Market Failure in Kidney Exchange*†

Nikhil Agarwal, Itai Ashlagi, Eduardo Azevedo,
Clayton R. Featherstone, and Ömer Karaduman‡

June 27, 2018

Abstract

We show that kidney exchange markets suffer from traditional market failures that can be fixed to increase transplants by 25%-55%. First, we document that the market is fragmented and inefficient: most transplants are arranged by hospitals instead of national platforms. Second, we propose a model to show two sources of inefficiency: hospitals do not internalize their patients' benefits from exchange, and current mechanisms sub-optimally reward hospitals for submitting patients and donors. Third, we estimate a production function and show that individual hospitals operate below efficient scale. Eliminating this inefficiency requires a combined approach using new mechanisms and solving agency problems.

JEL: D47, D42, L11
Keywords: Kidney Exchange

*First version: September 20, 2017. We thank Alex Garza, Sarah Taranto, and Jennifer Wainwright of UNOS; Jonathan Kopke and Michael Rees of APD; Cathi Murphey and Adam Bingaman of Methodist at San Antonio; and Garet Hill of NKR for their expertise and access to data. We thank Glenn Ellison, Amy Finkelstein, Parag Pathak, Al Roth, Jean Tirole, Utku Ünver, Mike Whinston, and seminar participants for helpful discussions. Abigail Ostriker provided excellent research assistance. Azevedo acknowledges support from Wharton's Dean's Research Fund, Agarwal from the NSF (SES-1729090), and Ashlagi from the NSF (SES-1254768).

†Disclaimer: The data reported here have been supplied by UNOS as the contractor for the Organ Procurement and Transplantation Network (OPTN). The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy of or interpretation by the OPTN or the U.S. Government.

‡Agarwal: MIT and NBER, agarwaln@mit.edu. Ashlagi: Stanford University, iashlagi@stanford.edu. Azevedo: Wharton, eazevedo@wharton.upenn.edu. Featherstone: Wharton, claytonf@wharton.upenn.edu. Karaduman: MIT, omerk@mit.edu.

1 Introduction

The kidney exchange market in the US enables approximately 800 transplants per year for kidney patients that have a willing but incompatible live donor. Exchanges are organized by matching these patient-donor pairs in swaps to enable transplants. Each such transplant extends and improves the patient's quality of life and saves hundreds of thousands of dollars in medical costs, ultimately creating an economic value that is estimated at more than a million dollars.[1] Since monetary compensation for living donors is forbidden and deceased donors are in increasingly short supply, kidney exchange markets play an important role in mitigating the shortage of organs available for transplant.[2]

This paper shows that, despite significant success, the kidney exchange market suffers from market failures that result in hundreds of lost transplants per year. Our descriptive evidence shows that the market is fragmented and operates inefficiently. The inefficiency arises from two standard market failures. First, kidney exchange platforms use inefficient mechanisms: hospitals are not rewarded for submitting high social value patients and donors to the platform. Second, there are significant agency problems: hospitals face most of the costs of participating in national platforms but receive only a fraction of the benefits. Both of these problems give hospitals inefficient incentives, which result in kidney exchange taking place at an inefficiently small scale. These market failures are serious, but fixable. We show how to combine theory and data to design efficient mechanisms, and discuss policies that address the agency problems. Our estimates suggest that fixing these problems would generate between 200 to 440 additional transplants per year (25% to 55% of the current total).

Our argument has three parts. First, we use administrative datasets to show that the market is fragmented, inefficient and shows signs of agency problems. Second, we develop a simple model to explain the market failures and propose solutions. Third, we combine the model and data to estimate the magnitude of the inefficiencies and to design practical alternative mechanisms and policies.

The first part documents three key facts using data on all transplants in the United States and proprietary data from the three largest kidney exchange platforms. First, the market is highly fragmented. Instead of most transactions being arranged by a few large platforms,

[1]Kidney exchange is amongst a handful of recent innovations that clearly improve health care delivery while savings costs (see Chandra and Skinner, 2012, for a perspective on trends in health care costs in the US). Transplantation roughly doubles the life expectancy of patients with end-stage kidney disease, and is cheaper than the alternative treatment of dialysis. Medicare provides nearly universal coverage, irrespective of age, for patients with End-Stage Renal Disease. This program comprises of about 7% of Medicare's annual budget (see USRDS, United States Renal Data System, 2016). The cost savings of transplantation relative to dialysis alone have been estimated to be over $270,000 (see Section 2).

[2]There are over 97,000 patients currently waiting for a kidney from a deceased donor, but less than a fifth are expected to be transplanted in the next year. Becker and Elias (2007) argue that this waitlist could be completely eliminated if there was monetary compensation for live donors and advocate for the creation of this market. However, this type of transaction is widely panned by bioethicists, and almost all countries forbid monetary compensation for organs. The National Organ Transplantation Act prohibits compensating donors in exchange for acquiring organs in the United States. The motivation for kidney exchange is to use donor swaps to help patients find an organ in an ethically and legally acceptable way (Roth, 2007).

62% of kidney exchange transplants involved patients and donors from the same hospital. Second, we find direct evidence of inefficient exchanges in the market. Kidney exchanges performed within hospitals often transplant kidneys from easy-to-match donors to easy-to-match patients. Existing theory has shown that these matches are inefficient (Roth et al., 2007). Third, hospital behavior is inconsistent with pure maximization of patient welfare. Many hospitals do not participate in national platforms. In addition, evidence suggests that hospitals are sensitive to financial and administrative costs of participating in kidney exchange, even though these costs are small relative to the social value of transplants. Even when hospitals do participate, the typical hospital does not conduct all kidney exchanges through a national platform. Instead, most hospitals continue to operate as competing small kidney exchange platforms.

The second part develops a model to explain these facts and design policy responses. Although kidney exchange markets do not directly use monetary incentives to acquire organs, we can analyze them with standard neoclassical producer theory. A kidney exchange platform produces a final good (transplants) from intermediate goods (submissions of patients and donors) supplied by a competitive fringe (hospitals) according to a production function. This model is motivated by three key institutional features. First, hospitals are the key decision-makers steering participants towards kidney exchange (Roth et al., 2005; Ashlagi and Roth, 2014; Rees et al., 2009). Second, due to biological compatibility constraints, some patients and donors generate considerably more transplants than others when they join a platform. Third, the structure of optimal matches make transplants a natural numeraire good. Platforms can effectively transfer transplants from one hospital to another by choosing which hard-to-match patients to transplant.

Much of the economics of kidney exchange markets is determined by the shape of the production function. Returns to scale determine whether it is efficient to match patients in large platforms, or whether a fragmented market can be efficient. Marginal products determine the value of different types of patients and donors to the platform, which is a key determinant of efficient mechanisms.

Theorem 1 shows that inefficiency comes from the two market failures we discussed. The first market failure is that platforms use inefficient mechanisms. When a hospital submits a patient or a donor to a platform, current mechanisms reward hospitals according to the *probability* with which that hospital's patient is matched. But the theorem shows that, to maximize hospital welfare, hospitals should be rewarded with the *marginal product* (the expected number of additional transplants enabled by the submission) of their submissions plus a small adjustment term. Because existing platforms do not reward hospitals based on the marginal product of their submissions, even a hospital that maximizes the number of its own patients that are transplanted has to perform socially inefficient matches. This problem can be addressed by using points mechanisms that reward hospitals according to marginal products. The second market failure is that hospital objectives may differ from pure social welfare maximization, which we refer to as an agency problem. For example, hospitals may participate too little in kidney exchange because they face most of the costs but only receive a fraction of the benefits. This problem can be addressed with subsidy policies and mandates.

The third part combines the theory and data to quantify the inefficiency in the market and to suggest policy responses. To do so, we estimate the production function using administrative data from the largest kidney exchange platform and detailed information on matching algorithms and operational procedures.

The estimated production function yields three sets of results. First, we measure the returns to scale of the production function and estimate the inefficiency from market fragmentation. We find that the largest kidney exchange platform is well above the minimum efficient scale. At the same time, almost all single-hospital hospital platforms are far below the efficient scale. We estimate that the gains from moving all the production to the efficient scale is at least 200 transplants per year, and likely closer to 400. These improvements correspond to an economic value of between \$220 and \$440 million annually, of which approximately a quarter is due to savings on healthcare costs. Thus, consistent with the descriptive evidence and the shape of the production function, fragmentation has a large efficiency cost.

Second, we use the estimated production function to design more efficient mechanisms. Optimal mechanisms should reward submissions approximately according to marginal products, while current mechanisms reward submissions according to probabilities of matching. We find that marginal products are considerably different from probabilities of matching, which implies that existing mechanisms are far from optimal. We discuss how optimal points mechanisms based on our estimates could be used to improve hospital incentives.

Third, we study the importance of the two market failures. The loss in hospital welfare due to the inefficient mechanism depends on the wedge between current and optimal rewards, and on the elasticity of supply from hospitals. We have estimated the wedges and the marginal products, but our data do not have enough information to credibly estimate supply elasticities. Therefore, we calculate this deadweight loss under a broad range of assumptions on elasticities. Except under extreme assumptions, the deadweight loss is significant but lower than the inefficiency due to market fragmentation. Hence, both the current mechanism and agency problems cause significant inefficiency in the market. These results motivate a combined approach that improves the mechanism design and implements policies that encourage hospital participation in the national platforms.

2 Background and Data

2.1 Basics of Kidney Exchange

End-Stage Renal Disease (ESRD) afflicts more than half a million Americans. The disease is almost universally covered by Medicare, even for patients under the age of 65. The Medicare ESRD program accounts for 7% of its budget, mostly spent on patients undergoing dialysis (USRDS, United States Renal Data System, 2016). The preferred treatment for ESRD patients is transplantation, which increases the quality and length of life by several years and is cheaper than dialysis. Transplantation saves approximately \$270,000 per Medicare beneficiary and even more for privately insured patients (Wolfe et al., 1999; Irwin et al., 2012; Held et al., 2016). Moreover, the health risks to living donors are small. Taken together, these facts indicate that living donor kidney transplants have large economic value. Held et al. (2016) estimated the economic value of a kidney transplant at $1.1 million with a detailed cost-benefit analysis.[3]

There is a severe shortage of organs for transplantation. Each year, approximately 13,000 patients are transplanted using organs from deceased donors and another 5,500 from living donors. Demand far outstrips this supply with approximately 35,000 patients added to the deceased donor kidney waitlist in each of the recent few years. The shortage has resulted in the kidney waitlist growing to almost 100,000 patients and about 8,000 patients on the list dying or being categorized as too sick to transplant while waiting on the list.[4] Monetary compensation cannot be used to address this shortage because of ethical and legal reasons, and compensation is forbidden in almost every country (Becker and Elias, 2007), including the US.

Kidney exchange is an innovative way to ameliorate this shortage (Roth et al., 2004; Sönmez and Ünver, 2013a). It serves patients who have a willing live donor with whom they are not biologically compatible. Such patients can swap donors with others in the same situation, enabling transplants for many patients. These swaps are organized by **kidney exchange platforms** that match patients and donors registered with them. The platforms receive three types of **submissions**. The most common type is a **pair**, consisting of a patient and a living but incompatible donor. The second type is an **altruistic donor**, who is willing to donate a kidney to a stranger without requiring a transplant for an associated patient. Finally, there are some **unpaired patients**, who do not have a willing live donor.

Platforms organize transplants in two ways. The first, called a **cycle**, involves a set of pairs. The kidney from one pair's donor is transplanted into the patient in the next pair until the cycle is closed. All transplants are carried out simultaneously to reduce the risk that a pair donates a kidney without also receiving one. Cycles are usually limited to at most three pairs due to logistical constraints. The second type, called a **chain**, is initiated when an altruistic donor donates to a patient in an incompatible pair. The donor from this pair can then continue the chain by donating to the next pair and so on until the chain terminates with an unpaired patient. Chains can be very long in principle because transplants do not have to be performed simultaneously. However, our data from the National Kidney Registry (NKR), the largest kidney exchange platform, show that most chains involve four to five transplants. Initially, cycles were the most common type of transaction, but chains became more important over time and now account for about 90% of the transplants.

There are two types of biological compatibility constraints on kidney transplants: **blood-type** and **tissue-type** compatibility (Danovitch, 2009). A donor is blood-type incompatible with a patient if the donor has a blood antigen that the patient lacks. There are two blood antigens, known as A and B. Blood type is A or B if the blood has only the A or the B antigen, respectively, AB if it has both, and O if it has neither. A donor is tissue-type

[3]Most of the value comes from gains in quantity and quality of life. The cost savings on dialysis are also significant. In 2014, Medicare paid $87,638 per year per dialysis patient but only $32,586 in post-transplant costs per year per patient (USRDS, United States Renal Data System, 2016, Chapters 7 and 11).

[4]Statistics taken from https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/ (accessed December 21, 2017).

incompatible with a patient if the donor has certain human leukocyte antigens to which the patient has an immune response.[5] The most common measure of **sensitization**, that is, how likely a patient is to reject a transplant due to tissue-type incompatibility, is the **Panel Reactive Antibody (PRA)** score. A patient's PRA is between 0 and 100 and denotes the percentage of a representative population of donors with whom a patient is tissue-type incompatible. Because this measure depends on the choice of representative population, the NKR's algorithm uses an alternative measurement tailored to its pool called **match power**. It measures, for a given recipient (donor), the fraction of donors (recipients) on the platform that are blood-type and tissue-type compatible.

