-point, all decisions must necessarily be rejections.

The deterioration that results in an organ being discarded can be caused by delays in retrieving the organ (warm ischaemia) or by subsequent delays in transplantation (cold ischaemia). The U.K. is much more conservative with regard to organ deterioration than the U.S. and, hence, sequence lengths tend to be short. When sequence lengths are short, most decisions are concentrated in early positions, and this is what we observe in Figure 3. For livers, 35% of observations are in first position, with a steep decline in the fraction of decisions at higher positions. For kidneys, 40% of observed decisions are in first position, followed by an even steeper decline in the fraction of decisions at higher positions. There are relatively few decisions past the eighth position for either type of organ.

Another useful way to describe the data would be to plot the fraction of rejections and the fraction of discarded organs (conditional on rejection) at each position. We see in Figure 4 that organs are discarded as early as the first position, presumably because such organs enter the offering sequence in relatively poor condition. Discard rates remain fairly stable at higher positions, except for a spike at positions 7 and 8 for livers. By contrast, rejection rates, which start at around 60% in first position for both livers and kidneys, increase steadily with position. Notice that rejection rates are systematically lower for kidneys; this results in shorter sequence lengths, explaining the steeper decline in the fraction of organs by position that we documented for kidneys in Figure 3.

4.2 Center Ability

The measure of center ability that we use for the herding tests is based on the center's probability of rejection when in first position. Figure 5 plots this statistic for all transplantation centers in the United Kingdom,

[2] A new National Kidney Allocation Scheme was initiated in 2006 and a new National Liver Allocation Scheme was initiated in 2015. The analysis thus covers a period during which both livers and kidneys were allocated in a uniform manner.

Figure 3: Proportion of Observations by Queue Position

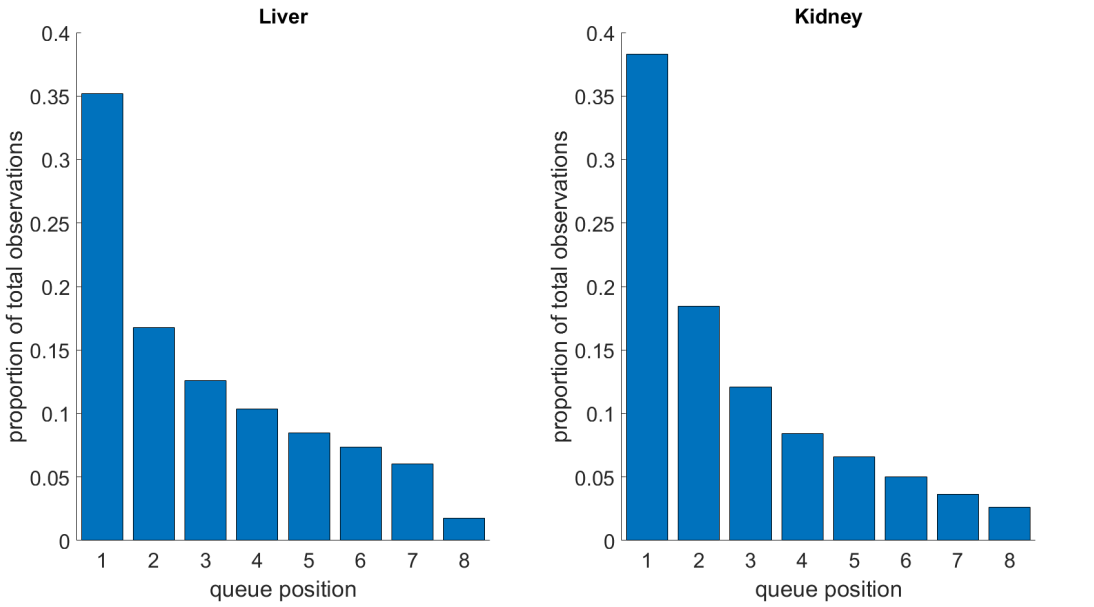
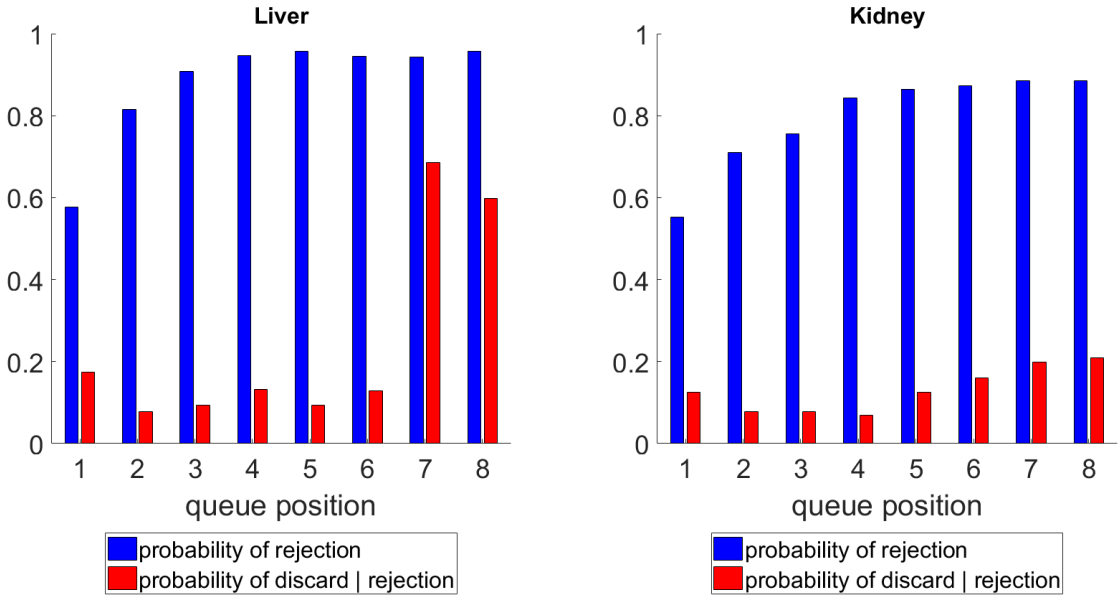
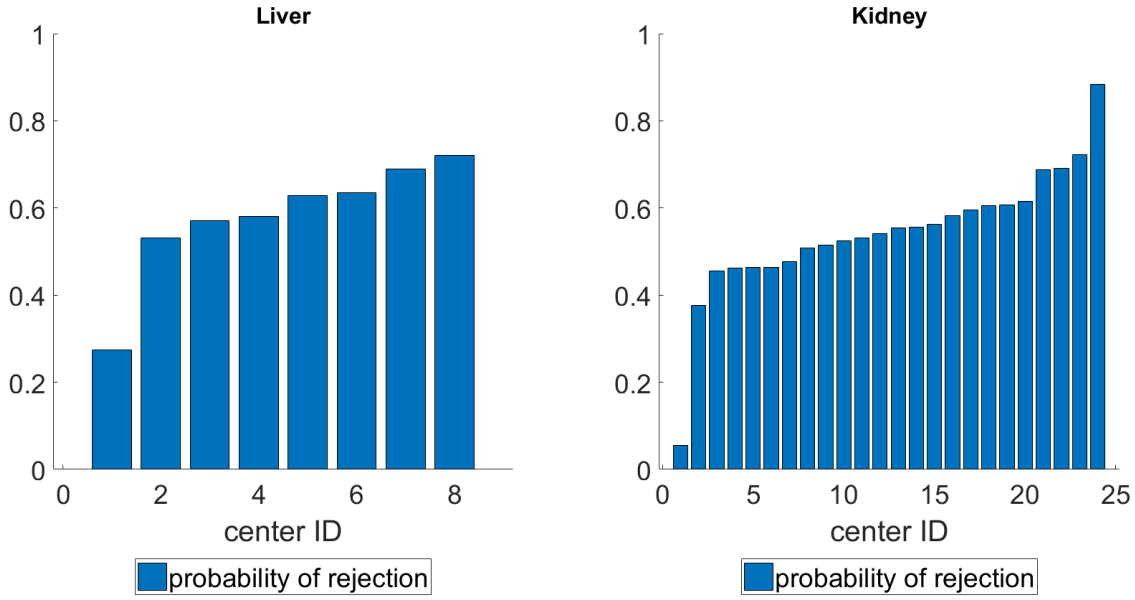


Figure 4: Probability of Rejection and Discarding by Queue Position



separately for livers and kidneys. There are eight liver transplant centers in total, and we see that their probability of rejection in first position ranges from 0.3 to nearly 0.8. This range is quite wide, even if we ignore one outlying center with a relatively low rejection probability. There are many more kidney transplantation centers (twenty-four), and here again we see wide variation in the probability of rejection. Ignoring two centers with relatively low or high probabilities, the rejection probabilities range from around 0.4 to 0.8. Thus, there appears to be substantial variation in our measure of ability across transplantation centers.

Figure 5: Rejections by Centers in First Position



As noted, the rejection probability in first position can be positively or negatively associated with center ability, depending on the sign of expression (7). We determine the sign of the expression by estimating the relationship between the probability of rejection in second position and center 1's overall probability of rejection when in first position: $p_2 = \alpha_0 + \alpha_1 \overline{p}_1$. If $\hat{\alpha}_1$ is positive (negative), then center ability is increasing (decreasing) in $\overline{p}_1$.

When estimating the preceding equation, the dependent variable is the decision in second position (reject=1, accept=0) for each organ that reaches that position. For a given center in first position, there is no variation in its decisions; they must necessarily be rejections for the organs to proceed to the second position. Nevertheless, that center will generate variation in the quality (G versus B) of the organs that it passes forward on account of the mistakes that it makes. While higher ability centers make fewer mistakes on average; i.e. they pass on a greater fraction of B organs, mistakes are independent across organs. Conditional on organ quality, signals received by center 2 are also independent across organs. It follows that the omitted (residual) term in the preceding estimating equation, for a given center 1, is i.i.d. By the same argument, the error term will be independent across organs with different centers in first position. Indeed, we could go

beyond the model and allow centers to incorporate recipient characteristics and the organ-recipient match in their decisions without altering the preceding argument, since these considerations will be independent across organs (even for the same center). The estimated standard errors do not require adjustment and this is also true, for the same reasons, for the tests of herding that follow.

Table 1: Measuring Center Ability

Dependent variable:	decision in second position			
Organ:	liver		kidney	
	(1)	(2)	(3)	(4)
$\bar{p_1}$	0.300***	0.513***	-0.281***	-0.412***
	(0.068)	(0.071)	(0.036)	(0.037)
Constant	0.652***	0.537***	0.888***	1.265***
	(0.041)	(0.076)	(0.022)	(0.042)
center 2 fixed effects	No	Yes	No	Yes
N	6383	6383	9257	9257

Note: standard errors in parentheses

Decision in second position: reject = 1, accept = 0

$\bar{p_1}$ measures center 1's overall probability of rejection in first position

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The relationship between the decision in second position (reject=1, accept=0) for each organ that reaches that position and center 1's probability of rejection when in first position, $\overline{p}_1$, is reported in Table 1, Columns 1 and 3, for livers and kidneys respectively. Results with an augmented specification that includes center 2 fixed effects, effectively assessing variation in position 2 decisions for a given center, are reported in Columns 2 and 4. The coefficient on $\overline{p}_1$ is positive and significant for livers, and negative and significant for kidneys. Centers in second position are unambiguously more likely to reject when they follow a higher ability center from expression (9). The results in Table 1 thus imply that center ability is increasing (decreasing) in the first-position probability of rejection for livers (kidneys). We measure center ability by this probability for livers and by one minus the probability for kidneys in the tests of herding that follow.

4.3 Testing for Herding

Our first test of herding is based on equation (2): $p_2 = \alpha_0 + \alpha_1 q_1 + \alpha_2 q_2 + \alpha_3 q_1 \cdot q_2$, where p_2 is the probability of rejection in second position and q_1, q_2 measure the ability of center 1 and center 2, respectively. As described in Test 1, we expect the cross-partial effect (i.e. the effect of the interaction term $q_1 \cdot q_2$) to be negative when each center's ability is increasing in its first-position rejection probability. Based on the preceding results, this will be the case with livers. In contrast, we expect the cross-partial effect to be positive for kidneys, as center ability is negatively associated with the first-position rejection probability.

Table 2 reports the estimated relationship between the decision in second position (reject=1, accept=0)

for each organ that reaches that position and center ability in first and second position, together with the interaction term. To interpret the estimated coefficients, it is convenient to normalize so that $\underline{q} = 0$. The coefficient on q_1 then applies to the case where q_2 equals zero. Without herding, a center with $q_2 = \underline{q}$ effectively chooses to accept or reject randomly, independently of q_1. However, with herding, the probability that these centers reject for sure is increasing in q_1. We find that β_1, the coefficient on q_1, is small and imprecisely estimated for kidneys and much larger and significant at the one percent level for livers. β_2, the coefficient on q_2, applies to the case in which q_1 equals zero. In this case, the first center's decision has no effect on the quality of the organ pool that is passed on and, moreover, does not increase the likelihood that the center that follows will herd. When $q_1 = \underline{q}$, center 2 effectively behaves as if it is in first position and β_2 corresponds to the intercept in Figures 1 and 2, which is increasing (decreasing) in q_2 for livers (kidneys). As predicted, the coefficient on q_2 is positive and significant for livers and negative and significant for kidneys. Our test of herding, however, is based on the interaction coefficient. As predicted by the model, the interaction coefficient is negative and significant for livers, and positive and significant for kidneys. By contrast, the interaction coefficient would be positive and significant for both livers and kidneys in the absence of herding.

Table 2: First Test of Herding (based on decisions in second position)

Dependent variable:	decision in second position	
Organ:	liver	kidney
	(1)	(2)
Center 1 ability (q_1)	2.135**	-0.0245
	(0.713)	(0.097)
Center 2 ability (q_2)	2.588***	-0.999***
	(0.688)	(0.102)
($q_1 \times q_2$)	-2.668**	1.003***
	(1.103)	(0.246)
Constant	-1.081**	0.977***
	(0.447)	(0.040)
N	6383	9257

Note: standard errors in parentheses.
Decision in second position: reject = 1, accept = 0
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.05$

Our second test of herding is based on the rejection probability of centers in third position, as specified in equation (3): $p_3 = \tilde{\alpha}_0 + \tilde{\alpha}_1 q_1 + \tilde{\alpha}_2 q_2$. As described in Test 2, we expect these decisions to vary more with first-position ability, q_1, than with second-position ability, q_2, when there is some amount of herding at position 2. The implicit assumption, which we verify below, is that centers always follow their signals in position 1. By contrast, decisions in third position would be equally responsive to q_1 and q_2 in the absence of herding (subject to the additional condition that $q_1 = q_2$).

Table 3 reports the estimated relationship between the decision in third position (reject=1, accept=0), for each organ that reached that position, and q_1, q_2. Center ability, for livers and kidneys, is measured as in Table 2. As predicted by the model when herding is present, the coefficient on q_1 is substantially larger than the coefficient on q_2; it is twice as large for livers and 50% larger for kidneys. The coefficients on q_1 and q_2 (β_1 and β_2 respectively) are imprecisely estimated for livers, and we cannot reject the hypothesis that $\beta_1 \leq \beta_2$. The corresponding coefficients for kidneys are, however, statistically significant; we can reject the hypothesis that $\beta_1 \leq \beta_2$ at the 5 per cent level.

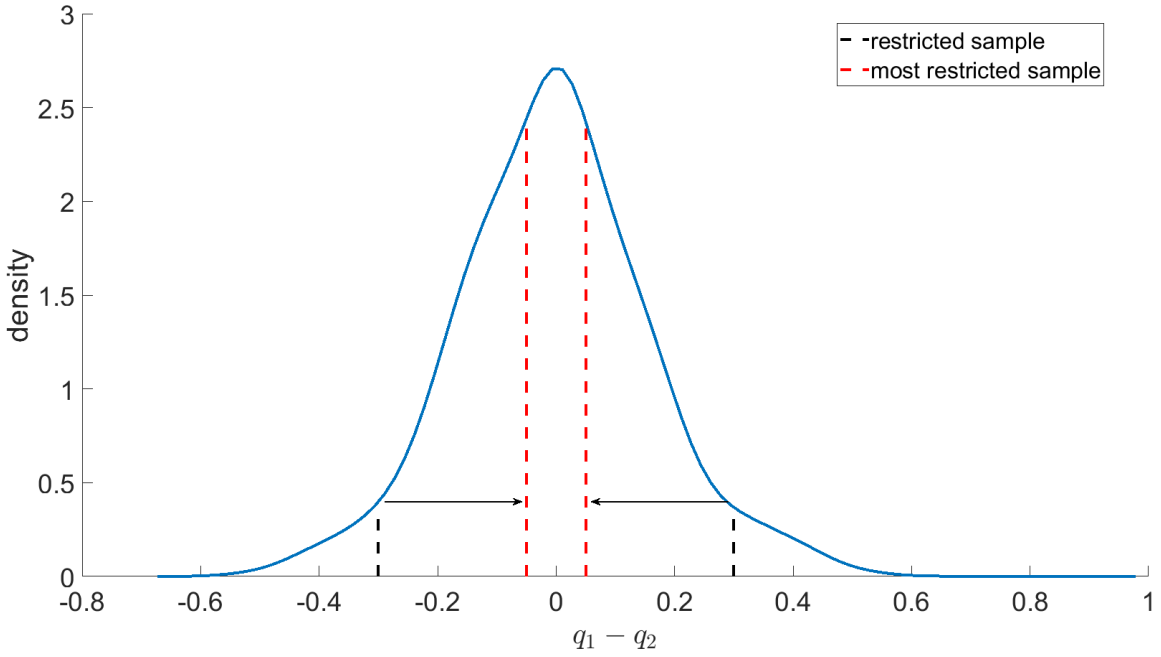
Table 3: Second Test of Herding (based on decisions in third position)

Dependent variable:	decision in third position	
Organ:	liver	kidney
	(1)	(2)
Center 1 ability (q_1)	0.104	0.352***
	(0.066)	(0.045)
Center 2 ability (q_2)	0.0529	0.220***
	(0.064)	(0.046)
Constant	0.820***	0.541***
	(0.067)	(0.024)
F-statistic ($\beta_1 \leq \beta_2$)	0.47	3.43
p-value	[0.247]	[0.032]
$\bar{q_1}$	0.60	0.40
$\bar{q_2}$	0.65	0.40
N	4819	6084

Note: β_1, β_2 are the coefficients on q_1, q_2, respectively
$\bar{q_1}$ and $\bar{q_2}$ denote the sample means of q_1 and q_2, respectively.
Standard errors in parentheses
Decision in third position: reject = 1, accept = 0
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The data requirements to implement the second test of herding are quite stringent: (i) A substantial fraction of centers should herd in second position. (ii) A substantial fraction of centers should *not* herd in third position (if they did, then variation in q_1, q_2 would have no consequence for their decisions). (iii) There should be substantial variation in decisions – accept versus reject – in third position for the test to have statistical power. Based on the structural parameter estimates, we will see below that conditions (i) and (ii) are satisfied for both livers and kidneys. The important difference between the two organ types is that by the third position, over 90% of decisions for livers are rejections. This lack of variation might explain why the coefficients on q_1 and q_2 are imprecisely estimated in Column 1. While livers are most useful for identifying

Figure 6: Ability Differential $(q_1 - q_2)$ Distribution



Note: sample includes all kidneys that reached third position.