2.2 Key Institutional Features and the Economics of Kidney Exchange

There are three institutional features that are crucial for the economics of kidney exchange. First, kidney exchange takes place both in large, national platforms and within individual hospitals. There are three major national platforms currently operating in the United States: the National Kidney Registry (NKR), which is the largest; the Alliance for Paired Kidney Donation (APD); and the United Network for Organ Sharing (UNOS) KPD Pilot Program. These large platforms match patients using optimization software that maximizes a weighted number of transplants. They differ in terms of exact algorithms and operational details.[6] Besides these major platforms, there are small regional platforms and individual hospitals that also organize kidney exchanges. Most hospitals that participate in large national platforms also match patients outside those platforms.

Hospitals are not forced to participate in platforms.[7] Platforms effectively reward hospitals with transplants in order to receive submissions, as hospitals perform the transplants on patients they submit to a platform. Rewards can also be explicit. For example, most platforms reward hospitals that submit altruistic donors by matching one of their unpaired patients.

Second, there is substantial variation in the social value of different submissions due to biological compatibility. One reason for this variation is blood-type compatibility. To simplify

[5] The immune system recognizes foreign cells based on certain cell-surface proteins, known as antigens. The organism has antibodies that bind to these antigens, tagging foreign cells which are then attacked. Hence, if we put a cell with an antigen in the body of a person who has antibodies for that antigen, the immune system will attack it. The primary antigens that lead to kidney transplant rejection are located at the A, B, and DR loci of the human leukocyte antigen (HLA). Each donor has up to 2 possible HLA antigens at each of these loci, out of a list of hundreds. Similarly, a recipient has a list of antibodies to some, possibly large, subset of the HLA antigens. If the recipient has an antibody to one of the donor kidney's antigens, the recipient's immune system will attack the kidney, leading to rejection. A recipient is tissue-type compatible with a donor's kidney if she has no antibodies corresponding the antigens of the donor's kidney (Danovitch, 2009). Note that a transplant between certain incompatible patients and donors has become possible due to development of desensitization technologies (Orandi et al., 2014).

[6] See Abraham et al. (2007); Ashlagi et al. (2016); Anderson et al. (2014); Dickerson et al. (2012).

[7] Roth et al. (2005) and Ashlagi and Roth (2014) argue that hospitals are the key decision-makers and have incentives to perform potentially inefficient within hospital transplants. Ashlagi and Roth (2014) analyze the related mechanism design problem using a static model with stylized hospital behavior.

exposition, assume that there are only two blood types, O and A. These two types together are a significant majority of patients and donors in the U.S. Denote a pair with patient blood type X and donor blood type Y as X-Y, and let $q_{\text{X-Y}}$ be the number of such pairs in a pool. Assume that $q_{\text{A-O}} < q_{\text{O-A}}$, which is the empirically relevant case.[8] For this simplified case, Roth et al. (2007) showed that the number of transplants that can be performed, $f(\boldsymbol{q})$, is approximately

$$f(\boldsymbol{q}) = 2 \cdot q_{\text{A-O}} + 1 \cdot (q_{\text{A-A}} + q_{\text{O-O}}) + 0 \cdot q_{\text{O-A}}. \tag{1}$$

This result follows because A-A and O-O pairs can be matched with pairs of the same type. Roth et al. (2007) call these pairs **self-demanded**. Self-demanded pairs have a marginal product of 1, in the sense that they generate 1 additional transplant when they join the pool. However, an O-A pair can only be transplanted using a cycle with one of the valuable A-O pairs. Thus, there will be many leftover O-A pairs, and they can only be transplanted if more A-O pairs join the pool. A-O pairs are called **over-demanded** and have a marginal product of 2. O-A pairs are called **under-demanded** and have a marginal product of 0. An under-demanded pair competes with another under-demanded pair and adds no value to the pool. Roth et al. (2007) showed that this qualitative pattern holds even in a model with all possible blood types.

Current platform rules largely ignore this variation in the social value of submissions, inducing hospitals to perform socially inefficient matches. Consider a hospital with two over-demanded pairs. The hospital could perform a pairwise exchange to conduct two transplants. However, if the hospital submits the pairs to the platform, then in expectation, the hospital receives a number of transplants equal to twice the probability that a pair is matched. According to our data, this probability is 0.8, so the hospital expects only 1.6 transplants from submitting these two pairs to the platform. This expectation pushes the hospital to match its patients outside the platform. However, each pair the hospital submits to the platform generates its marginal product, which the Roth et al. (2007) model puts at 2. This suggests that the platform could generate 4 transplants if the hospital would submit both its pairs. Using a more realistic empirical model, we estimate only three additional transplants (Section 5). Either way, matching these two pairs within the hospital is socially inefficient despite the hospital's desire act in the best interest of their patients.

An important corollary of Roth et al. (2007)'s results is that transplants are a natural numeraire in a kidney exchange platform. Because hospitals have a large number of under-demanded pairs, it is easy for a platform to transfer transplants from one hospital to another without compromising efficiency, simply by choosing which under-demanded pairs to match.

Third, hospitals do not necessarily maximize a utilitarian measure of the welfare of the patients and third-party payers who they represent. We refer to such behavior as a broadly defined agency problem, since hospitals incur most of the costs of kidney exchange. The social value from one transplant is more than $1,000,000, mostly accruing to gains in quality-adjusted life years to patients and savings in health-care costs. But hospital revenues are between $100,000 to $160,000 per transplant.[9] Variable profits are likely much smaller.

[8] This fact is confirmed for patients and donors registered in the NKR. See Table 2 below.

[9] See Held et al. (2016); USRDS, United States Renal Data System (2013). The revenues include payments

Thus, even socially insignificant costs of performing kidney exchange through a platform can be important for hospitals. Conversations with hospital staff indicate that participation in kidney exchange platforms involves logistical and administrative hassle in addition to direct costs arising from biological testing and platform fees.[10] Previous surveys and interviews have found that these logistical and financial costs are commonly cited barriers to participation (Ellison, 2014; American Society of Transplant Surgeons, 2016). Besides costs, hospitals may have behavioral reasons for not perfectly maximizing patient welfare. For example, there is considerable heterogeneity regarding hospital sophistication: some hospitals use optimization software to match patients while others manually search for matches.

2.3 Data

We assembled two datasets for this paper. The first, the transplant dataset, records all kidney exchange transplants in the United States. We use this dataset to document fragmentation, inefficiency and participation in the market for kidney exchange. The second, the NKR dataset, records all patients and donors that registered with the largest kidney exchange platform, the NKR. We use this dataset to estimate a transplant production function.

The transplant dataset consists of anonymized records of every kidney transplant conducted in the US between January 1, 2008 and December 4, 2014. We obtained this dataset from the Organ Procurement and Transplantation Network (OPTN), a contractor for the U.S. Department of Health and Human Services.[11] The OPTN dataset includes each transplant's date and location; whether it is part of a kidney exchange; the age, sex, weight, height, body mass index (BMI), blood-type, and HLA antigens of the donor and recipient; and the unacceptable antigens and number of days on dialysis of the recipient. See Appendix B for details.

Although a comprehensive source for data on transplants conducted, the only field in the OPTN dataset that specifically pertains to kidney exchange is an indicator for which transplants were part of such an exchange. Therefore, the OPTN dataset does not identify which, if any, multi-hospital kidney exchange platform organized a given transplant.

To address this limitation, we separately obtained anonymized records of all transplants organized by each of the three largest platforms for multi-hospital kidney exchange platforms in the US: NKR, APD, and UNOS. These platform data include many of the fields as the

for surgery teams, drugs, equipment, and capital.

[10]Platforms require extensive biological testing, which is particularly complicated because donors and patients are in different hospitals. Platforms also charge fees, which are paid by hospitals. NKR charges annual fees of about $10,000 plus about $4,000 per transplant. See National Kidney Registry (2016) for NKR's fees, and Rees et al. (2012) and Wall et al. (2017) for a broader discussion kidney exchange costs borne by hospitals.

[11]This study uses data from the Organ Procurement and Transplantation Network (OPTN). The OPTN data system includes data on all donor, wait-listed candidates, and transplant recipients in the US, submitted by members of the Organ Procurement and Transplantation Network (OPTN). The Health Resources and Services Administration (HRSA), US Department of Health and Human Services provides oversight to the activities of the OPTN contractor.

OPTN dataset. By merging the data from these platforms with the OPTN data, we identified which transplants were organized through NKR, APD, UNOS, or through other avenues. This merge is not straightforward because all of our datasets are anonymized. Fortunately, the rich biological data allows us to match transplants across datasets on the blood type, sex, and HLA antigens of the recipient and donor, as well as the date and location of the transplant. More details on the merge procedure are provided in Appendix B. We were able to match approximately 94% of transplants at these platforms to the corresponding OPTN data with a high degree of certainty.[12]

The transplant dataset contains information on transplants that were performed, but not on the pool of patients and donors that were available for kidney exchange. This information is needed to estimate a platform's transplant production function. Therefore, we assembled the NKR dataset. It records all patients and donors that registered with the NKR between April 2, 2012 to December 4, 2014. These data are sourced from the administrative records the NKR uses to organize transplants. It includes the registration date, blood type, age, sex, HLA antigens for both patients and donors. It also records whether the patient or donor left NKR's system, and the date and reason for departure (transplantation or otherwise). In addition, it includes information on which donor is paired with which patient (if any), unacceptable antigens, and all the restrictions a patient places on which organs are acceptable. These data allow us to determine the set of transplants the NKR considers acceptable and medically feasible. We also have detailed data on how the transplants were organized, including the donors and patients involved, and the chain or cycle configuration. Appendix B provides details on how we assembled the NKR dataset.

3 Descriptive Evidence

We now document three key facts: the kidney exchange market is highly fragmented, this fragmentation leads to inefficiency, and there is evidence of agency problems between hospitals and patients.

3.1 Fragmentation

We first document that the market is highly fragmented. Most kidney exchange transactions are matched internally by individual hospitals, as opposed to by large, national kidney exchange platforms. A kidney exchange transplant is defined as **within hospital** if the donor's operation took place in the same hospital as the patient's, and **across hospitals** if the donor's and patient's operations took place in different hospitals.[13] We also classify transplants based on which platform coordinated the exchange: NKR, APD, or UNOS. Transplants that were

[12] 90% of the matches were within 1 day on the transplant date, within 5 years on donor and recipient age, and agreed on the hospital where the transplant was conducted as well as the blood type, sex, and all six major HLA alleles (2 alleles each at the HLA-A, B and DR loci) of both the donor and recipient.

[13] The common practice is to transport the organ after recovery instead of transporting the donor and recovering the organ elsewhere. A primary motivation for this practice is to safeguard the donor's interests and

not organized by one of these platforms are classified as being performed by other platforms, including single-hospital kidney exchange programs and by small regional platforms.

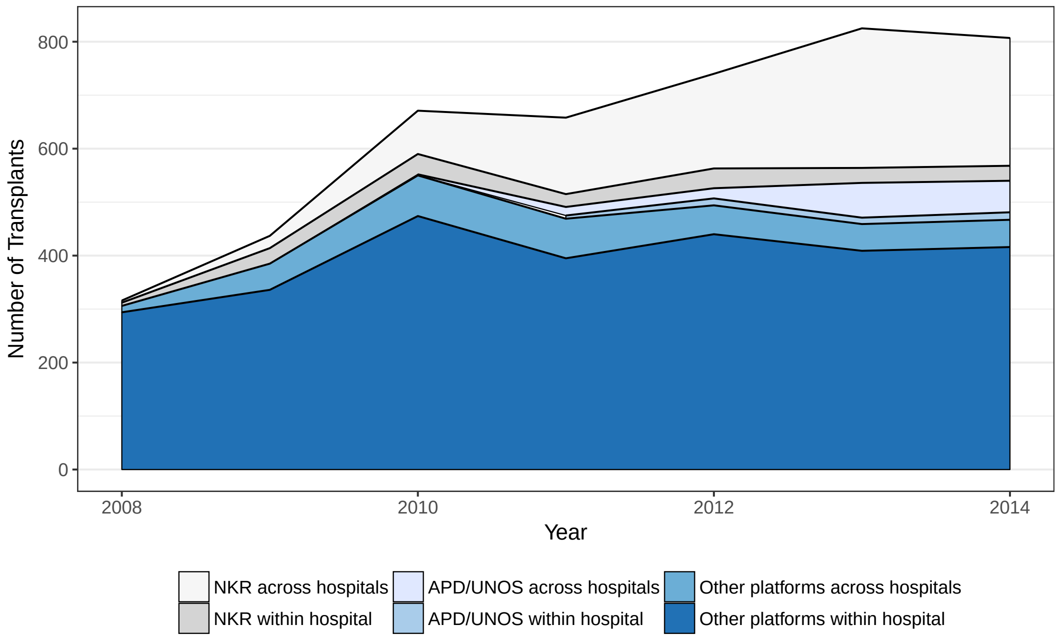


Figure 1: Market fragmentation and trends in kidney exchange

Notes: The figure displays the number of kidney exchange transplants in different categories. The category "Other" represents transplants that were not facilitated by NKR, APD, or UNOS. Single-hospital platforms fall under this category. Within hospital and across hospital classify a transplant into whether the donor's hospital was the same as the patient's hospital.

Figure 1 shows that the market is highly fragmented. The three largest multi-hospital platforms together only account for a minority share of the kidney exchange market. 62% of kidney exchange transplants are within hospital transplants that are not facilitated by the NKR, APD or UNOS. Moreover, over 100 hospitals performed kidney exchanges outside these three platforms during this period.

Unlike the dominance of within hospital exchanges in the overall market, a large majority of the transplants facilitated by multi-hospital platforms are across hospitals. This contrast between the overall market and the platforms is striking because the platforms do not prioritize across hospital exchanges as a rule. Rather, the predominance of across hospital exchanges in the national platforms is a by-product of maximizing the total number of transplants. This suggests that coordinating across hospitals has potential gains.