Table 4: Second Test of Herding (restricted samples)

Dependent variable:	decision in third position			
$(q_1 - q_2)$ range:	[-0.30,0.30] (1)	[-0.20,0.20] (2)	[-0.10,0.10] (3)	[-0.05,0.05] (4)
Center 1 ability (q_1)	0.396***	0.427***	0.638***	1.020**
	(0.053)	(0.064)	(0.136)	(0.468)
Center 2 ability (q_2)	0.153**	0.132**	0.00181	-0.437
	(0.055)	(0.065)	(0.140)	(0.470)
Constant	0.554***	0.550***	0.513***	0.515***
	(0.025)	(0.027)	(0.039)	(0.045)
F-statistic $(\beta_1 \leq \beta_2)$	7.27	6.81	5.94	2.44
p-value	[0.004]	[0.004]	[0.007]	[0.059]
N	5603	5071	3069	1665

Note: standard errors in parentheses

Alternative samples restricted to kidneys within a pre-specified ability differential $(q_1 - q_2)$ range

Decision in third position: reject = 1, accept = 0

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

herding with the first test, we thus focus on kidneys for the second test.

Test 2 is derived for the case where centers at position 1 and 2 have equal ability; i.e. $q_1 = q_2$. Although average ability in first and second position, $\overline{q}_1$ and $\overline{q}_2$, respectively, are equal for kidneys in Column 2, the more stringent requirement to test the model is that these abilities should be equal for each organ. Figure 6 describes the distribution of the ability differential, $q_1 - q_2$, for all kidneys that reached at least third position (and are thus used for the second test of herding). Although the distribution is centered at zero, there is substantial variation in the ability differential statistic. Table 4 takes account of this variation in $q_1 - q_2$ by implementing the test of herding with an increasingly restricted sample of kidneys; i.e. by gradually narrowing the ability differential range. We see that the key result from Table 3, which is that the coefficient on q_1 is significantly larger than the coefficient on q_2 for kidneys, is retained as we reduce the sample. Indeed, this result is even obtained with the most stringent ability-differential restriction in Column 4, by which point the sample is just one-quarter of the full sample of kidneys.

5 Structural Estimation and Quantification

5.1 Estimation

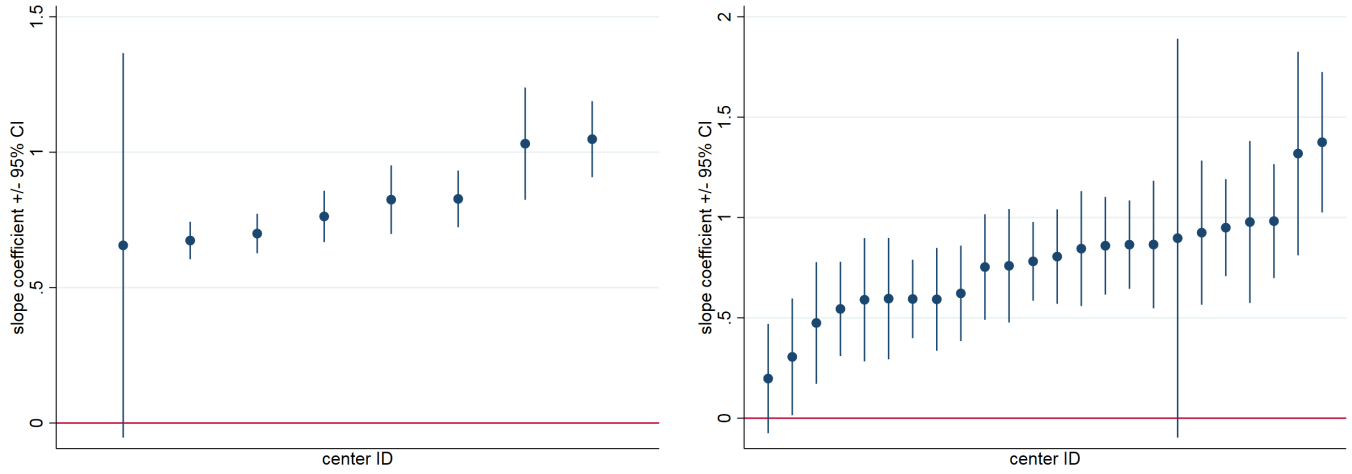
There are two types of organs in our model: good (G) and bad (B). Centers are heterogeneous in their ability to identify organ type; β_j is the probability that center j receives a b signal when the organ is bad and γ_j is the probability that it receives a g signal when the organ is good. The fraction of good organs in the population of organs is π and the cutoff belief (that the organ is good) above which centers accept an organ is $\tilde{\pi}$.

We begin by describing how β_j and γ_j are constructed. The reduced-form tests of the model are based on a single ability measure, which is derived from each center's probability of rejection in first position. The structural estimation, in contrast, is based on the complete specification of the model; i.e. with two ability measures. In order to construct separate measures of ability, β_j and γ_j, we take advantage of two indices of organ quality that have previously been developed specifically for the United Kingdom: the UK KDRI (Kidney Donor Risk Index) and the UK DLI (Donor Liver Index).

Indices of liver and kidney quality were first constructed in the United States, but have recently been adapted to the U.K. population. The UK KDRI is based on U.K. National Transplant Registry data covering over 7000 recipients who received deceased-donor kidneys between January 1, 2000 and December 31, 2007 (Watson et al., 2012). Various recipient and transplant factors were included in a model of transplant success, measured by patient survival, and the UK KDRI consists of those donor and organ characteristics that were found to be significant determinants of success (with optimal, estimated weights on each of those characteristics). More recently, data from all liver transplants from deceased donors between January 1, 2000 and December 31, 2014 have been used to construct the UK DLI (Collett et al., 2017). As with the UK KDRI, donor, recipient, and transplant data were used to identify factors associated with graft survival. Those donor and organ characteristics that were found to be significant determinants of transplant success, appropriately weighted, are included in the UK DLI.

We begin by using the probit model to estimate the relationship between the probability that an organ is rejected and its risk index, by center, restricting the sample to decisions that were made when centers were in first position (and therefore, by assumption, following their signals, as verified below). The risk indices were originally developed to aid centers in their decision-making, and a proposal to incorporate the UK KDRI into the National Kidney Offering Scheme was presented at the 2018 Blood and Transplantation Congress. At the time of writing, however, neither the UK KDRI nor the UK DLI, the latter of which was developed in 2017, are made available to transplant surgeons when they make their decisions. While centers thus may not have had explicit knowledge of the risk indices during the period of our analysis (2006-2015), we expect them to have put weight on many of the factors incorporated in the UK KDRI and the UK DLI. Figure 7 reports probit estimates of the Risk Index coefficient, with the corresponding 95% confidence interval, by center. As expected, the coefficient is positive and significant, almost without exception, both for livers and kidneys.

Figure 7: Center-Specific Probit Slope Coefficient Estimates



(a) liver (b) kidney

Note: estimates based on the relationship between the probability of rejection in first position and the risk index.

The risk indices are optimally constructed on the basis of outcomes generated by thousands of transplants over many years. This information is not available to transplant centers, who must base their decisions on past experiences with a limited set of outcomes. Higher ability centers will, nevertheless, make decisions that track more closely with the risk indices, and see in Figure 7 that there is substantial heterogeneity in the estimated slope coefficients. Although we do not explicitly map the risk index onto underlying organ types (G versus B), in the limit, extremely high levels of the risk index would correspond to a B organ, while extremely low levels would correspond to a G organ. For a given center j, the predicted probability of rejection at the top of the risk index distribution thus provides an estimate of β_j; i.e. the probability of receiving a b signal with a B organ. Similarly, at the bottom of the risk index distribution, the predicted probability of rejection provides an estimate of $1 - \gamma_j$; i.e. the probability of receiving a b signal with a G organ.

Higher ability centers will have larger probit slope estimates (associated with the ability to identify bad

Figure 8: Cross-Validating Ability Measures (intercept)

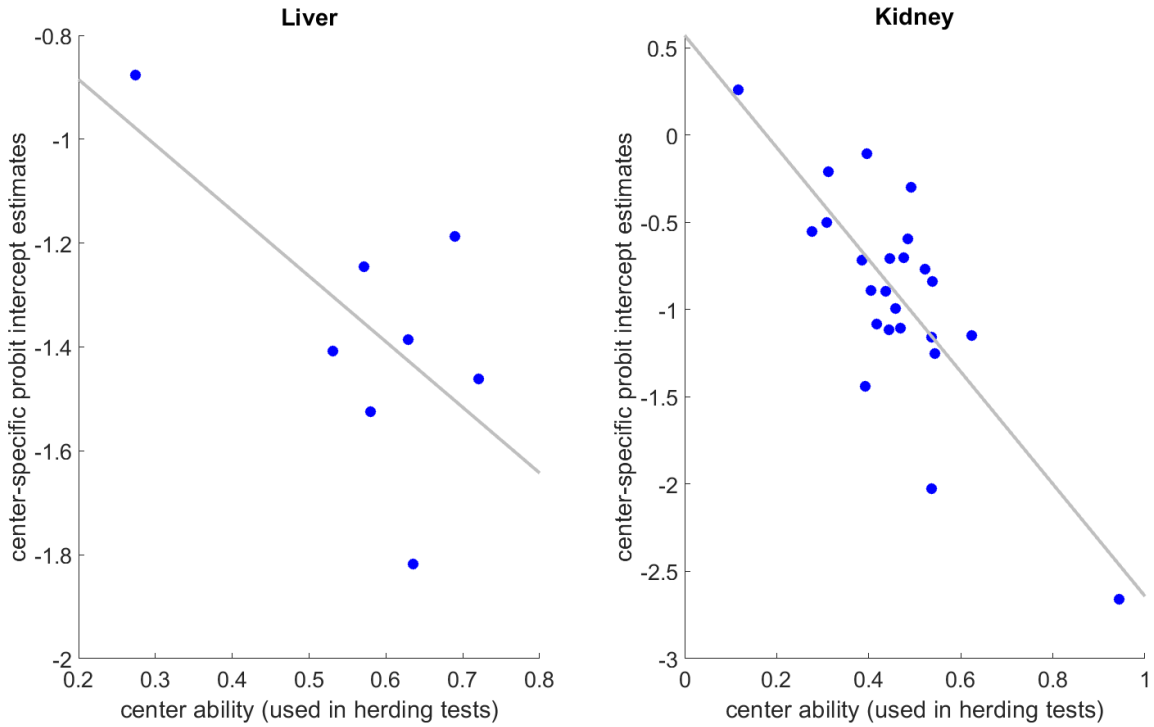
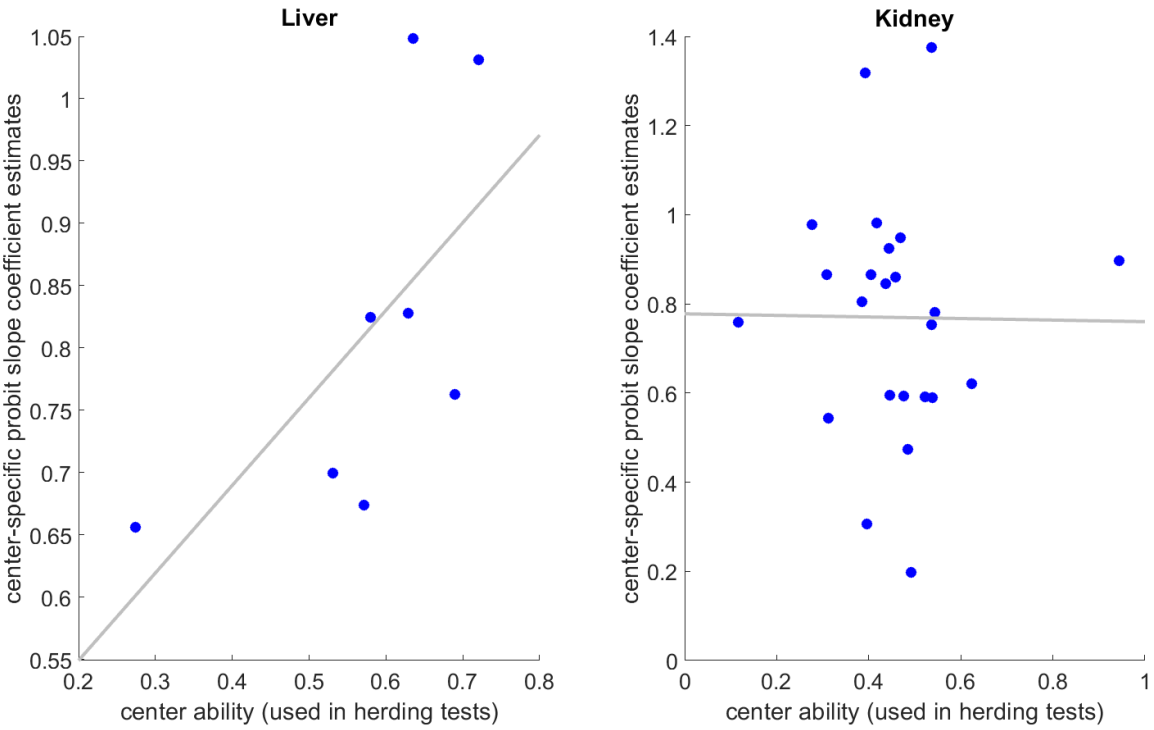


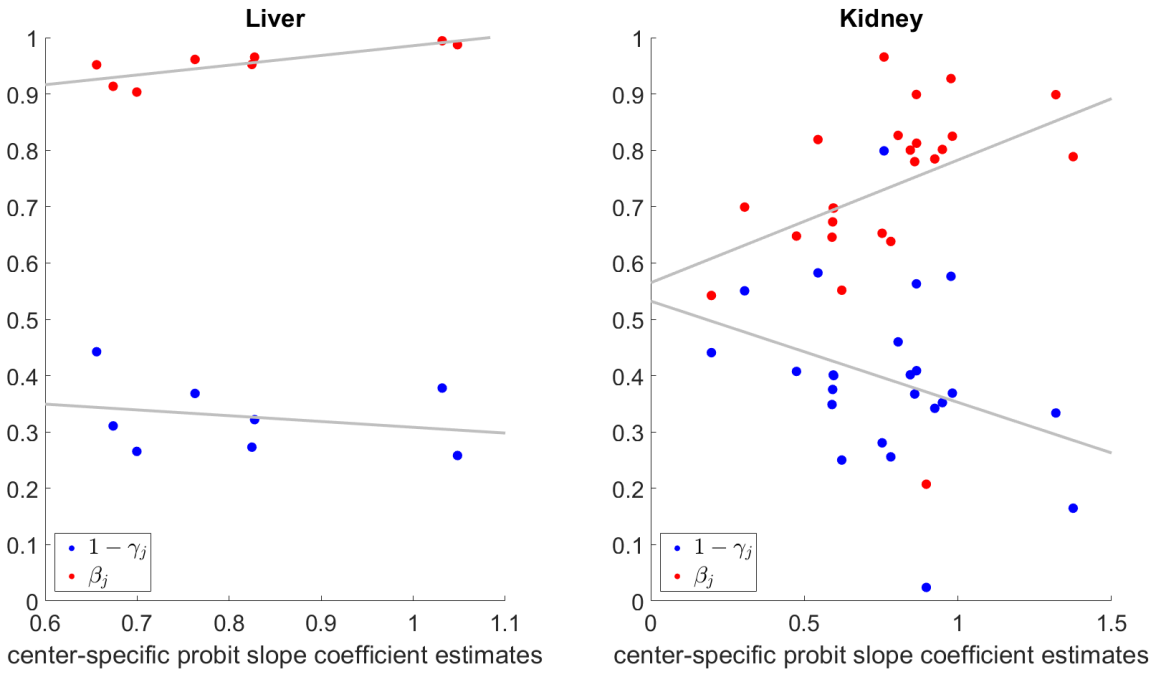
Figure 9: Cross-Validating Ability Measures (slope)



organs; i.e. larger β) and smaller intercepts (associated with the ability to identify good organs; i.e. smaller $1-\gamma$). We verify that this is the case by separately plotting the relationship between the center-specific slope coefficients and intercepts against the measure of ability that we constructed above for the tests of herding. We expect that the slope coefficient will be positively associated with the ability measure, while the intercept will be negatively associated with that measure. Cross-validating our independently constructed ability measures, we see in Figures 8 and 9 that that this is indeed the case (except that there appears to be little variation in the ability of centers to detect bad kidneys).

Figure 10 reports the estimated β_j, $1-\gamma_j$, for all centers, separately for livers and kidneys. These predictions are computed at the 95^{th} percentile and the 5^{th} percentile of the risk index distribution, respectively. Assumption 1 in the model states that $\beta_j > 1-\gamma_j$; i.e. that centers are not systematically misinformed. We see in Figure 10 that this assumption is satisfied for each center, both for livers and for kidneys.[3]

Figure 10: Estimates of Center Ability



The risk indices can also be used to compute π (the fraction of G organs in the population of organs). In constructing β_j and γ_j we assumed that an organ at the top of the risk index distribution is a B organ, while an organ at the bottom of the distribution is a G organ. At intermediate levels, the probability that an organ is a G organ will be decreasing in the risk index. We thus characterize decision-making in first position as follows:

$$p_j(R_i) = \pi_i(1-\gamma_j) + (1-\pi_i)\beta_j, \tag{17}$$

[3]Notice also that, while β_j is increasing and $(1-\gamma_j)$ is decreasing in the estimated risk index coefficient, the relationships are not monotone. This is because the β_j, γ_j measures that we construct are based on both the risk index coefficients – i.e. the slope and constant (intercept) terms of the estimated probit model. The slope alone cannot be used to construct our measures of ability.

where $p_j(R_i)$ is the probability that center j rejects organs with risk index R_i when in first position, while π_i is the probability that the underlying organ is of type G. $p_j(R_i)$ can be predicted for a given organ, with risk index R_i, for each center j with our β_j, γ_j estimates. Although, in principle, the π_i corresponding to a given R_i should be the same for all centers, noise in the estimated β_j and γ_j could generate some variation in practice. Our best estimate of π_i is thus the average across all centers.