Figure 1 also shows that the total number of kidney exchange transplants grew from about 400 transplants in 2008 to about 800 in 2014.[14] However, overall market growth seems to

because, by the time of the transplant, the donor has built a relationship with his or her hospital and surgeon. The surgery performed on the donor requires extensive pre-planning and follow-up care. Conversations with surgeons suggest that these factors severely limit willingness to transport the donor and conduct surgery in another hospital.

[14] Our data for the NKR extend until December 4, 2014. This censoring may account for the slight drop

have slowed in recent years. The total number in 2017 remains at around 800,[15] well below some estimates of the potential size of the kidney exchange market (Bingaman et al., 2012; Massie et al., 2013).

The growth in kidney exchange between 2010 and 2014 is concurrent with the NKR becoming the dominant kidney exchange platform. It accounted for 33.1% of all kidney exchange transplants in 2014, and facilitated more than 5 times as many transplants as the APD and UNOS combined.[16] The importance of the NKR during our sample period motivates our focus on the platform in the subsequent sections.

3.2 Evidence of Inefficiency

Market fragmentation creates inefficiency if there are increasing returns to scale in matching patients and donors, and hospitals are operating below efficient scale. We now present direct evidence of hospitals conducting exchanges that are inefficient from a social perspective.

One easily detectable inefficiency is a transplant between an O blood type donor and a non-O blood type patient. As explained in Roth et al. (2007) and in Section 2, O donors are scarce while O patients are abundant. If all transplants are of equal social value, optimal matches in a large market should only transplant organs from O donors to O patients because O patients cannot accept other blood types.[17] The exception to this rule is for a highly sensitized patient, that is, one with a very high PRA. The platform may want to use an O donor to transplant such a patient if it is the only way to get the patient transplanted.

Figure 2 displays the fraction of O donors that are used to transplant non-O patients, categorized into NKR transplants, APD/UNOS transplants, across hospital transplants at other platforms, and within hospital transplants at other platforms. Among NKR transplants, only 6.5% of O donors are used for non-O patients. In contrast, among within hospital transplants outside the three platforms, this percentage is 22.8%. This difference is statistically significant ($p < 0.01$) and constitutes strong evidence that hospitals often perform inefficient matches outside the platform. Transplants at APD, UNOS and across-hospital transplants at other platforms are in between these two categories but are much closer to the NKR.

An alternative explanation for inefficient matching is that within hospital transplants use O donors to help highly sensitized patients who would otherwise remain untransplanted. However, Figure 2 shows that almost none of the potentially inefficient transplants in the Other (within hospital) category involve highly sensitized patients. In contrast, about half of the potentially inefficient NKR transplants involve highly sensitized patients.

---

in transplants in the last year of this figure.

[15]Source: https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/ (accessed December 21, 2017).

[16]The APD has grown in recent years and has significantly closed the gap.

[17]Strictly speaking, efficiency as discussed here means maximizing the total number of transplants. However, transplanting an O donor to a non-O patient is also likely to be Pareto inefficient. To see this, consider a pairwise exchange between two overdemanded A-O pairs. This exchange results in two transplants. It would be more efficient to transplant each of the A-O pairs to an underdemanded O-A pair, which otherwise would be left unmatched.

Table 1: Summary Statistics for Kidney Exchange Transplants

	NKR	APD / UNOS	Other platforms	
			Across Hospital	Within Hospital
N	1118	198	341	2719
Patient Blood Type				
A	34.7%	36.4%	37.2%	37.1%
B	19.0%	21.2%	17.6%	17.0%
AB	5.7%	3.5%	7.0%	5.7%
O	40.6%	38.9%	38.1%	40.2%
Donor Blood Type				
A	36.8%	35.4%	37.5%	33.4%
B	18.2%	20.2%	14.7%	13.8%
AB	3.9%	1.5%	6.7%	2.9%
O	41.1%	42.9%	41.1%	49.9%
Panel Reactive Antibody (PRA) (Sensitization)				
Mean	35.0	43.0	30.4	17.6
Standard deviation	39.7	40.8	37.5	30.8
Percent >90	16.4%	20.6%	12.0%	5.1%
Transplant Ourcomes and Quality Measures				
Donor Age				
Mean	44.1	44.6	44.1	43.2
Standard deviation	11.8	11.1	11.3	11.8
Donor Body Mass Index (BMI)				
Mean	26.5	27.0	26.6	26.5
Standard deviation	4.0	4.0	4.1	4.2
Donor Height (cm)				
Mean	169.4	168.0	169.6	169.3
Standard deviation	9.8	9.6	10.3	9.8
Donor Weight (kg)				
Mean	76.3	76.3	76.9	76.3
Standard deviation	15.1	13.9	15.4	15.1
Tissue Type Mismatch (0-6)				
Mean	4.2	4.2	4.2	4.4
Standard deviation	1.3	1.4	1.2	1.2
Mean Days on Dialysis				
Mean	1026.6	1048.4	1063.1	969.1
Standard deviation	1088.1	848.1	1269.5	990.9

Notes: Sample of all Kidney Exchange Transplants between January 1, 2008 and December 4, 2014.

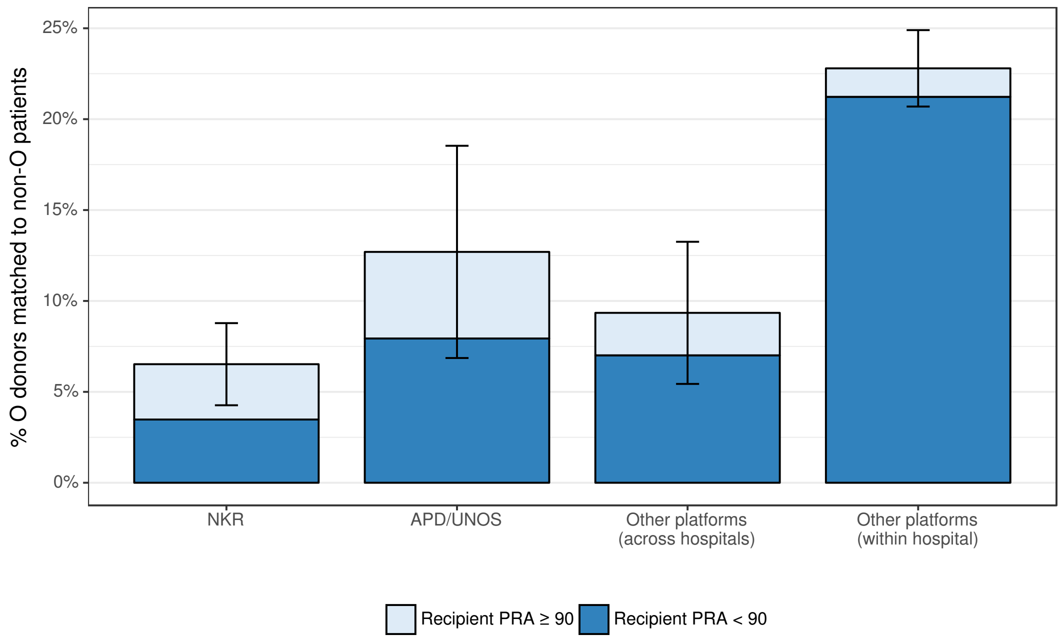


Figure 2: Evidence of hospitals performing inefficient matches

Notes: The bars display the percentage of transplanted O donors whose kidneys were transplanted into non-O patients for different categories of transplants. Other indicates a transplant not organized by NKR, APD, or UNOS. This category includes transplants organized by single-hospital platforms. Within hospital and across hospital classify a transplant into whether the donor hospital was the same as the patient hospital. The colors decompose this total into highly sensitized patients (PRA >90) and non-highly sensitized patients. The error bars depict 95% confidence intervals for the totals.

This exercise is based on the assumption that the value of a kidney exchange transplant does not depend on how it was organized. Section 2.1 argues that that logistical costs of conducting transplants through a platform are negligible relative to the value of transplants lost by using organs from O donors to transplant non-O patients. However, there may be dimensions on which within hospital transplants are superior to transplants organized by national platforms. For example, a transplant through a national platform could involve a longer wait on dialysis or a lower-quality donor. However, Table 1 shows that patients who receive a transplant through a platform typically spend only two more months on dialysis than patients who receive a within hospital transplant outside these three platforms. Given that the average patient wait is about 32 months, this difference represents an 8% longer waiting time. The longer waiting time at the platforms should be expected because, as we discuss below, patients transplanted through the platform are, on average, harder to match. Further, there do not seem to be differences in how desirable donors might be to patients. Donor quality indicators such as age, weight, height, and BMI are similar across platforms. One reason why patients considering a multi-hospital platform need not worry about donor quality is that the platforms typically allow patients and doctors to specify donor acceptability criteria. They also allow patients to refuse proposed transplants if the donor is unsuitable.

If each of these inefficient transplants from O donors to non-O patients comes at the cost of one other transplant, as in the Roth et al. (2007) model, then achieving the level of efficiency obtained by the NKR would have resulted in about 250 additional transplants between 2008 and 2014.[18] The advantage of considering only the clearly inefficient transplants is that the results provide transparent evidence of inefficiency. The total inefficiency, of course, can be much larger.

3.3 Hospital Participation Behavior and Evidence of Agency Problems

The results on market fragmentation and inefficiency lead us to ask why hospitals do not participate more in national platforms. We start by documenting key facts about hospital behavior and argue that hospitals do not purely maximize the number of transplanted patients. Instead, hospitals seem to maximize complex and heterogeneous objectives, including, but not limited to, profits and patient welfare.

3.3.1 Descriptive Evidence

We focus on participation behavior at the NKR because it is the primary multi-hospital kidney exchange platform during our sample period (Table 1). Figure 3 depicts the extensive margin of participation among hospitals conducting kidney exchange transplants. A hospital is considered an NKR participant if it has ever submitted a patient or donor to the NKR. The figure is a binned scatterplot of the fraction of hospitals that participate in the NKR versus hospital size in terms of the total number of kidney transplants performed (living and deceased).[19] Figure 4 depicts the intensive margin of participation. The vertical axis in this scatterplot is the fraction of kidney exchange transplants that a hospital performs through the NKR. The results are qualitatively similar if we consider participation at any of the three largest kidney exchange platforms.

The figures reveal four key facts about participation. First, both the extensive and intensive margins are important drivers of market fragmentation. Only 46.3% of hospitals participate in the NKR. Within those participating hospitals, only 52.9% of transplants are conducted through the NKR. Second, larger hospitals are considerably more likely to participate in the NKR. The probability of participating at all is about 80% for a hospital that performs approximately 250 transplants per year but only about 35% for a hospital that performs about 50 transplants per year (Figure 3). Third, conditional on participating, large hospitals conduct more of their matches outside the platform (Figure 4). Although size positively correlates

[18] Table 1 shows that there is a larger gap between the fraction of donors and patients that are blood type O for within hospital platforms as compared to the NKR. The difference in this gap multiplied by the number of transplants arranged within hospital is a measure of transplants lost due to inefficient use of O donors in within hospital transplants.

[19] This broad measure of size limits the endogenous effect of participation in the NKR on hospital size because deceased donor and direct living donor transplants form the bulk of kidney transplants conducted by hospitals. Moreover, during our sample period, the total number of kidney transplants has remained stable relative to the growth in kidney exchange.