The characterization of decision-making in equation (17) is not inconsistent with our description of decision-making in the model. In the model, each organ is of a particular type – G or B – and, while centers may not know this type, the signal that they receive (which is all that matters for decision-making) correctly reflects the underlying organ type. We do not observe the underlying organ type either, but the risk index allows us to estimate the probability that an organ is a G organ (given R_i). This probability, π_i, averaged over all risk indices, provides an estimate of π, the fraction of G organs in the population of organs.

Having constructed measures of center ability (β_j and γ_j) and estimated π, all that now remains is to estimate $\tilde{\pi}$, the cutoff belief that an organ is a G organ above which centers accept the offer. All centers follow their signal in first position in the model, which implies that their belief following a g (b) signal lies above (below) $\tilde{\pi}$. While some centers continue to follow their signals in later positions, others will start to herd (i.e. to reject offers regardless of whether they receive a g or a b signal). This is because their beliefs always lie below $\tilde{\pi}$. As $\tilde{\pi}$ increases, the fraction of centers that herd thus increases, with an accompanying increase in the rejection rate (the decisions of centers that follow their signals remain unchanged). To estimate $\tilde{\pi}$ we thus match the overall rejection rate in the data to the rejection rate predicted by the model; there is a unique value of $\tilde{\pi}$ at which the actual and predicted rejection rates match and this will be our best estimate of the $\tilde{\pi}$ parameter.

The simulated method of moments is used to estimate $\tilde{\pi}$. To draw signals for the estimation, we take advantage of the fact that our probit estimates (based on the risk index) allow us to predict the probability of rejection in first position for any organ-center pair. Since centers always follow their signals in first position, as verified below, this provides us with the probability that the center would receive a b signal in first position and, for that matter, in any position. We draw signals in this way, and then predict decisions at each position (given the previously-estimated values of β_j, γ_j, and π). The average over multiple draws of the signals predicts the overall rejection rate for a given $\tilde{\pi}$, and we then search over all $\tilde{\pi}$ to find the value at which the actual and predicted rejection rates match.

The data are effectively generated by a single draw from the signal distribution. Even if the model was correctly specified and the correct value of $\tilde{\pi}$ was selected by the econometrician, actual decisions and predicted decisions would evidently not match at each organ-position. However, as long as a large number of centers follow their signals, this sampling error will wash out and actual and predicted rejection rates, overall, will match when the correct $\tilde{\pi}$ is selected.[4] Table 5 reports $\tilde{\pi}$ estimates, with bootstrapped standard errors, separately for livers and kidneys. As we estimate a single parameter, we need only match on a single moment;

[4] A special feature of our data is that the order of centers who would have been approached, past the position where an organ is accepted or discarded, is unavailable. Thus, if the model predicts a rejection at the final position in a sequence where an organ was accepted in the data, then we can go no further. To be consistent, if the model predicts an acceptance at a position where an organ was rejected in the data, we proceed no further (and subsequent positions are not utilized for estimation).

we can utilize the rejection rate at any position for this purpose and the benchmark specification matches on the rejection rate in second position. As noted, the organ-recipient mismatch as well as organ deterioration are likely to be less relevant in the U.K. If they were, then centers would behave differently, using a different $\tilde{\pi}$, at higher positions. Reassuringly, we see in Table 5 that the $\tilde{\pi}$ estimate remains very stable when we match on additional positions (moments) up to position 5, with one exception (adding position 3 for estimation with kidneys).[5]

Table 5: Structural Parameter Estimates

Organ:	liver				kidney			
Number of matched moments:	one (1)	two (2)	three (3)	four (4)	one (1)	two (2)	three (3)	four (4)
$\tilde{\pi}$	0.87 (0.0038)	0.86 (0.0058)	0.86 (0.0055)	0.86 (0.005)	0.58 (0.0044)	0.52 (0.0018)	0.50 (0.0066)	0.50 (0.0015)
N	5029	3780	3109	2548	7691	5031	3508	2747

Note: when matching moments, we begin with the probability of rejection in second position, and sequentially add the corresponding probabilities in third, fourth and fifth positions.
Bootstrapped standard errors in parentheses.

As multiple positions (moments) are available for estimation, we might think it possible to simultaneously estimate both $\tilde{\pi}$ and π. To see why this is infeasible, however, suppose that we choose a value of π, but leave the estimated β_j and γ_j unchanged. The generated signals will no longer be organ-specific but, nevertheless, as π increases, the rejection rate among centers that follow their signals will decrease. To bring the overall rejection rate back in line with the data, $\tilde{\pi}$ must increase (and, with it, the fraction of centers who herd). It follows that, for every value of π, there exists a $\tilde{\pi}$ such that the actual and predicted rejection rates match. Therefore, $\tilde{\pi}$ and π cannot be estimated simultaneously, highlighting the important role played by the risk index in our analysis.

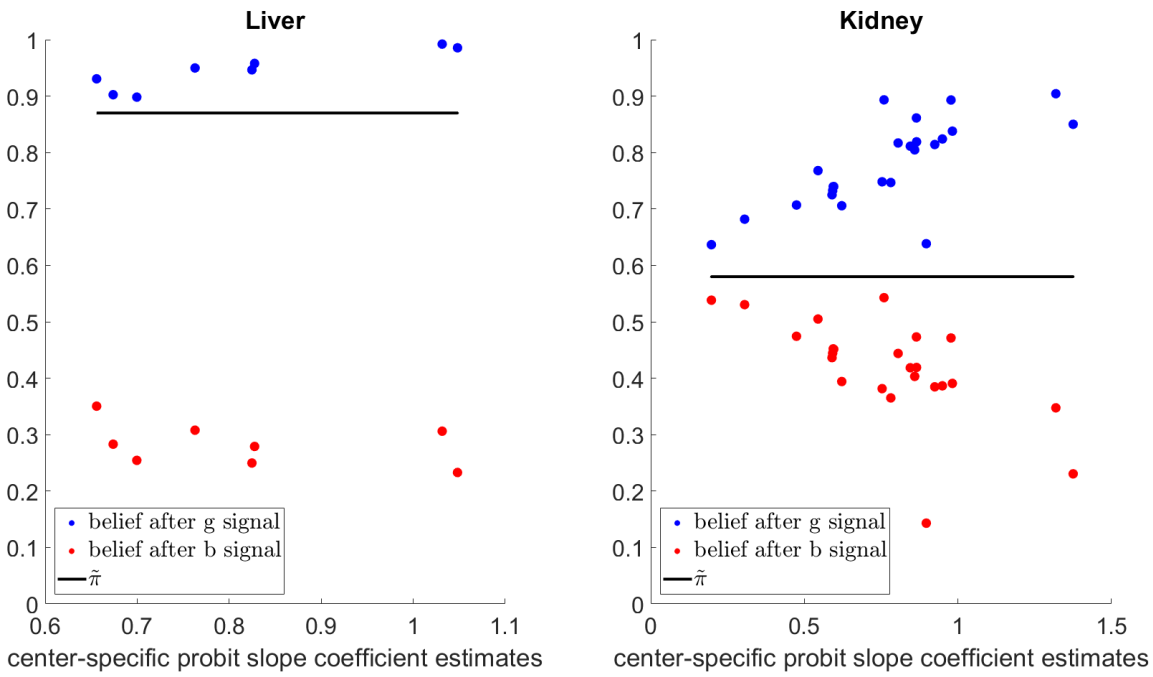
5.2 Validation and Goodness of Fit

As we need only estimate a single parameter, $\tilde{\pi}$, it is straightforward to establish that this parameter is identified - i.e. that a unique value exists at which the model best fits the data. This value must, however, also satisfy additional restrictions implied by the model. In particular, Assumption 2 requires that each center's belief that the organ is good must lie above (below) $\tilde{\pi}$ when it receives a g (b) signal in first position. This ensures that it follows its own signal. Figure 11 verifies that this important assumption is satisfied for each center, both for livers and for kidneys.

Having estimated the model and validated its key assumptions, the next step is to assess the model's goodness of fit with the data. As we use the overall rejection rate to estimate $\tilde{\pi}$, we begin by comparing

[5]When matching on multiple positions, we compute the error in the rejection rate at each position and then take the unweighted average across all positions to compute the overall error. Our best estimate of $\tilde{\pi}$ is the value that minimizes the overall error.

Figure 11: Updated Beliefs in First Position, by Center



the actual rejection rate and the predicted rejection rate, by position, separately for livers and kidneys. The benchmark $\tilde{\pi}$ estimate is obtained by matching rejection rates in second position, and all the results that follow are based on this estimate. Thus, we would expect a close match in second position, but not necessarily at higher positions. As a basis for comparison, we also report rejection rates from an alternative model with no learning. Center ability and the signal-generating process are determined in the same way as in our model, but decisions are now based exclusively on the signals received by each center (without regard to the decisions of preceding centers). We see in Figure 12 that rejection rates predicted by our herding model exceed the corresponding rates in the data, particularly at higher positions, while the alternative no-learning model systematically under-predicts rejection. The error associated with our model is not, however, substantial: overall, predicted rejection rates exceed actual rejection rates by just 3% for kidneys and 7% for livers

An alternative metric to compare the performance of these models is the fraction of correctly-predicted decisions – that is, acceptances (rejections) in the data that are predicted to be acceptances (rejections). Based on this metric, the herding model clearly out-performs the no-learning model, both for livers and for kidneys, as observed in Figure 13. This result complements the reduced-form tests of the model, providing independent evidence that herding is an important component of centers' decision-making process.

5.3 Quantification and Counter-Factual Simulations

While both the reduced-form tests and tests of the structural model against the no-learning alternative indicate that centers are herding, we would like to quantify the prevalence of this behavior. Given the estimates of β_j, γ_j, π, $\tilde{\pi}$ and the sequence of centers associated with each organ, we can determine whether

Figure 12: Goodness of Fit (Probability of Rejection)

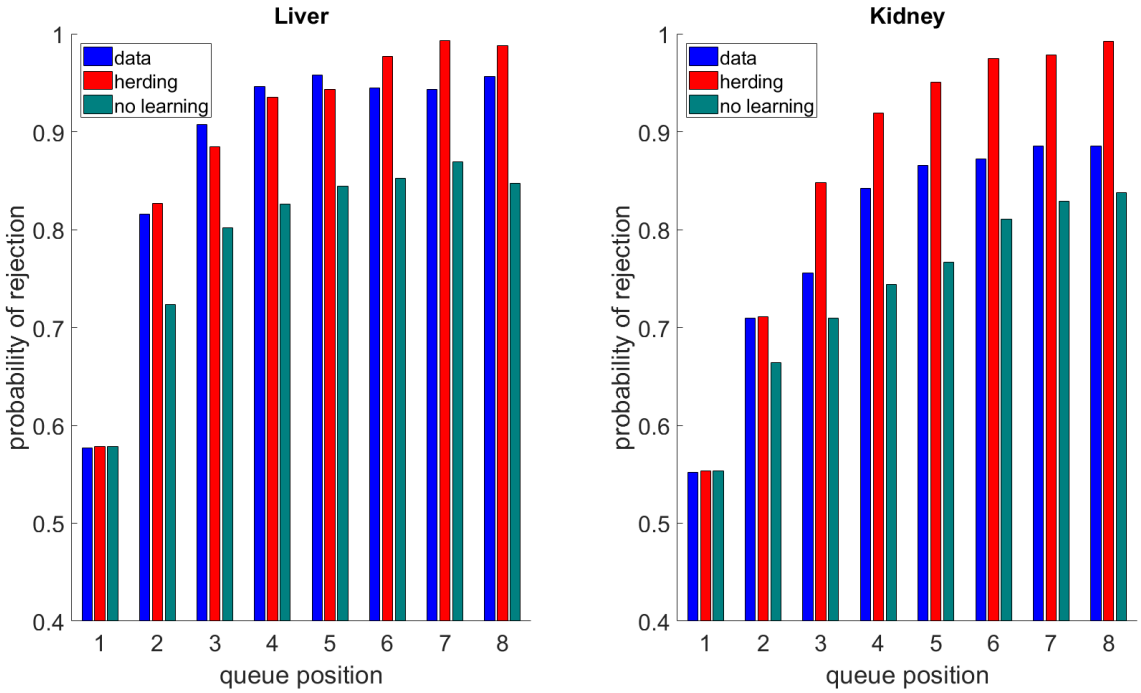
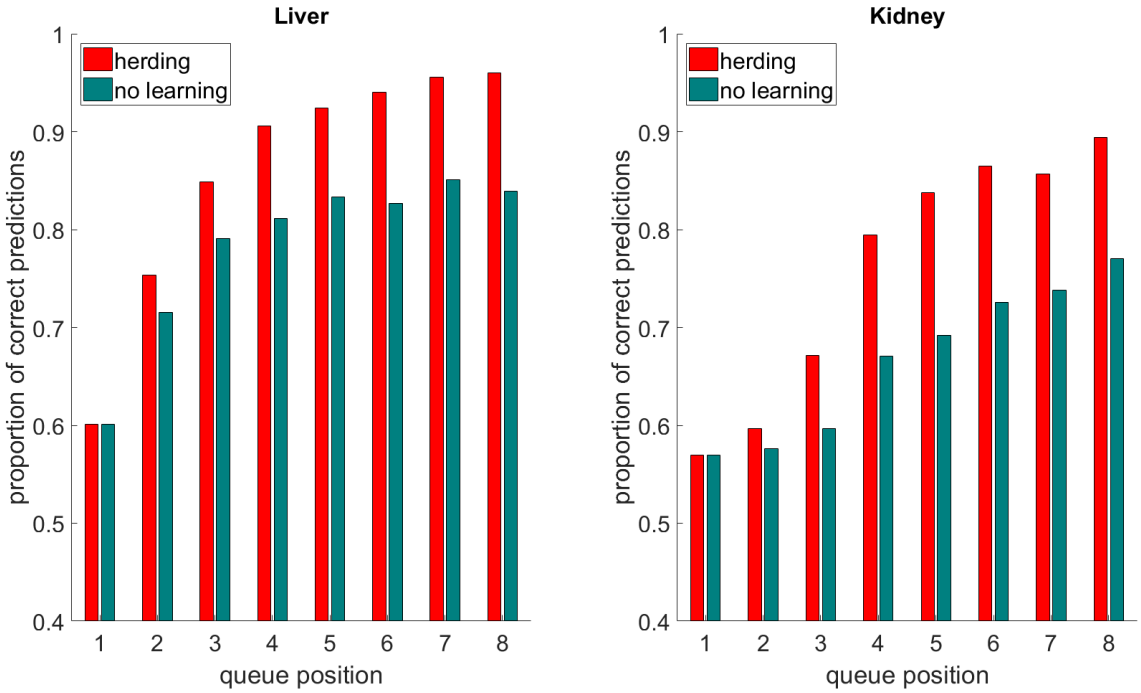
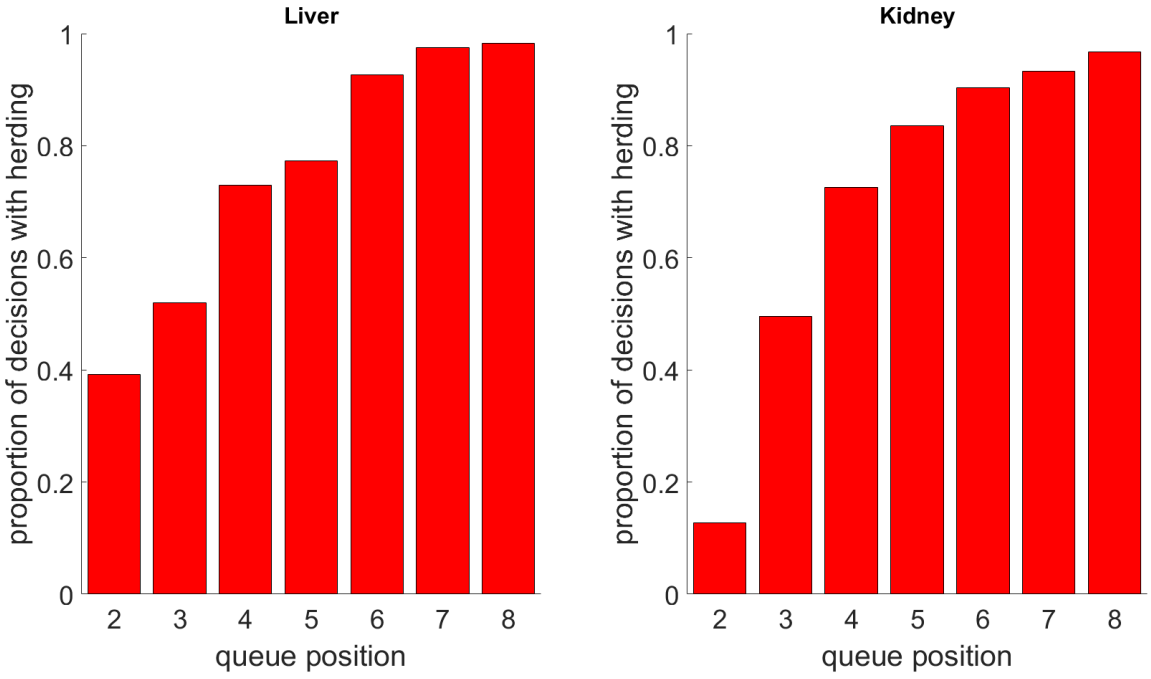


Figure 13: Goodness of Fit (Proportion of Correct Predictions)



a given center at a given position is herding – i.e. that its prior belief based on preceding decisions is so far below $\tilde{\pi}$ that it will reject regardless of the signal it receives. Figure 14 reports the prevalence of such herding, by position, for livers and kidneys. Herding is very common. For livers, about 40% of centers in second position herd. There is a steep increase in herding at higher positions and, by the sixth position, almost all centers herd. Herding is less prevalent, on average, for kidneys. Nevertheless, over 10% of centers in second position herd, there is a sharp increase to 50% at third position, and over 90% of centers herd by the sixth position.

Figure 14: Prevalence of Herding



Having measured the prevalence of herding in organ transplant decisions, we next consider the efficiency consequences of such behavior. We do this by comparing organ discard rates under the status quo (with herding) with discard rates under pooled information (where centers observe, and use, the signals of all their predecessors) and under no-learning (where centers rely exclusively on their own signals). The standard practice with counter-factual analyses is to compare outcomes predicted by the model with simulated outcomes under alternative scenarios. The complication that arises when predicting discards with our model is that the order of centers who would have followed, past the point where an organ is accepted or discarded in the data, is unavailable. Given that predictions are thus based on observed data and given that observed sequences consist exclusively of rejections, with the possible exception of an acceptance at the end of a sequence, the model will under-predict discards, and sequence lengths, even if it is correctly specified (by predicting acceptances in place of rejections just by chance).[6] Given the special structure of our data, we thus treat the

[6]Some acceptances in the data will be predicted by the model to be rejections, but this will be less frequent than reversals in the opposite direction since rejections are more common. Moreover, we cannot infer that an organ for which an acceptance in the data is reversed by the model will necessarily be subsequently discarded.

actual discard rate as the benchmark in the counter-factual analysis.