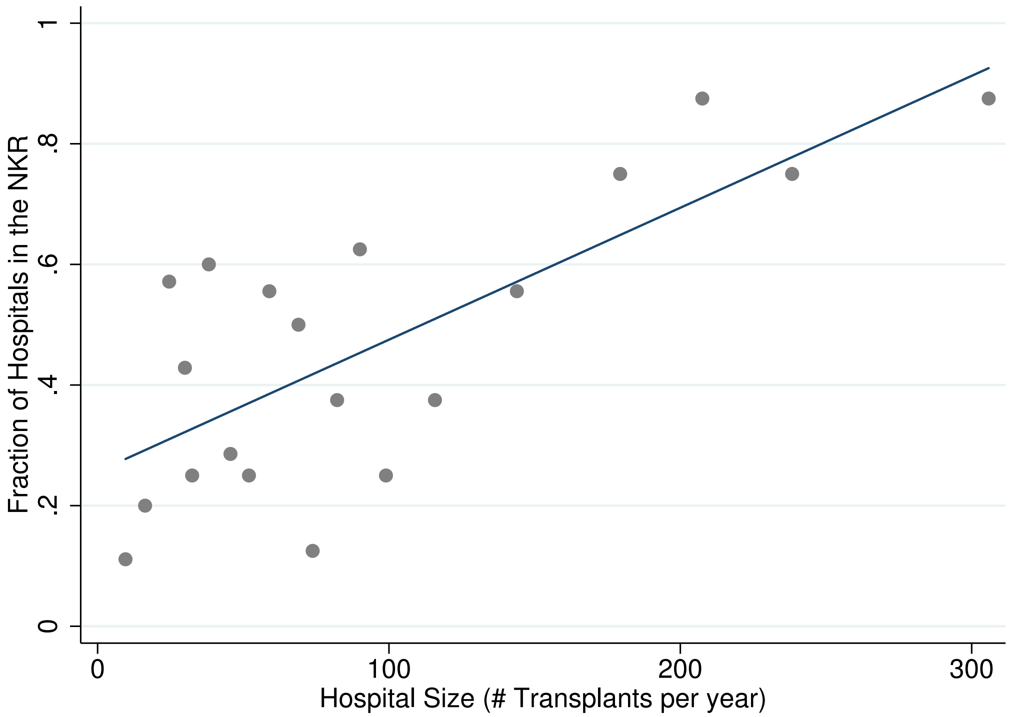


Figure 3: Heterogeneity in participation in the NKR

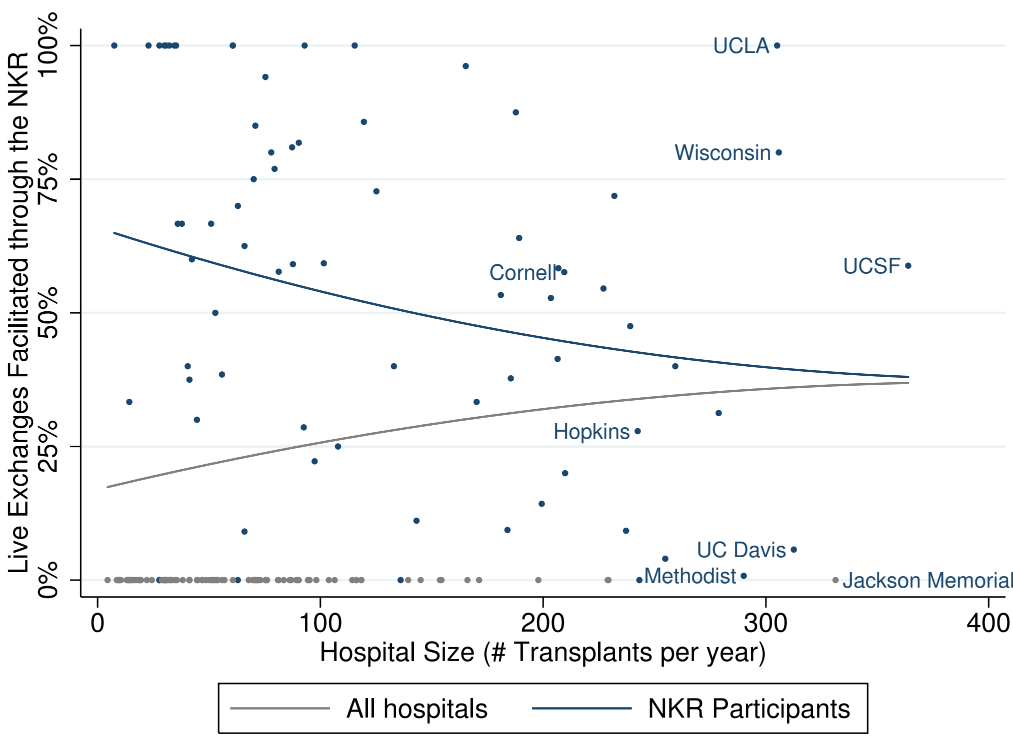


Figure 4: Reliance on the NKR for live-donor exchanges

with the fraction of kidney exchange transplants performed in the NKR, the relationship is negative if we focus exclusively on hospitals that participate at all (Figure 4). Fourth, there is a high degree of heterogeneity in intensive margin participation. Even among hospitals with similar size, participation varies considerably (Figure 4). For example, among the five transplant hospitals that perform more than 300 transplants per year, one does not participate at all (Jackson Memorial), one participates close to zero percent (UC Davis Medical Center), two participate in the 50-60% range (UCSF Medical Center and the University of Wisconsin Hospital), and one participates more than 80% (UCLA Medical Center).

In addition to revealing the decision to participate, the data provide information on the characteristics of patients submitted to the NKR and the characteristics of patients transplanted by each hospital categorized by how the transplant was facilitated. Tables 1 and 2 reveal three main facts.

First, the NKR receives submissions that are very hard to match compared to the general population (Table 2). The blood types of both altruistic and paired donors skew away from O donors and toward A donors relative to the US population. The deceased donor population has about 45% O donors and 40% A donors. In contrast, patients in pairs are disproportionately likely to have blood type O (58.6%), and their related donors are unlikely to have blood type O (31.9%). Only a small fraction of pairs (13.8%) are overdemanded. Interestingly, unpaired patients are much more likely to have an easy-to-match blood type, with the majority having blood type A. The average PRA for patients registered with the NKR is high. At a mean PRA of 48.8%, the average patient in the NKR is tissue-type incompatible with approximately half the reference donor population.

Second, the NKR transplants patients who are considerably harder to match than patients transplanted by single hospitals (Table 1). Approximately 40% of the patients and 41% of donors transplanted through the NKR are blood type O. The PRA of the patients transplanted through the NKR is approximately 35%, and about one in six patients have a PRA above 90%. These statistics are similar for across hospital kidney exchanges not facilitated by the NKR and transplants facilitated by APD or UNOS. In contrast, among within hospital kidney exchanges not conducted by a large platform, almost 50% of the donors are blood type O, but only 40% of the patients are blood type O. The average PRA of patients transplanted through within hospital exchanges is only 18%. This is almost half the mean PRA for patients transplanted through one of the three national platforms.

Third, transplants on all platforms look similar in donor quality measures that do not affect compatibility, such as weight, body mass index, and age (Table 1). This supports our equal treatment of all transplants for welfare calculations irrespective of whether they are facilitated though a national platform.

3.3.2 Implications for Hospital Behavior

The facts above have implications for different hypotheses about hospital behavior. In the discussion that follows, we approximate total patient welfare with the total number of transplants. As we argued in Section 2, kidney exchange costs are small relative to the benefits

Table 2: Summary Statistics for NKR Submissions

	Altruistic Donors	Pairs	Unpaired Patients
N	164	1265	501
Patient Blood Type			
A	–	23.8%	51.1%
B	–	15.0%	16.0%
AB	–	2.6%	19.0%
O	–	58.6%	14.0%
Donor Blood Type			
A	44.5%	44.4%	–
B	14.0%	18.5%	–
AB	3.7%	5.2%	–
O	37.8%	31.9%	–
Match Power			
Recipient/Pair			
Mean	–	0.218	0.431
Standard deviation	–	0.210	0.392
Donor			
Mean	0.279	0.258	–
Standard deviation	0.162	0.159	–
Panel Reactive Antibody (PRA) (Sensitization)			
Mean	–	48.8	44.4
Standard deviation	–	41.1	45.1
Pair Type			
Overdemanded	–	13.8%	–
Underdemanded	–	41.9%	–

Notes: A pair is overdemanded if the patient is blood type compatible with the related donor, but not of the same blood type. Underdemanded pairs are either blood type O patients without blood type O donors or are blood type AB donors without blood type AB patients. Sample of all patients and donors registered in the NKR between April 2, 2012 and December 4, 2014.

of transplantation.

The first hypothesis is that that hospitals maximize the total welfare of all patients in the system, regardless of which hospital a patient belongs to. This hypothesis is strongly rejected by several features of the data. Most clearly, this hypothesis is inconsistent with the evidence of socially inefficient matches (Figure 2).

A second hypothesis is that hospitals only maximize the welfare of their own patients. This hypothesis was investigated theoretically by Ashlagi and Roth (2014) who argue that hospitals will try to match as many of their patients internally as possible and only submit the remaining patients to a multi-hospital kidney exchange platform. This hypothesis fits some qualitative patterns in the data, but not others. For example, it explains why larger hospitals in the NKR perform fewer transplants through the platform. These hospitals have more opportunities to match patients outside the platform (Figure 4). However, it does not explain why many hospitals do not participate in a national platform at all, even though all hospitals likely have patients who cannot be matched.[20] Moreover, many small hospitals do not participate in the NKR, even though these hospitals, due to their size, are precisely the ones least likely to find matches outside the platform. The patterns suggest that hospitals respond to fixed costs of participating in kidney exchange platforms, even though these costs are small relative to benefits to patients and cost savings from dialysis to health insurers.

A third hypothesis is that hospitals are profit maximizers. This hypothesis is consistent with the fact that small hospitals are less likely to participate in the NKR (Figure 3) because the fixed costs of participation may not compensate for the gains in profits from additional transplants. However, this theory alone cannot fully explain the large variation in the degree of participation, especially among large hospitals. For example, Cornell Medical Center is a large hospital with a high rate of participation in the NKR. Interviews with transplant coordinators at Cornell, reported in Ellison (2014), suggest that a primary reason for participating is the view that contributing to a national kidney exchange platform is important.

Taken together, the evidence on hospital participation suggests that hospitals maximize complex and heterogeneous objectives. This finding is consistent with the anecdotal evidence on kidney exchange reviewed in Section 2, as well as the standard view in healthcare economics (Arrow, 1963), and typical findings about the behavior of healthcare providers (Arrow, 1963; Kolstad, 2013; Clemens and Gottlieb, 2014).

The facts about selection into which patients and donors are submitted to the NKR also indicate that these two theories, maximizing profits and maximizing their own patients' welfare, can explain many hospitals' behavior. These theories' shared implication is that pairs submitted to national platforms are negatively selected, in the sense of being hard to match. In both cases, a hospital only submits a pair to a platform if an internal match is not possible. Unfortunately, we cannot directly test this prediction because we do not have data on the entire pool of patients available to individual hospitals. But, it is reassuring that the results on selection do not falsify the two theories that best fit the participation behavior.

[20] Recall that overdemanded pairs are typically scarce. We will see in Section 5 that even the NKR is able to match only approximately 50% of its donors.

To summarize, these findings have two important implications. First, there is clear evidence of agency problems, as we defined broadly in Section 2. The data disprove the hypothesis that hospitals purely and rationally maximize their own patients' welfare. Second, none of the simple models describes the behavior of all hospitals.

4 Theory

The evidence above shows that kidney exchange markets are fragmented and that this fragmentation leads to real efficiency loss. We now build a model of the kidney exchange market that is fundamentally similar to a traditional market in which the platform procures submissions (donors and patients) from hospitals and rewards these hospitals with transplants.[21] Thus, we can use standard economic theory to explain how inefficiency arises, quantify it, and develop responses.

4.1 Model

A kidney exchange platform procures submissions from hospitals and rewards hospitals with transplants. The platform's ability to produce transplants is described by a **production function** f. We consider **types of submissions** $i = 1, \ldots, I$. A vector of **quantities** $\boldsymbol{q} = (q_i)_{i=1,\ldots,I}$ in $\mathbb{R}^I_+$ specifies a quantity q_i of each submission type available to the platform, where $\mathbb{R}_+$ is the set of non-negative real numbers. Given a vector of quantities $\boldsymbol{q}$, the platform can produce $f(\boldsymbol{q})$ transplants. The model can be interpreted as either static or as a steady-state from a dynamic model. We will use the steady-state interpretation in the empirical analysis. All variables are measured in flows, such as transplants per year.

The production function $f(\boldsymbol{q})$ summarizes what matches are possible. Roth et al. (2007) calculated the production function using a simple model that we described in Section 2. Since that paper assumed that all submissions are pairs and that only blood type compatibility matters, its model has $I = 16$ types. Our analysis applies both to such theoretically tractable production functions as well as to more complex production functions. Section 5 uses an empirical production function that allows submissions to differ by whether they are patient-donor pairs, altruistic donors, or unpaired patients, and by a host of variables including blood types, antigens, and antibodies. Thus, the number of types I is potentially large.

We say that the production function f has **constant returns to scale at $\boldsymbol{q}$** if its elasticity with respect to scale at $\boldsymbol{q}$ is equal to one, that is, if $\frac{\alpha}{f(\alpha \boldsymbol{q})} \cdot \frac{\partial f(\alpha \boldsymbol{q})}{\partial \alpha}\Big|_{\alpha=1} = 1$. Mathematically, this property is equivalent to $\boldsymbol{\nabla f}(\boldsymbol{q}) \cdot \boldsymbol{q} = f(\boldsymbol{q})$. The Roth et al. (2007) model considers a large platform with a linear production function, implying constant returns to scale. Our empirical production function in Section 5 will measure the elasticity with respect to scale.

[21] Our setting is also closely related to the platforms literature, wherein platforms maximize a private or social goal by setting incentives for participants (Rochet and Tirole, 2003; Weyl, 2010).

The platform produces transplants using submissions provided by **hospitals** indexed by $h = 1, \dots, H$. Hospitals are rewarded for these submissions with transplants. We assume these rewards are linear in submissions and anonymous. That is, there exists a **vector of rewards** $\boldsymbol{p} = (p_i)_{i=1}^{I}$ in $\mathbb{R}^I$ where the i^{th} component denotes the (expected) number of transplants awarded to the hospital per submission of type i. The units of p_i are transplants per submission. A hospital that submits a flow $\boldsymbol{q}^h$ in $\mathbb{R}_+^I$ of submissions receives a flow $\boldsymbol{p} \cdot \boldsymbol{q}^h$ of transplants. Since all transplants that are performed must be allocated to some hospital, a platform must satisfy the constraint that $f\left(\sum_h \boldsymbol{q}^h\right) = \sum_h \boldsymbol{p} \cdot \boldsymbol{q}^h$.

This linear reward schedule is a good approximation of current platforms' rules because their matching algorithms maximize a weighted sum of the number of matches without considering the entire pool of patients and donors submitted by each hospital (Sönmez and Ünver, 2013b; Anderson et al., 2015). When a hospital submits an additional pair, the probability that the platform matches a different pair from the same hospital does not significantly change. Therefore, the current reward for submitting a type i pair is equal to the probability p_i that the pair is matched.

We assume that hospital preferences are equal to the number of transplants they receive from the platform minus the **private cost** of their submissions, $C^h\left(\boldsymbol{q}^h\right)$, measured in transplant units. For instance, if the hospital maximizes the number of its own patients that are transplanted, then $C^h(\boldsymbol{q}^h)$ is the number of within-hospital transplants that the hospital must forgo in order to submit $\boldsymbol{q}^h$.

Welfare is defined over an **allocation** $(\boldsymbol{q}^h)_{h=1,\dots,H}$ that specifies the quantity of pairs supplied by each hospital. We will use two welfare notions. Both welfare notions use transplants as a numeraire because platforms can effectively transfer transplants between hospitals by choosing which underdemanded submissions to match (see Section 2).