We begin by examining the decisions that underlie discard rates, under alternative scenarios. With pooled information, the signals received by centers who herd will be available to those that follow. Some of those signals will be g signals, in which case subsequent rejections, particularly by centers who herd (despite receiving g signals) could be reversed. Other unobserved signals will be b signals, in which case acceptances in the data could be reversed. Under the no-learning scenario, centers follow their own signals, ignoring the negative information contained in their predecessors' rejections. The resulting false acceptances are more frequent than with the herding model, where they only arise because b signals received by predecessors who herd (and reject) are ignored. False rejections can also arise with the no-learning model if a center that received a b signal and rejected would have reversed its decision if preceding signals were made available and utilized, and a sufficiently large number of those signals were g signals (these would have to be centers that were herding in the data). These are stringent conditions and, hence, false rejections are likely to be rare with no-learning.

To predict decisions with the counter-factual pooled information and no-learning models, we only draw signals where necessary. The presumption is that the estimated herding model correctly characterized the data. Thus, when a center accepts an organ, we assume that it must have followed its (g) signal. Similarly, when our model predicts that a center was following its signal, and it was observed to reject the organ, we assume that it received a b signal. It is only when our model predicts that a center was herding that we draw signals, using the organ's risk index and the center's ability to determine the probability that it received a b signal, as described above. Decisions generated by the pooled information model for each draw of the signals are compared with actual decisions at each organ-position, and the resulting discrepancy is then averaged over multiple draws to compute the fraction of false rejections and false acceptances with herding.[7] The same procedure is used to compare no-learning with pooled information, except that, with no-learning, centers always follow their signals.

Figure 15 reports the fraction of false rejections among all rejections by position in the herding and no-learning models, relative to the pooled information benchmark. As predicted, the proportion of false rejections is much greater for the herding model than under no-learning. Although the fraction of false rejections for the herding model is quite large at higher positions, especially for livers, most organ sequences are short. Thus, the efficiency consequences of the false rejections may not be substantial (as verified below). Figure 16 reports the fraction of false acceptances among all acceptances, comparing the herding model and the no-learning model to the pooled information benchmark. Once again, as predicted, we see in Figure 16 that false acceptances are more common with the no-learning model than with the herding model, beginning as early as position 2.

While the position-specific analysis described above paints a comprehensive picture of decision-making

[7] Observed decisions are determined by a single draw from an underlying signal-generating process. Thus, there will always be a discrepancy, due to sampling error, between the observed decision and the decision predicted by the model, based on repeated draws at a given organ-position (even when the model is correctly specified). This discrepancy washes out, however, when averaged over many organ-positions. By the same argument, we obtain a consistent estimate of the discrepancy between the alternative models under consideration as long as the number of organ-positions for which we draw signals is large. This will indeed be the case, due to the high prevalence of herding documented at each position in Figure 14.

Figure 15: False Rejections as a Proportion of All Rejections (relative to pooled information benchmark)

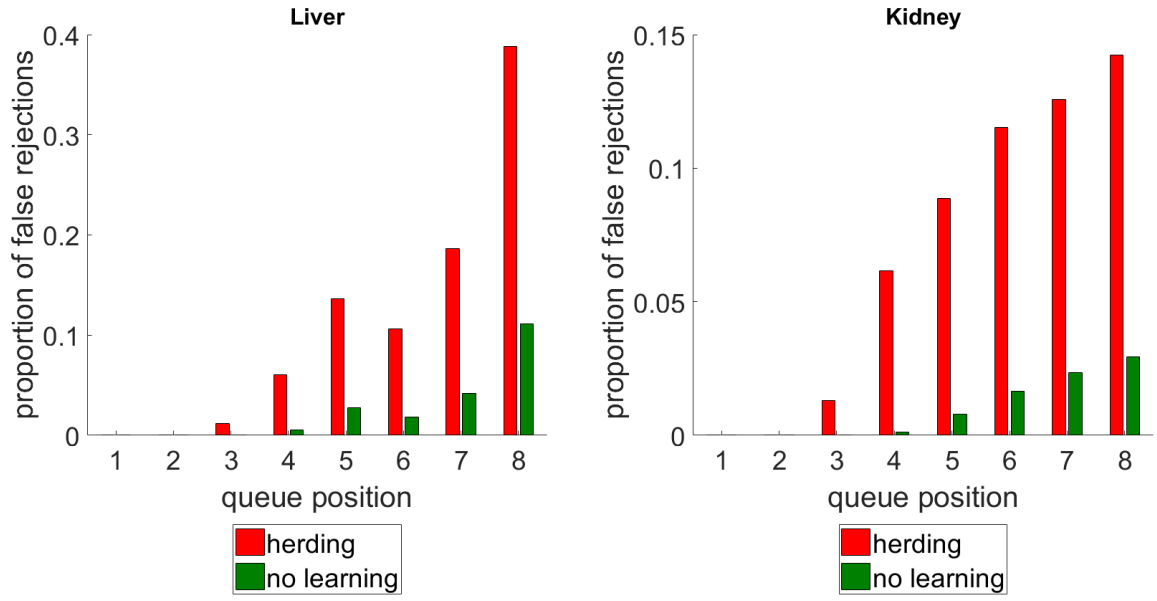
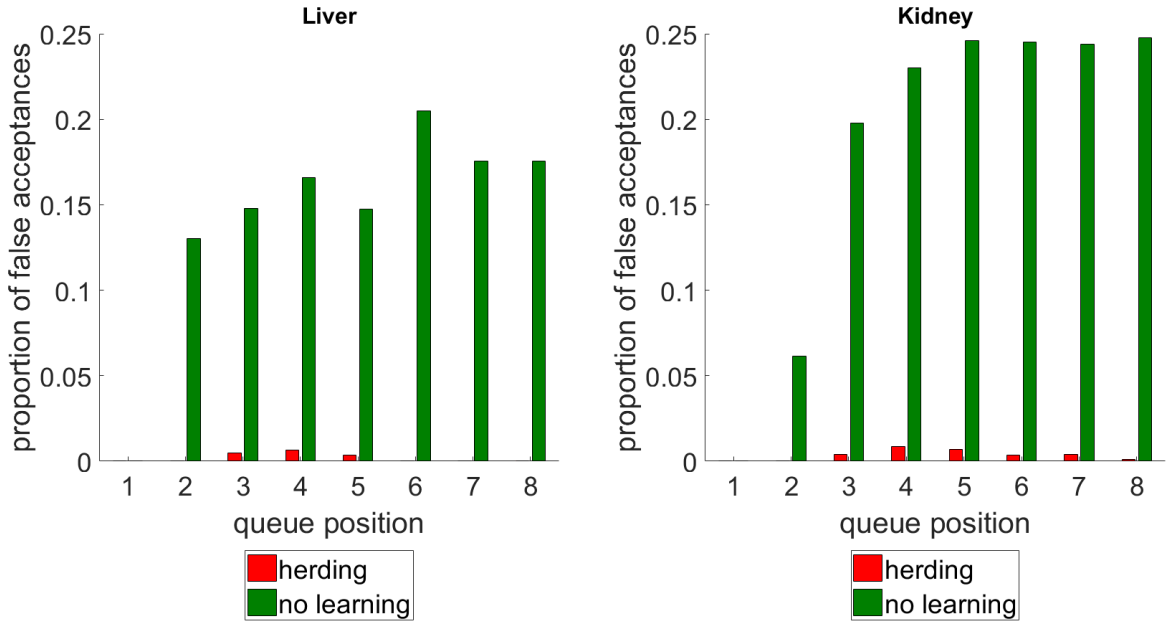


Figure 16: False Acceptances as a Proportion of All Acceptances (relative to pooled information benchmark)



under different models relative to the pooled information benchmark, the important consideration from a social welfare perspective is that good organs are accepted and bad organs discarded. For example, the case in which a center herds behind its predecessors and rejects a good organ is costly for its patient. There is, however, no welfare loss as long as patients are treated interchangeably and the organ is accepted further down the line. We thus complete the analysis by comparing discard rates with pooled information and no-learning against the data.