The first notion is **hospital welfare** $W^H(\boldsymbol{q}^1, \dots, \boldsymbol{q}^H)$, which is the total welfare measured from the point of view of hospitals. Hospital welfare equals the total number of transplants produced (which is the same number of transplants that hospitals receive) minus the private costs. That is,

$$W^H(\boldsymbol{q}^1, \dots, \boldsymbol{q}^H) = f\left(\sum_h \boldsymbol{q}^h\right) - \sum_{h=1}^{H} C^h(\boldsymbol{q}^h). \tag{2}$$

This is a compelling notion of welfare if the goal is to help key market participants (hospitals, in this case) achieve their objectives.

Hospital welfare is not compelling if there are agency problems, that is, if hospitals do not purely maximize patient and insurer welfare. As discussed in Sections 2.1 and 3, there is anecdotal and empirical evidence of agency problems. For this reason, we also consider a utilitarian welfare measure, which we term social welfare.

Define $SC^h(\boldsymbol{q}^h)$ as the **social cost** for hospital h to supply a vector $\boldsymbol{q}^h$ submissions. If there are agency problems, then social and private costs are different, and there is an agency **externality** from hospital h's submissions because

$$C^h(\boldsymbol{q}^h) \neq SC^h(\boldsymbol{q}^h).$$

For example, $C^h(\boldsymbol{q}^h)$ is larger than $SC^h(\boldsymbol{q}^h)$ if hospital h acts as though the financial and logistical costs of participating in kidney exchange platforms are significant relative to their private value of a transplant. The externality represents the benefits to stakeholders other than the hospital itself. In the particular case where there are no agency problems, we have $C^h(\boldsymbol{q}^h) = SC^h(\boldsymbol{q}^h)$. for all h. Define **social welfare** to be

$$SW(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^H) = f(\boldsymbol{q}) - \sum_{h=1}^{H} SC^h(\boldsymbol{q}^h).$$

Define first-best hospital welfare as the supremum of W^H and first-best social welfare as the supremum of SW.

Given these primitives, for a vector of rewards $\boldsymbol{p}$ the hospital supply of hospital h is given by

$$\boldsymbol{S}^h(\boldsymbol{p}) = \arg \max_{\boldsymbol{q}^h \in \mathbb{R}^I_+} \boldsymbol{p} \cdot \boldsymbol{q}^h - C^h(\boldsymbol{q}^h).$$

Define the **aggregate cost**, $C(\boldsymbol{q})$, to be the minimum sum of hospital private costs necessary to ensure that hospitals supply $\boldsymbol{q} \equiv \sum_h \boldsymbol{q}^h$ in aggregate. Let the aggregate supply correspondence be

$$\boldsymbol{S}(\boldsymbol{p}) = \arg \max_{\boldsymbol{q} \in \mathbb{R}^I_+} \boldsymbol{p} \cdot \boldsymbol{q} - C(\boldsymbol{q}).$$

We assume that the production function, social and private costs, and aggregate cost functions are defined over all non-negative real vectors and are smooth. The maximum of each hospital's objective is attained for some quantity for every vector of rewards. Quantities are column vectors, and vectors of rewards and gradients are row vectors. Further, assume that aggregate cost is strictly convex.

Appendix A shows that aggregating individual hospital supplies yields $\boldsymbol{S}(\boldsymbol{p})$. Denote the **aggregate inverse supply** with $\boldsymbol{P}_S(\boldsymbol{q}) = \{\boldsymbol{p} \in \mathbb{R}^I \mid \boldsymbol{q} \in \boldsymbol{S}(\boldsymbol{p})\}$. Further, Appendix A shows that, for strictly positive $\boldsymbol{q}$, the aggregate inverse supply is single-valued and $\boldsymbol{P}_S(\boldsymbol{q}) = \boldsymbol{\nabla C}(\boldsymbol{q})$. This result is similar to how firms supply at price equal to marginal cost in a competitive market.

4.2 Illustrative examples

4.2.1 Agency and the wedge between private and social costs

Our model of the kidney exchange market is framed in terms of transplants as a numeraire, and captures agency problems as a wedge between private and social costs. We now present a particular example of our general model to clarify definitions and these two features of the model. The specific assumptions in this section are not necessary for our results.

Let $K^h(\boldsymbol{q}^h)$ be the monetary costs borne by hospital h of sending $\boldsymbol{q}^h$ submissions to a kidney exchange platform. This cost can include platform fees, costs of rearranging the hospital's schedule around the platform, and funds for hiring additional transplant coordinators (see Section 2.1). Let $T^h(\boldsymbol{q}^h)$ be the flow of kidney exchange transplants that hospital h forgoes when submitting $\boldsymbol{q}^h$ to the platform because the hospital cannot match these patients and donors internally. The function T^h can depend on the patients and donors that are available to hospital h.

To combine the monetary costs and the transplant costs of submitting, we need a rate of exchange between the two. Let hospitals value each transplant at v dollars, which includes profits and the value that hospitals place on transplanting their patients. Gross revenues from a transplant are approximately \$150,000 (USRDS, United States Renal Data System, 2013; Held et al., 2016). For illustrative purposes, take v to be \$50,000, which represents a generous 50% mark-up on costs. In transplant units, hospital h's cost function is

$$C^h(\boldsymbol{q}^h) = T^h(\boldsymbol{q}^h) + \frac{K^h(\boldsymbol{q}^h)}{v}.$$

The private value of a transplant just discussed does not account for any benefits that fall to non-hospital stakeholders. Such benefits include the value a patient has for a transplant beyond the value the hospital places on it and the savings in healthcare costs to insurers. Hospitals contract with these agents, but may not take all of their benefits into account because they are not incentivized to do so. For this reason, the social value of a transplant may differ from the private value to a hospital, creating agency problems.

Let society value transplants at $V > v$ dollars. Following the cost-benefit analysis in Held et al. (2016), we take V as \$1.1 million.[22] This fits our model with social costs

$$SC^h(\boldsymbol{q}^h) = T^h(\boldsymbol{q}^h) + \frac{K^h(\boldsymbol{q}^h)}{V}.$$

Hence, the wedge between private and social costs equals

$$C^h(\boldsymbol{q}^h) - SC^h(\boldsymbol{q}^h) = \left(\frac{1}{v} - \frac{1}{V}\right) \cdot K(\boldsymbol{q}^h).$$

The difference is how much more hospitals care about the costs of participating in a kidney exchange platform than society does measured in transplant units.

To develop intuition for this wedge's magnitude, assume that the monetary cost is linear in the number of submissions, i.e. $K^h(\boldsymbol{q}^h) = k \sum_i q_i^h$. Then, the wedge is

$$C^h(\boldsymbol{q}^h) - SC^h(\boldsymbol{q}^h) = \left(\frac{k}{v} - \frac{k}{V}\right) \cdot \sum_i q_i^h \approx \frac{k}{v} \cdot \sum_i q_i^h,$$

where the approximation holds because the social value of a transplant V is much larger than the monetary cost k. The wedge is large because it depends on the platform participation costs borne by the hospitals as a fraction of a transplant's private value, not its social value.

[22] Some patients who receive a kidney exchange transplant would otherwise receive a kidney from a deceased donor. But, in each of those cases, a transplant through kidney enables another patient on the waitlist to receive a kidney. Therefore, the social benefit of each kidney exchange transplant should still be the same as the gain from a single transplant.

For example, if k is \$10,000 and v is \$50,000, then the wedge is $k/v = 0.20$ transplants per submission. Hospitals compare this wedge to the rewards vector $\boldsymbol{p}$, which is equal to the probability of matching various submissions in the current mechanism. In effect, the wedge creates an incentive for the hospital to not submit a patient or donor to a national platform. The calculation above suggests that, because of agency problems, rewards have to be 20 percentage points higher in order to induce a given submission. Therefore, it is likely that agency problems are an important part of the kidney exchange market.

4.2.2 Two sources of market failure

Figure 5 presents a graphical illustration to clarify the two sources of market failure: agency problems and inefficient platform incentives. The horizontal axis plots aggregate supply $\boldsymbol{q}$. The vertical axis plots marginal products, social costs, and social benefits. The current vector of rewards, which is equal to the probability of matching each pair, is denoted by $\boldsymbol{p}_0$. The current quantity supplied given these rewards is $\boldsymbol{q}_0$. The curve $\boldsymbol{\nabla SC}(\boldsymbol{q})$ is the marginal aggregate social cost if hospitals choose privately optimal quantities given rewards $\boldsymbol{P}_S(\boldsymbol{q})$.[23]

The figure shows that the current allocation is inefficient from both the hospital and social perspectives. The hospital-optimal quantity $\boldsymbol{q}^*$ equates $\boldsymbol{\nabla f}$ with marginal aggregate private costs, as required by the first-order conditions of the hospital welfare maximization problem. Thus, the first inefficiency is that the platform gives inefficient incentives, $\boldsymbol{p} \neq \boldsymbol{\nabla f}$. The second inefficiency is that there are agency problems because hospitals do not choose socially optimal quantities, or equivalently $C^h \neq SC^h$. The aggregate quantity $\boldsymbol{q}^{**}$ maximizes social welfare subject to hospitals optimizing given a rewards vector. It attains the first-best social and hospital welfare if we also solve agency problems so that $C^h = SC^h$ and the two welfare notions coincide. In the example above, agency problems can be solved by reimbursing hospitals for the costs of kidney exchange through the platform $K^h\left(q^h\right)$.

This intuitive explanation glossed over two subtleties, as will be made clear by the formal results. First, efficient platform incentives are only approximately equal to marginal products. A platform cannot set incentives equal to the marginal product because there are increasing returns to scale, and therefore, marginal products exceed average products. However, estimates in Section 5 will show that this adjustment is negligible for the NKR.

Second, it is not possible to reach the first-best social welfare by only improving the mechanism if there are agency problems. If agency problems are more severe in some hospitals than in others, then it is not possible to achieve first-best welfare with rewards that are the same for all hospitals. Moreover, even achieving the welfare level corresponding to $\boldsymbol{q}^{**}$ requires setting rewards equal to the marginal products plus the difference between marginal private and social costs. But, even with constant returns to scale, the number of transplants equals the sum of marginal products of submissions. Therefore, there are not enough transplants to

[23] Formally, $SC(\boldsymbol{q}) = \sum_{h=1}^{H} SC^h\left(\boldsymbol{S}^h\left(\boldsymbol{P}_S(\boldsymbol{q})\right)\right)$ is the reward-moderated social cost. The figure assumes that individual supply is uniquely defined and that SC is differentiable.

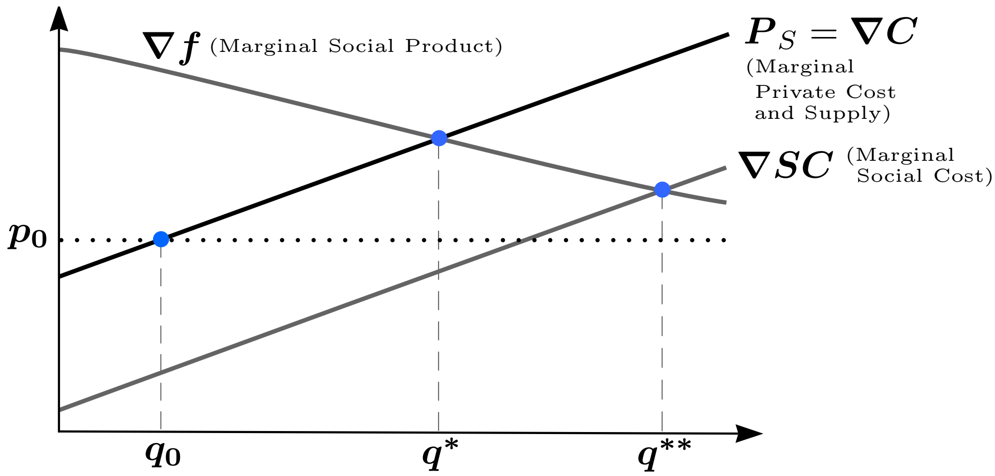


Figure 5: The Two Sources of Market Failure

Notes: The horizontal axis represents aggregate quantity of submissions into the kidney exchange platform. The curves represent the marginal product of submissions $\nabla f(q)$, the marginal private cost of submissions from hospital's perspective, $\nabla C(q)$ (which is equal to $P_S(q)$, the inverse aggregate supply), and the marginal social cost of submissions $\nabla SC(q)$. Both axes represent I-dimensional vectors. The figure depicts the current quantity q_0, with agency problems and a suboptimal mechanism, the quantity q^* from a hospital-optimal mechanism but with agency problems, and the first-best quantity q^{**} with an efficient mechanism, and no agency problems.

also reward hospitals to correct for agency problems (Proposition 2). These arguments suggest a two-pronged approach: design optimal mechanisms that help hospitals achieve their collective goals, and simultaneously implement policies to solve the agency problems.

4.3 Optimal Incentives

We now describe optimal reward vectors. The following theorem collects the main insights.

Theorem 1 (Optimal Rewards). *Consider a vector of rewards $\boldsymbol{p}$ and an allocation $(\boldsymbol{q}^h)_{h=1}^H$ with strictly positive aggregate quantity $\boldsymbol{q}$ that maximizes hospital welfare subject to all hospitals choosing $\boldsymbol{q}^h \in \boldsymbol{S}^h(\boldsymbol{p})$ and subject to the total rewards allocated being the same as the number of transplants produced, that is, $f(\boldsymbol{q}) = \boldsymbol{p} \cdot \boldsymbol{q}$. Then:*

1. *The platform rewards each type of submission with its marginal product minus an adjustment term,*
$$\boldsymbol{p} = \nabla \boldsymbol{f}(\boldsymbol{q}) - \boldsymbol{A}(\boldsymbol{q}),$$
where
$$\boldsymbol{A}(\boldsymbol{q}) = \left(\frac{\nabla \boldsymbol{f}(\boldsymbol{q}) \cdot \boldsymbol{q} - f(\boldsymbol{q})}{\boldsymbol{q}' \cdot \boldsymbol{DP}_S(\boldsymbol{q}) \cdot \boldsymbol{q}}\right) \boldsymbol{q}' \cdot \boldsymbol{DP}_S(\boldsymbol{q})$$
and $\boldsymbol{DP}_S(\boldsymbol{q})$ is the Jacobian matrix of the inverse supply.