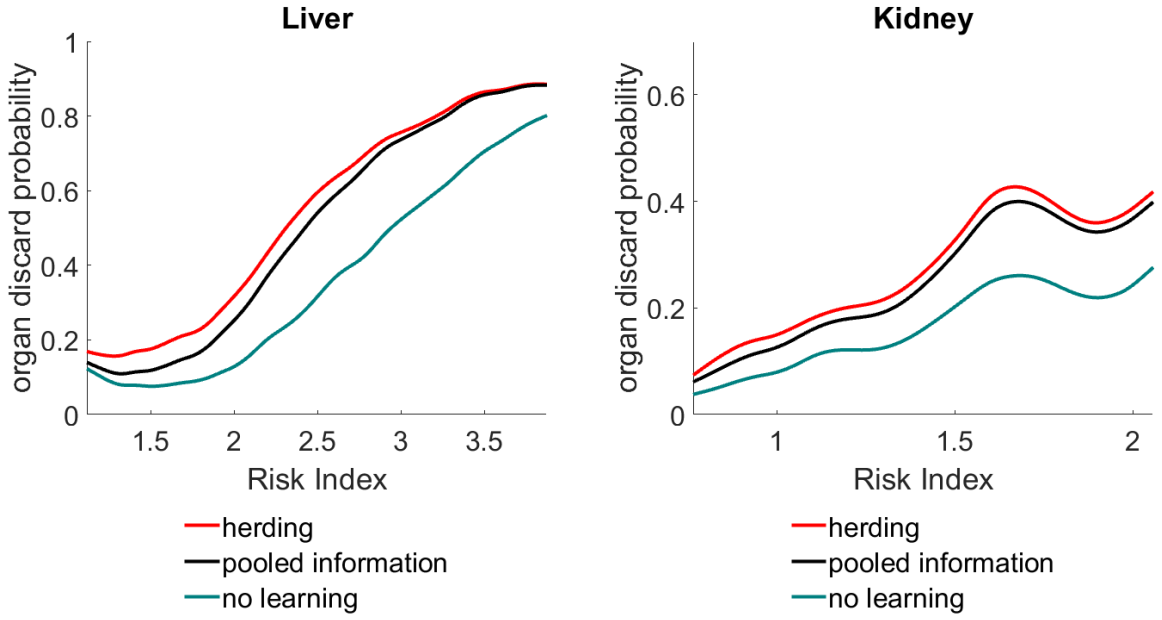
In our data, an organ is either accepted for transplantation or discarded (set aside for research). The decision to discard an organ is taken by NHSBT and is based on its usability, which depends, in turn, on its condition (which is distinct from its quality). Thus, the discard decision is treated as exogenously determined in our analysis. If we observe that an organ has been discarded in the data, but one of the alternative models (no-learning or pooled information) predicts that it would have been accepted at one or more earlier positions, we assume that the alternative model would have generated an acceptance for that organ. If the alternative model does not predict an acceptance at any position, then the organ is assumed to have been discarded. By the same logic, if an organ was accepted in the data and an alternative model predicts that it would have been accepted at one or more positions up to the point at which it was accepted, then the organ is assumed to have been accepted with the alternative model. Finally, if an organ was accepted in the data, but an alternative model does not predict an acceptance at any position up to that point, then we assume that it would have been accepted further down the line.[8]

Figure 17 reports the discard rate in the data (with herding), for the pooled information model, and for the no-learning model, across the range of risk indices and separately for livers and kidneys. The herding model and the pooled information model differ only when there are false rejections and false acceptances in the data. As observed in Figures 15 and 16, the former is much more likely. Nevertheless, discard rates with the herding model are only slightly higher than those in the pooled information model. This contrasts with discard rates for the no-learning model, which are substantially lower than those in the pooled information model. This is on account of the false acceptances in the no-learning model that were documented in Figure 16, which appear to be concentrated at higher risk indices, precisely where they are most dangerous. With organ transplantation, herding appears to protect centers from costly false acceptance decisions, without substantially raising the discard rate through false rejections.

The gain in welfare from a particular organ transplant program depends on both the number of organs that are transplanted and the subsequent transplantation outcomes. Table 6 reports discard rates and the average risk index of accepted organs, which is associated with worse outcomes on average, for the herding model, the pooled information model, and the no-learning model. Although centers did not have access to the risk indices when they made their decisions historically, we also examine the consequences of making this superior source of information available to them. Under the assumption that the risk index contains more information about an organ than is available to any individual center, all centers would now base their decisions on the risk index and ignore their own signals (assessments). The first center in line would accept the organ if its belief that the organ was a good organ exceeded $\tilde{\pi}$. If not, it would reject, and all centers

[8] We make this assumption in order to generate a lower bound on the discard rate for the pooled information model. Nevertheless, we will see that the discard rate predicted by this model is not too far below what we observe in the data.

Figure 17: Organ Discard Rate



that followed would do the same.[9]

Table 6: Welfare Analysis

	livers		kidneys	
	average risk index	discard rate	average risk index	discard rate
Herding	1.75	0.40	1.21	0.22
Pooled information	1.75	0.36	1.21	0.20
No learning	1.86	0.25	1.23	0.13
Risk index made available	1.72	0.37	1.19	0.15

We see in Table 6 that the average risk index of accepted organs is the same with herding and pooled information (up to two decimal places) and substantially higher with no-learning, particularly for livers. Making the risk index available to all centers would result in the selection of higher quality (lower risk) organs on average, both for livers and kidneys. Turning to the second ingredient in the welfare analysis, the discard rate relative to the status quo (herding) would decline by 4 percentage points for livers and 2 percentage points for kidneys with pooled information. However, much greater gains would be realized by making the risk index available to all centers. In particular, the discard rate would decline from 0.22 to 0.15 for kidneys, which is almost as low as with no-learning, despite the fact that a superior (safer) organ pool is being selected.

[9]The belief that an organ is a good organ, π_i, given risk index, R_i, is determined from equation (17). In principle, it should be the same for all centers.

6 Conclusion

We develop tests to detect information-based herding in environments where prospective buyers, acting sequentially, must choose whether or not to acquire an object. These tests are applied to organ transplant decisions in the United Kingdom. Our results indicate that herding, with centers ignoring their own assessment of the organ and simply following their predecessors, is common and contributes substantially to the high rate at which organs (livers and kidneys) are discarded in the data. However, this increase in the discard rate is not necessarily inefficient. The reliance by centers on their predecessors' decisions actually protects them from accepting bad organs.

In our research setting, prospective buyers (transplant centers) are experts at distinguishing between good and bad organs. Moreover, within this highly trained pool, centers are more likely to herd behind more able predecessors who are better than they are in detecting bad organs. It is consequently not surprising that the usual pathologies associated with herding are absent. In other environments, where buyers are less informed, potential gains from trade could fail to be realized when agents herd.

Counter-factual analyses, based on the estimated structural parameters of the model, indicate that making private assessments of organ quality publicly available would have little impact on both the discard rate and the quality of accepted organs. An alternative, easily implementable, policy that is effective on both dimensions would make recently developed indices of organ quality available to all centers. These indices contain more information about an organ's quality than the private assessment of any individual center and their utilization would reduce discard rates for kidneys, in particular, by over 30%, while simultaneously improving the quality of accepted organs.

References

Agarwal, Nikhil, Itai Ashlagi, Michael A Rees, Paulo J Somaini, and Daniel C Waldinger, "An empirical framework for sequential assignment: The allocation of deceased donor kidneys," Technical Report, National Bureau of Economic Research 2019.

Anderson, Lisa R and Charles A Holt, "Information cascades in the laboratory," *The American Economic Review*, 1997, pp. 847–862.

Banerjee, Abhijit V, "A simple model of herd behavior," *The Quarterly Journal of Economics*, 1992, *107* (3), 797–817.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, 1992, *100* (5), 992–1026.

Bohren, J Aislinn, "Informational herding with model misspecification," *Journal of Economic Theory*, 2016, *163*, 222–247.

Çelen, Boğaçhan and Shachar Kariv, "Distinguishing informational cascades from herd behavior in the laboratory," *American Economic Review*, 2004, *94* (3), 484–498.

Cipriani, Marco and Antonio Guarino, "Estimating a structural model of herd behavior in financial markets," *American Economic Review*, 2014, *104* (1), 224–51.

Collett, David, Peter J Friend, and Christopher JE Watson, "Factors associated with short-and long-term liver graft survival in the United Kingdom: development of a UK donor liver index," *Transplantation*, 2017, *101* (4), 786–792.

Conley, Timothy G and Christopher R Udry, "Learning about a new technology: Pineapple in Ghana," *American Economic Review*, 2010, *100* (1), 35–69.

Dupas, Pascaline, "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment," *Econometrica*, 2014, *82* (1), 197–228.

Foster, Andrew D and Mark R Rosenzweig, "Learning by doing and learning from others: Human capital and technical change in agriculture," *Journal of Political Economy*, 1995, *103* (6), 1176–1209.

Frick, Mira, Ryota Iijima, and Yuhta Ishii, "Misinterpreting Others and the Fragility of Social Learning," 2019.

Gale, Douglas, "What have we learned from social learning?," *European Economic Review*, 1996, *40* (3-5), 617–628.

Golub, Benjamin and Evan Sadler, "Learning in social networks," *The Oxford Handbook of the Economics of Networks*, 2016.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior," *The American Economic Review*, 1995, pp. 1088–1105.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, "The impact of institutional trading on stock prices," *Journal of Financial Economics*, 1992, *32* (1), 23–43.

Munshi, Kaivan, "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of Development Economics*, 2004, *73* (1), 185–213.

Smith, Lones and Peter Sørensen, "Pathological outcomes of observational learning," *Econometrica*, 2000, *68* (2), 371–398.

Vives, Xavier, "Social learning and rational expectations," *European Economic Review*, 1996, *40* (3-5), 589–601.

Watson, Christopher JE, Rachel J Johnson, Rhiannon Birch, Dave Collett, and J Andrew Bradley, "A simplified donor risk index for predicting outcome after deceased donor kidney transplantation," *Transplantation*, 2012, *93* (3), 314–318.

Watson, CJE and JH Dark, "Organ transplantation: historical perspective and current practice," *British Journal of Anaesthesia*, 2012, *108* (suppl_1), i29–i42.

Wermers, Russ, "Mutual fund herding and the impact on stock prices," *the Journal of Finance*, 1999, *54* (2), 581–622.

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