2. *If the production function has constant returns to scale at $\boldsymbol{q}$, then the reward for each type of submission is equal to its marginal product, $\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q})$. Moreover, the allocation $(\boldsymbol{q}^h)_{h=1}^H$ attains first-best hospital welfare.*

3. *If, in addition, social cost is equal to private cost ($C^h\left(q^h\right) = SC^h\left(q^h\right)$ for all h), then this allocation attains first-best social welfare.*

The theorem characterizes rewards in a mechanism that maximizes hospital welfare. The first part shows that the reward for each submission in an optimal mechanism is approximately equal to its marginal product. The intuition is simple if we ignore the constraint that the platform cannot allocate more transplants than it produces. The platform is similar to a firm that produces a consumption good (transplants) using intermediate goods (submissions). The supply of intermediate goods is efficient when prices $\boldsymbol{p}$ are equal to marginal products $\boldsymbol{\nabla f}$. The proof is identical for kidney exchange platforms, even though there are no monetary prices paid to acquire submissions. The first order condition for the first-best aggregate supply is $\boldsymbol{\nabla C} = \boldsymbol{\nabla f}$. The marginal cost curve, which governs hospital incentives, equals the supply curve; therefore, optimal rewards are $\boldsymbol{p} = \boldsymbol{\nabla f}$.

The only complication is the constraint that a platform must allocate total rewards that equal the number of transplants produced. This constraint affects the optimal rewards vector if f does not exhibit constant returns to scale. If this is the case, the optimal rewards deviate from marginal products. The optimal level of shading for each type of submission is given by the adjustment term $\boldsymbol{A}\left(\boldsymbol{q}\right)$, which says that the platform should shade more aggressively on submissions with less elastic supply. To see why, consider the case when the cross-elasticities of supply are zero so that $\boldsymbol{DP}_S$ is a diagonal matrix. Then, for each type i, the reward is marked down from marginal product according to

$$\frac{\frac{\partial f}{\partial q_i}\left(\boldsymbol{q}\right) - p_i}{p_i} = \frac{\lambda}{\varepsilon_i},$$

where ε_i is the own-price supply elasticity and λ is the Lagrange multiplier on the constraint that all transplants produced must be given out as rewards, that is, $f\left(\boldsymbol{q}\right) = \boldsymbol{p}\cdot\boldsymbol{q}$. Our general formula is an inverse-elasticity rule, as in Ramsey (1927)'s work on commodity taxation, Boiteux (1956)'s work on regulation of monopolies, and Lerner (1934)'s work on optimal pricing with market power.

The theorem shows that current platform rules are inefficient. Instead of rewarding submissions with their marginal products, current rules reward submissions with the probability of being transplanted. Therefore, there is a wedge between the social and private benefits of submissions. Under current rules, a hospital chooses between serving their own patients or providing a service to the system as a whole. A clear example of this dilemma, described in Section 2.2, is of a hospital with two overdemanded pairs. This hospital could match the pairs internally instead of submitting them to a platform, but doing so would cause the type of inefficiency documented in Section 3.

The second part of the theorem shows that, when returns to scale are constant, the optimal mechanism rewards submissions exactly according to marginal products. The adjustment term in this case equals zero, and optimal rewards achieve first-best hospital welfare. As we will show in Section 5, this case is empirically relevant because the NKR is well within the region of approximately constant returns to scale. Therefore, optimal mechanisms can be calculated in practice by estimating marginal products.

Moreover, there is no need to consider non-linear rewards because we can achieve first-best hospital welfare by rewarding hospitals linearly. One approach for using these results in practice is to introduce a simple dynamic points mechanism. For example, for each submission, a platform can credit a hospital points equal to the marginal product. Then a point can be subtracted whenever a hospital conducts a transplant. The platform performs optimal matches with a constraint that no balance falls below a certain level. Naturally, there are important theoretical issues related to implementing incentives in this kind of mechanism without compromising efficiency. We return to these issues in Section 6.

The third part of the theorem states that if the production function exhibits constant returns to scale and there are no agency problems, then the optimal mechanism achieves first-best social welfare. This result clarifies that there are two possible sources of inefficiency: inefficient platform incentives and agency problems. Platform incentives are inefficient if rewards deviate from marginal products, $\boldsymbol{p} \neq \boldsymbol{\nabla f}$. In the platforms literature, this problem is usually attributed to wedges between the platform's goals and society's (Rochet and Tirole, 2003; Armstrong, 2006; Weyl, 2010). Agency problems exist if hospitals do not fully internalize the welfare of the parties they represent, i.e. social cost is not equal to private cost. The market functions efficiently if platform incentives are optimal ($\boldsymbol{p} = \boldsymbol{\nabla f}$) and there are no agency problems (social cost equals private cost).

Figure 5 depicts these two market failures under some regularity conditions. The current aggregate supply is $\boldsymbol{q}_0$, which is determined by rewards that equal matching probabilities. If a platform switches to an efficient mechanism, aggregate supply moves to $\boldsymbol{q}^*$. If agency problems are also solved, the market moves to the first-best aggregate supply. This quantity is denoted by $\boldsymbol{q}^{**}$ if $C(\boldsymbol{q}) = SC(\boldsymbol{q})$. The deadweight loss at any of these points is given by a (multi-dimensional) Harberger triangle between the marginal product and the marginal social cost curves.

The upshot of this analysis is that, much like in more traditional markets, many key questions about kidney exchange depend on the production function, which we turn to next.

5 Production Function Estimates and Results

We now estimate the production function using data from the largest kidney exchange platform, the NKR. We focus on the NKR because it is the dominant kidney exchange platform during our sample period (Table 1). We use these estimates to measure the total inefficiency due to market fragmentation, calculate the rewards in an optimal mechanism, and to measure the efficiency gain from moving to an optimal mechanism.

5.1 Estimation

Production functions are commonly estimated using data on inputs and outputs from several firms. The key econometric challenges in this literature are endogeneity in the chosen inputs and selection in the set of operating firms (see Marschak and Andrews, 1944; Olley and Pakes, 1996). Unfortunately, this approach is not appropriate in our setting for three reasons. First, the standard methods are best suited for low-dimensional production functions that only depend on a few inputs, such as capital and labor. These methods suffer from a curse of dimensionality if there are many input types. In our case, the vector of inputs is high-dimensional because submissions can vary in many ways. Second, commonly used functional forms such as Cobb-Douglas restrict all inputs to be substitutes, a property that is not appropriate for a matching context. Third, the standard methods depend on a panel dataset with inputs and outputs of multiple firms and exogenous variation of inputs. However, we only have data from a single large platform.

We circumvent these econometric issues by using an engineering approach based on detailed institutional knowledge and administrative data on the processes involved in organizing kidney exchange. We have detailed institutional knowledge of the operational procedures and algorithms used by kidney exchange platforms. One of us (Ashlagi) has developed the matching software for several platforms, and has worked with the NKR. Moreover, we have detailed data on NKR operations and the composition and biological compatibility of its patient pool. We use this experience and knowledge to develop a detailed simulation model of a kidney exchange platform.

We simulate the various steps involved in organizing kidney exchange to evaluate the number of transplants, $f(\boldsymbol{q};\theta)$, that can be produced with a flow of inputs $\boldsymbol{q}$ and parameters θ. The simulation is dynamic, with each period representing one day. There are four steps that take place: submissions to the platform, transplant proposal, final review and transplantation, and departure from the platform. Each step is described in detail below, and is governed by a set of parameters. The parameters governing the first and last steps are directly estimated from the NKR data; the parameters involved in the second step are known; and the parameters from the third step are calibrated to fit observed transplantation probabilities for various patient and donors types, and the average length of chains. Our estimation and calibration methods are described below, with details provided in Appendix C.

These steps and their associated parameters are as follows:

1. **Submissions, $\boldsymbol{q}$:** Hospitals submit patients and donors, either individually or in pairs, to the platform. These submissions are added to the current pool of patients and donors already registered with the exchange. Patients and doctors, at this time, can submit minimal acceptance criteria for a donor.

   Submissions arrive according to a Poisson process. The baseline arrival rates at the NKR are represented by a vector $\boldsymbol{q}_0$, with dimension equal to the number of submission types I. We estimate the daily arrival rate of each submission type i as average number of arrivals per year. An identical arrival process with Poisson arrival rates $\boldsymbol{q}$ allows us to calculate the production function at other arrival rates $\boldsymbol{q} \neq \boldsymbol{q}_0$.

Our exercises will start by treating each submission as a separate type (I =1930). We will then aggregate types to best predict probability of matching and marginal products using biological characteristics that are relevant for kidney exchange (e.g. blood-type and patient PRA).

2. **Transplant Proposal:** Each day, the NKR identifies an optimal weighted set of potential exchanges within the stock of patients and donors registered with the platform. This algorithm incorporates four constraints. First, none of the proposed transplants should be (known to be) biologically incompatible or ruled out by pre-set acceptance criteria. These constraints are directly observed in the data. Second, no donor or recipient can be involved in more than one transplant. Third, a donor who is part of a pair is only asked to donate an organ if the intended recipient has been proposed a transplant. Finally, kidney exchange platforms limit the cycle size because of logistical difficulties in organizing many simultaneous surgeries.[24]

   The parameters of this algorithm are the weights w_{jk} used by the NKR for a transplant involving donor k and patient j and the maximum cycle size. Consistent with NKR policy and observed data, we prohibit all cycles of length four or greater. The weights are known to one of the authors (Ashlagi) and are detailed in Appendix C. They prioritize unlikely matches in an attempt to utilize hard-to-match donors and transplant hard-to-match patients whenever possible. The weights typically only break ties between two matches with the same number of transplants in favor of retaining patients and donors who are likely to match in the future.

3. **Final Review and Transplantation:** Each proposed transplant is reviewed by doctors, patients, and donors, and approved before it is performed. Both approval and biological testing can take several days. Moreover, patients and donors in proposed transplants that are under review on a given day are excluded from the maximal matching algorithm on that day. This step also involves a final set of blood-tests to ensure biological compatibility.[25] Cycles in which any patient refuses or is found to be incompatible with the proposed donor are abandoned. NKR usually abandons chains in which the second patient cannot be transplanted. For other chains, all proposals until

[24] Formally, the NKR maximizes $\sum_{jk} c_{jk} w_{jk} x_{jk}$ by picking $x_{jk} \in \{0, 1\}$, where $x_{jk} = 1$ denotes a proposed transplant from donor k to patient j; w_{jk} is the weight accorded to each such transplant by the NKR; and $c_{jk} = 1$ if a transplant from k to j is feasible (biologically compatible and acceptable) and 0 otherwise. This problem is subject to three additional constraints. First, no donor or patient is involved in more than one transplant, i.e. $\sum_j x_{jk} \leq 1$ and $\sum_k x_{jk} \leq 1$. Second, if donor k and patient j belong to a pair, then $x_{j'k} = 1$ for some j' only if $x_{jk'} = 1$ for some donor k'. To write the third constraint, note that a cycle of length n is an ordered tuple, $(j_1, j_2, \ldots, j_n)$ where $x_{j_k j_{k+1}} = 1$ for $k < n$ and $x_{j_n j_1} = 1$. We impose the constraint $n \leq 3$. Because there are a very large number of cycle length constraints, we first solve a relaxed problem without this last constraint and iteratively add the constraints to prohibit large cycles. Appendix C provides further details on the algorithm.

[25] These failures are recorded by setting $c_{jk} = 0$ for future iterations if the donor k was refused by patient j.

the first failure are consummated. The donor belonging to the final patient-donor pair in such a chain may initiate new chains in the future, much like an altruistic donor. This donor is often referred to as the "bridge" donor. Consistent with NKR policy, unpaired patients are prioritized according to the net difference between altruistic donors and unpaired patients previously transplanted by the patient's hospital.

This step results in frictions within the system that reduce transplantation rates (Agarwal et al., 2018). The parameters that govern these frictions are the time required for each of the two approval steps, the probability that a proposed transplant is abandoned in each step, and the duration for which a bridge donor is retained in the pool before donating her kidney to a patient on the deceased donor list.

Unfortunately, we do not have detailed data on which transplants were refused, how often transplants were aborted due to biological testing, or how long each review phase takes. Additionally, the NKR does not seem to have clear-cut algorithmic policies on how to use bridge donors. Chains would be indefinitely long if bridge donors were allowed to initiate new chains forever but too short if bridge donors were not used. Although cases of donors reneging are rare (Cowan et al., 2017), platforms try to transplant bridge-donors quickly, to an unpaired patient if necessary, to avoid these cases.

We calibrate these parameters by simulating our model to find values that most closely replicate the match probabilities, durations, pool size and chain lengths observed in our data. We match average values of each of these variables, except for chain length, by the following submission types: altruistic donor, patient-donor pairs, and unpaired patients.[26]

Our simulations suggest that a two-week period for both the acceptance and the biological testing phases; and a one-fifth failure rate for each phase best fit these moments. Reducing the failure rates in simulations primarily increases chain length and transplantation rates, while reducing the duration of either phase increases the transplantation rates without having a large effect on chain length. For the bridge donor policy, we find that a hold-period of 30 days best fits the data.

Details on the fit of our calibrated parameters are provided in Appendix C.5.1. Further, Appendix D repeats all of our analyses under alternative parameters to examine robustness of our results.

4. **Departure:** Patients and donors often depart the NKR without a transplant. A patient and his/her associated donor may leave the platform because the patient dies, becomes too sick to transplant, or receives a kidney transplant elsewhere. Therefore, we need to estimate the probability that a patient or a donor leaves the NKR without

[26] In principle, we could have estimated these parameters using simulated minimum distance. However, a simulation for each parameter value can take weeks, making optimization over the parameter set infeasible.

a transplant.[27]

We estimate a model of departures using the registration and transplantation dates (if transplanted) for each patient and donor. Additionally, we use regular data snapshots of the patients and donors registered at the NKR to determine how long the patient or donor was registered in the NKR without a transplant. We estimate an exponential hazards model for the departure process using maximum likelihood.[28] The departure rates in the model depend on the fraction of donors (patients) ever registered with the NKR who are compatible with a patient (donor), blood-type dummies for the donor and the patient, and the patient and donor ages at registration. Appendix C.2.2 presents the estimates for the model.

This procedure allows us to evaluate a transplant production function for any vector of inputs $\boldsymbol{q}$ by simulating each of these events for each calendar day. Given any initial pool of patients and donors in the NKR, these simulations generate a Markov chain with a sequence of registrations, transplants, and departures. We initialize the NKR pool with the set of patients and donors registered on April 1, 2012, and burn-in 2,000 simulation days in each run. The dependence on the initial pool eventually fades away. We compute the time average of the total number of transplants to estimate f:

$$\hat{f}(\boldsymbol{q}) = \frac{1}{T}\sum_{t=1}^{T} y_t,$$

where T is the total number of days simulated and y_t is the total number of transplants in period t of our simulation. In what follows, we report estimates based on an average of 100 simulations. Standard errors are calculated using the non-overlapping batch means estimator described in Appendix C.4.

[27] Our approach will treat all donor departures as a lost opportunity for a transplant if a better design can use that donor for a transplant. To validate this assumption, we tried to determine the outcome of patients that were paired with a donor that leave the NKR without a transplant by matching them to the OPTN data on all living and deceased donor transplants. Our ability to follow these patients is not perfect, but approximately three-quarters of patients could be perfectly matched on the HLA-A, B, and DR loci; gender; and blood type. A majority of patients either remained untransplanted or received a deceased donor, effectively crowding out a kidney from another patient. Of those that received a living donor transplant, most received direct donations and the vast majority did not utilize a multi-hospital kidney exchange platform. These facts support our treatment of departures as an appropriate approximation.

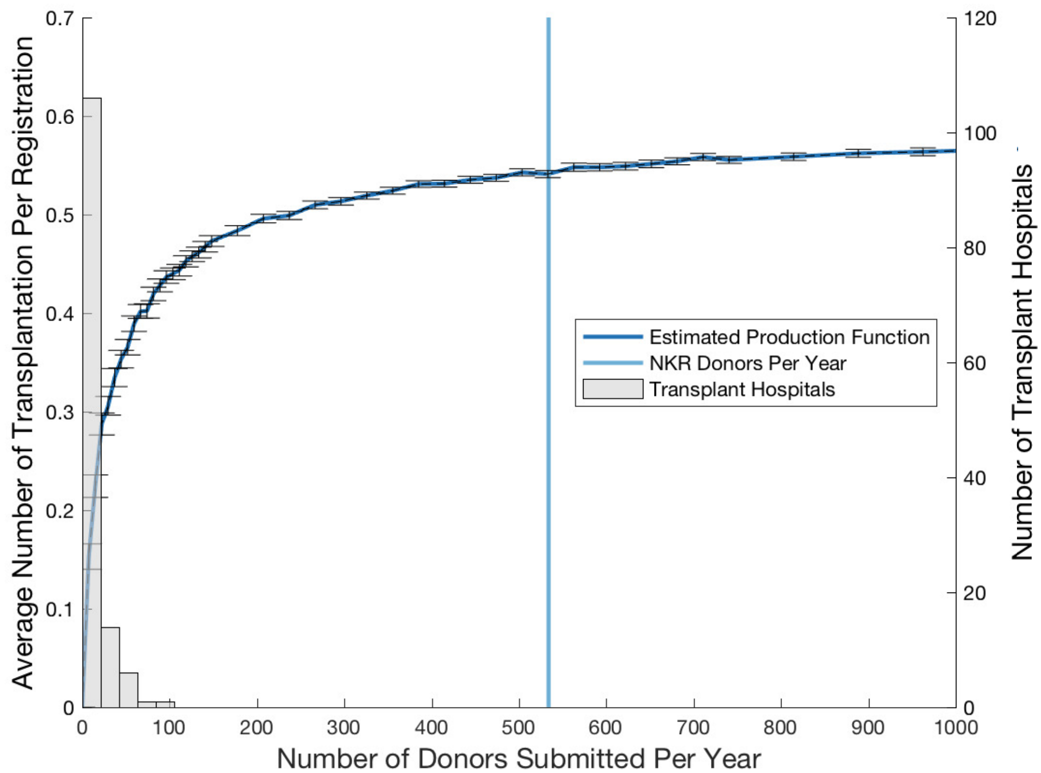
[28] Specifically, the departure rate for registration j is given by $\lambda_{g_j} \exp(z_j\beta)$, where g_j denotes whether j is an altruistic donor, a patient-donor pair, or an unpaired patient; λ_{g_j} is a group-specific constant departure risk; z_j denotes a vector of characteristics for j; and β is a conformable vector of coefficients. We use maximum likelihood using the (censored) observations of departure times for each registration in the NKR. Censoring in our dataset can occur because we only observe a lower bound for the departure time if j was transplanted or remained in the NKR pool at the end of our sample period.

5.2 Returns to scale and misallocation

5.2.1 Returns to scale

We first document the estimated returns to scale in the transplant production function, that is, how the average product changes with platform size. We evaluate the production function for pools of submissions $\boldsymbol{q}$ with the same composition as the NKR but with different scales as measured by the total flow of donors submitted per year, which we denote as $x(\boldsymbol{q})$. Figure 6 depicts average products, equal to $f(\boldsymbol{q})/x(\boldsymbol{q})$, as a function of the total flow of submitted donors $x(\boldsymbol{q})$. We choose the total flow of donors $x(\boldsymbol{q})$ for the denominator because it is the flow of transplants that a platform could perform if all donors were used in an exchange.

Figure 6: Production Efficiency versus Scale



Notes: The line plot represents the average product of a kidney exchange platform versus its scale. The histogram is based on the estimated scale of various hospitals. The left vertical axis represents average products, defined as the share of pairs and altruists who are transplanted. The right vertical axis is the scale for the histogram. The horizontal axis represents scale, measured as the yearly arrival rate of pairs and altruists. The error bars on the estimated production function show a 95% confidence interval. The plot uses the baseline parameters and the pool composition from the NKR.

The figure shows that there are increasing returns to scale, but that productivity eventually plateaus. With a scale of 534 donor arrivals per year, the NKR is well within the region of approximately constant returns to scale. The NKR has an average product of 0.54, which varies only slightly once the scale is sufficiently large. A platform that is half the size of NKR has an average product of 0.51, while a platform that is double the size has an average product of 0.57. Therefore, the market can operate at a high level of efficiency even if there are a handful of competing platforms. These estimates suggest that mergers of sufficiently large platforms would have small effects on efficiency.

Next, we use these estimates to calculate whether individual hospital platforms operate at an efficient scale. Recall that within hospital transplants collectively account for the majority of kidney exchanges. A challenge with this exercise is that we observe neither the number nor the composition of patients and donors available to a hospital. We only observe the kidney exchange transplants conducted by a hospital both through the NKR and outside. To make progress, assume, for the moment, that hospitals have the same production technology and composition as the NKR. Further, assume that hospitals conducting within-hospital transplants do not participate in the NKR. Under these assumptions, one can use the observed rate of kidney exchange transplants at individual hospitals to infer the scale for each hospital. Specifically, let y^h be the flow of within hospital kidney exchange transplants conducted at hospital h. We estimate the flow x^h of donors available to hospital h as the flow necessary to produce y^h with the same composition and technology as the NKR. That is, x^h solves $y^h = \hat{f}\left(x^h \cdot \frac{\boldsymbol{q}_0}{x(\boldsymbol{q}_0)}\right)$, where $\boldsymbol{q}_0$ is the flow of submissions received by the NKR.

This exercise suggests that almost all individual hospitals operate far below the efficient scale. The histogram in Figure 6 shows the estimated distribution of hospital scale. The median hospital has a scale of 9 donor arrivals per year. The 90th percentile is 26 donor arrivals per year. The largest, Methodist Hospital in San Antonio, has a scale of 104 donor arrivals per year. The average product at these efficient scales is 0.16, 0.30 and 0.44 transplants per donor, respectively. Thus, at our estimated production function, even the largest single-hospital platform does not operate at an efficient scale. UNOS and APD have estimated average products of 0.41 and 0.42 respectively. Hence, the implied efficiency losses are considerable even for the largest platform other than the NKR. These results are consistent with the evidence presented in Section 3.2 that hospitals often perform matches that are socially inefficient, and that UNOS and APD are also somewhat less efficient than the NKR.

5.2.2 Misallocation: inefficiency due to small production scale

We start by using the baseline approach in the previous section to estimate inefficiency due to market fragmentation. That is, we estimate how many additional transplants would be performed if the entire kidney exchange market functioned at NKR's efficiency. We use a hospital's estimated scale to calculate the difference in average product between the hospital and NKR. Because NKR operates at constant returns to scale, this difference multiplied by the hospital scale is the total number of transplants that are lost due to the hospital conducting kidney exchange at an inefficiently small scale. The aggregate lost transplants equal the total deadweight loss because our social welfare function is the total number of transplants nationwide. The estimated deadweight loss presented in Table 3 shows that

447.7 transplants are lost per year due to market fragmentation (panel A, column (1)). This number is large relative to the 800 transplants conducted through kidney exchange each year. The economic value of the lost transplants approach $500 million per year based on the Held et al. (2016) estimates of a value of a transplant. The cost savings alone are on the order of $140 million per year.

This baseline approach is simple but suffers from four potential biases. First, the composition of submissions in hospitals may differ from that in the NKR. We assess robustness to this assumption by estimating inefficiency using patient and donor compositions based on submissions from three different groups of hospitals: all hospitals, hospitals in the top quartile of intensive margin participation rate, and hospitals in the bottom quartile.[29] Hospitals in the top quartile of participation should submit a less selected pool because they conduct a larger fraction of kidney exchanges through the NKR. Estimates from a production function using patients and donors from only these hospitals allows us to assess robustness to potential compositional differences between single hospital platforms and the NKR. Columns (2) and (3) present estimates under alternative assumptions on the composition of patients and donors available to a hospital. Comparing estimates suggests that overall inefficiency exceeds 380 transplants per year under alternative assumptions on composition differences.

Second, our baseline approach assumes that all within hospital transplants are produced by hospitals in isolation of the rest of the market. The bias due to hospitals that also participate in national platforms does not have a clear direction. We address this issue by disaggregating the efficiency losses by whether a hospital participates in the NKR, APD and UNOS and by the fraction of the hospital's paired kidney exchanges that are conducted through the NKR. If we restrict attention only to the 96 hospitals that do not participate in NKR, the efficiency loss in column (1) is 212.9 transplants per year (panel C, excluding the NKR row). Some of these hospitals participate in UNOS or APD and may be producing transplants at a more efficient scale. Even if we assume that each of these hospitals that particpate in UNOS or APD produce transplants at the estimated scales for the two platforms, we estimate that the deadweight loss in column (1) would be 127.0.[30] However, this extremely conservative calculation is likely at slack for two reasons. First, even among the non-NKR hospitals that participate in either UNOS or APD, two-thirds of kidney exchange transplants are performed within hospital (panel C). The deadweight loss lower bound of 127.0 assumes that all transplants are produced at the APD/UNOS scale. Second, it ignores deadweight loss from hospitals that participate in NKR. Among the set of NKR participants, the 17 hospitals that are in the lowest quartile of fraction of transplants performed in NKR alone contribute to an efficiency loss of 94.7 transplants per year (panel D). In summary, despite potential bias due to some hospitals participating in large platforms, this decomposition suggests that a loss of 200 transplants per year is a conservative estimate for the costs of market fragmentation.

[29]We measure participation rate as the number of donors submitted to the NKR as a fraction of donors submitted to the NKR or transplanted in a within hospital kidney exchange.

[30]The deadweight loss from hospitals that do not participate in any of the three national platforms alone is 106.9. In addition, we estimate that the deadweight loss for hospitals that participate only in UNOS or APD is 20.1, assuming that all kidney exchange transplants from these hospitals are produced at a scale corresponding to the platform in which they participate.

Table 3: Total Efficiency Loss

	Number of Hospitals	Kidney Exchange Transplants Per Year	Within Hospital Kidney Exchange Transplants Per Year	Efficiency Loss Additional Kidney Exchange Transplants		
				(1)	(2)	(3)
			Panel A: All Hospitals			
All Hospitals	164	800.5	465.4	447.7	386.1	539.7
			Panel B: By hospital size (number of PKEs per year)			
Top Quartile	42	598.8	358.3	237.5	186.2	285.2
2nd Quartile	48	143.2	73.4	132.7	111.4	148.1
3rd Quartile	40	45.7	27.7	57.9	64.6	81.6
Bottom Quartile	34	12.7	6.0	19.7	23.9	24.8
			Panel C: By Platform Membership			
NKR	68	580.5	297.2	234.8	191.9	280.8
Only UNOS or APD	45	133.0	90.7	106.0	92.7	126.9
None	51	86.9	77.6	106.9	101.5	132.1
			Panel D: By NKR Participation Rate (Fraction of PKEs facilitated through the NKR)			
Top Quartile	17	65.2	8.2	14.5	13.9	17.9
2nd Quartile	17	102.3	27.0	44.2	38.8	51.3
3rd Quartile	17	196.7	98.2	81.5	66.6	97.4
Bottom Quartile	17	216.2	163.8	94.7	72.7	114.1

Notes: Column (1) assumes that the typical transplant hospital has a composition of patient-donor pairs and altruistic donors given by the average registration in the NKR. Column (2) assumes the composition in transplant hospitals using only the hospitals with the top quartile of participation rates in the NKR. Column (3) assumes a composition based on hospitals with the lowest quartile of participation rates. Transplants per year is calculated using data between April 1, 2012 and December 4, 2014.

Third, hospitals may use a different matching technology than the NKR. For example, Bingaman et al. (2012) report that Methodist Hospital in San Antonio, which is now perhaps the most sophisticated single-hospital program, initially used a Microsoft Access Database and that their algorithm was "stratified by ABO compatibility and then by HLA compatibility." Such algorithms are less efficient than the linear-programming algorithms used by the NKR.[31] On the other hand, single-hospital programs face simpler logistical constraints, which may increase their productivity vis-à-vis our estimates. The direction of this bias is not signed in general, but it is more likely that single-hospital platforms are less efficient than our estimated production function.[32]

Fourth, these exercises keep the patients and donors interested in kidney exchange fixed. However, this flow is endogenous and affects the magnitude of the deadweight loss. Although

[31] In 2013, Methodist Hospital in San Antonio adopted software written by one of us (Ashlagi).

[32] See Agarwal et al. (2018) for an analysis of how various logistics influence the productivity of a kidney exchange platform. NKR's practices are optimized to maximize the number of transplants given the available patients and donors.

the direction of this bias is ambiguous, our baseline approach likely yields a conservative estimate of overall market inefficiency. The most likely bias is due to hospitals valuing transplants at less than the social value and, due to administrative costs, being likely to expend inefficiently low effort in recruiting patients and donors. If incentives were optimal, hospitals may try to recruit more – and more valuable – donors into kidney exchange. Our approach does not account for this margin because we do not observe recruitment efforts and we are therefore likely to underestimate overall market inefficiency.

Table 3 also points to which types of hospitals concentrate most of the inefficiency. Consider column (1) and, for the purposes of this decomposition, ignore the biases discussed above. Even though they perform internal exchanges more efficiently, large hospitals account for most of the inefficiency because their market share is higher (panel B). Indeed, 53.0% of the losses come from hospitals in the top quartile of kidney exchange transplant numbers. Moreover, both the intensive and extensive margins of participation are important. A little less than half of the efficiency losses are due to hospitals that do not participate in the NKR at all, and a quarter from hospitals that do not participate in any of the national platforms (panel C). Among hospitals that do participate in the NKR, a large share of the efficiency loss is due to the hospitals with low participation (panel D).

To summarize, although the baseline estimate of 447.7 lost transplants is potentially biased, a battery of robustness exercises suggest the deadweight loss from market fragmentation is large. These losses arise from all types of hospitals. The most conservative estimates place this loss at over 200 transplants a year. Additionally, these estimates do not appear to be driven by potential compositional differences in the kidney exchange pool. Table D6 in Appendix D further evaluates these results' robustness to alternative choices for the production function parameters that were calibrated. Across various specifications, we continue to find that an estimated 200 lost transplants is conservative. These results are consistent with our descriptive finding that hospitals often perform inefficient matches.

5.3 Optimal rewards

5.3.1 Inefficiency of current mechanisms

Theorem 1 shows that optimal rewards are approximately equal to marginal products. That is, $\boldsymbol{p}^* = \boldsymbol{\nabla f}(\boldsymbol{q}^*) - \boldsymbol{A}(\boldsymbol{q}^*)$, where $\boldsymbol{q}^*$ and $\boldsymbol{p}^*$ are the aggregate quantities and rewards that maximize hospital welfare. We will test this equality at the aggregate supply $\boldsymbol{q}_0$ and rewards $\boldsymbol{p}_0$ in our data.

Current rewards, $\boldsymbol{p}_0$, equal the probabilities of matching each kind of submission. These probabilities can be easily estimated from our simulations, and the estimated probabilities closely match the probabilities in the data (see Appendix C.5.1). Marginal products $\boldsymbol{\nabla f}(\boldsymbol{q}_0)$ can be estimated by numerically differentiating the production function. In principle, calculating the adjustment term requires estimates of the supply elasticity matrix, which is not feasible with our data. But the adjustment term is small because returns to scale are approximately constant for NKR's size. Therefore, optimal rewards are approximately equal to marginal products. Formally, Theorem 1 implies that the quantity-weighted average of the adjustment term is

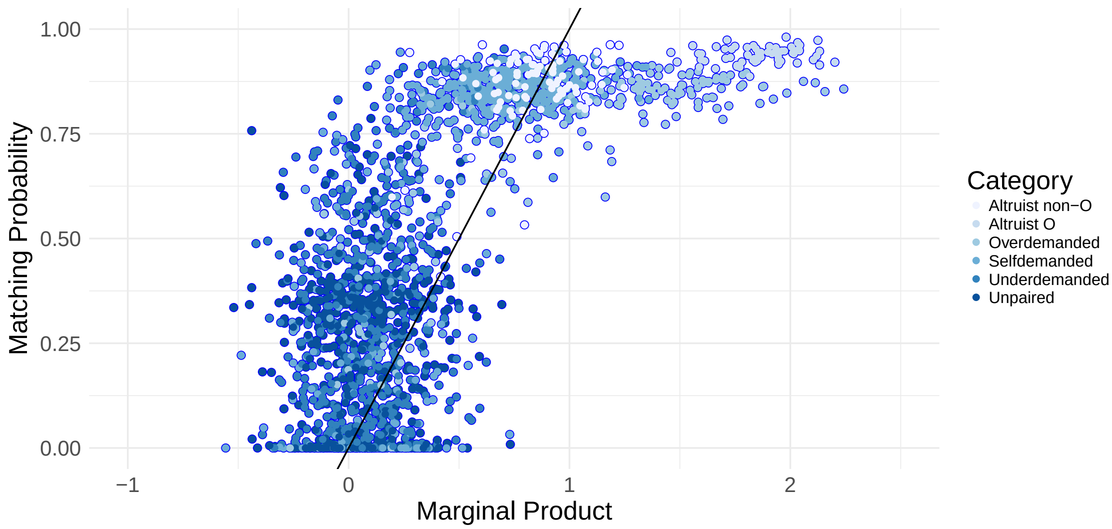
$$\frac{\boldsymbol{A}(\boldsymbol{q})\cdot\boldsymbol{q}}{\|\boldsymbol{q}\|_1}=\frac{\boldsymbol{\nabla f}(\boldsymbol{q})\cdot\boldsymbol{q}-f(\boldsymbol{q})}{\|\boldsymbol{q}\|_1}.$$

That is, the average level of shading is the difference between the average marginal product and the average product. Evaluating this formula using using the estimated production function and numerical derivatives for each of the 1930 submission types yields an average shading of only 2.16×10^{-4}. In what follows, we simply approximate optimal rewards with marginal products.

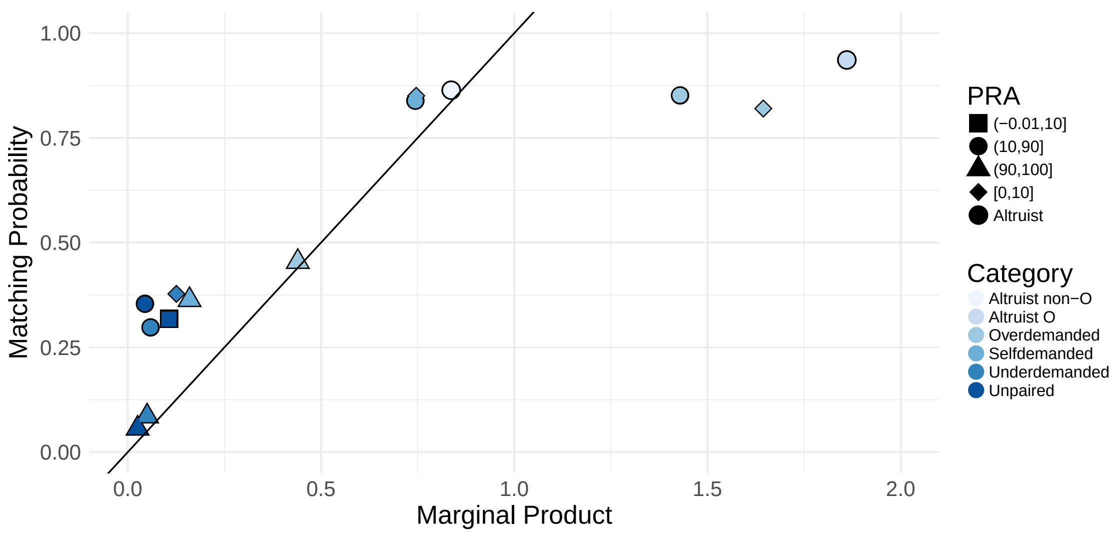
Figure 7a plots current rewards (the probabilities of matching $\boldsymbol{p}_0$) versus optimal rewards (marginal products $\boldsymbol{\nabla f}(\boldsymbol{q}_0)$). Some of the 1930 types have negative estimated marginal products because the matching algorithms are myopic, which can result in crowding out of future transplants. However, negative point estimates can also result from noise due to simulation error. Figure 7b aggregates these estimates with categories constructed by using the classification into under-demanded, over-demanded and self-demanded types based on Roth et al. (2007), split by the immune sensitization levels of the patient. These aggregated marginal products and match probabilities are estimated more precisely.

The marginal products are qualitatively similar to the Roth et al. (2007) theoretical predictions discussed in Section 2. The marginal product of an underdemanded pair is 0, both in our estimates and in the Roth et al. (2007) model. The estimates differ for other types. For example, the marginal product of an overdemanded pair with low sensitization is 2 in the Roth et al. (2007) model, but 1.64 in our estimates. One reason for this difference is that, in our data, these pairs are only matched with probability 0.82. Our empirical model also refines the predictions from the theoretical models by showing how marginal products vary with sensitization. For example, the marginal products of overdemanded and self-demanded pairs are considerably lower if these pairs are sensitized. These finer results can be important when designing practical mechanisms.

Both figures show a large wedge between current and optimal rewards. If current rewards were optimal, all points on these two figures would be on the 45° line. Altruistic donors and over-demanded pairs with low PRA are far below this line. Over-demanded pairs with low sensitization have marginal products of 1.64, but the probability of matching them is only 0.82. Even more extreme, altruistic O donors have a marginal product of 1.86, but their probability of matching is only 0.94. Therefore, hospitals are not rewarded enough for submitting these types, and it may explain why we see relatively few of these types are submitted to the NKR. Other submission types are drastically over-priced. Under-demanded pairs with low sensitization have marginal products of approximately 0.13 but have a probability of being matched of around 0.38. Similarly, unpaired patients have low marginal products but a significant probability of being matched. These differences suggest the platform can do considerably better by increasing rewards to the productive and undervalued submissions while reducing rewards to the unproductive submissions.



(a)



(b)

Figure 7: Private versus Socially Optimal Rewards for Submission Types

Notes: The vertical axis is the probability of a submission being matched, which are the private rewards that hospitals receive according to current exchange rules. The horizontal axis plots the marginal product of a submission, which equals the social contribution of the submission in terms of transplants. Each point corresponds to a submission in the data. Matching probabilities and marginal products are calculated in the baseline simulation. Marginal products are measured with substantial noise at the individual level because, due to computational reasons, each individual derivative uses a small number of simulation days. In the aggregated version different dots of the same color correspond to the different PRA levels. Figure 7a shrinks the estimated marginal products and match probabilities towards to the group means following the procedure recommended by Morris (1983).

5.3.2 Approximately efficient point mechanisms

Theorem 1 and the small adjustment term $\boldsymbol{A}$ suggest that platforms should set rewards close to marginal products. We will now show that marginal products are highly predictable using a small number of patient and donor categories. Then, we discuss how to design point mechanisms that are both approximately efficient and simple enough for practical application.

We use a regression tree to construct categories that best predict marginal products. We allowed the tree to depend on the patient's PRA, submission type (altruistic, patient-donor pair, unpaired patient), and ABO blood type. Figure 8 shows the categories found by a standard algorithm for finding the best cross-validated predictor for the marginal products. These categories are intuitive as they split submissions based primarily on submission type, whether or not the patient/donor is blood type O, and on immune sensitivity. The procedure chose a tree with few leaves. The within-category mean marginal products $\boldsymbol{\nabla f}$ and probabilities of matching $\boldsymbol{p}_0$ are dispersed relative to the (appropriately shrunk) within-category standard deviation. This suggests that marginal products and probabilities of matching are approximated with a small number of categories.

A mechanism that assigns points based on these categories can be explained to participants with this tree or a simple table (for example, Table C5). Points could be awarded when the hospital submits a patient and/or donor to the NKR, or at the time of transplantation. A point should then be subtracted whenever the hospital conducts a transplant for one of its patient because it is the numeraire in our model. Rewards at submission raise the possibility that hospitals will make shill submissions. This reasoning suggests that points should be awarded at the time of transplantation. In this case, the marginal products should be divided by the probability of matching $\boldsymbol{p}_0$ in order to implement identical expected rewards. As before, a point is deducted for each transplant the hospital conducts. The optimal points awarded at the time of transplantation is denoted as $\boldsymbol{r}^*$ in Figure 8. We postpone the discussion of implementation details to Section 6.

5.3.3 Welfare gains from optimal point mechanisms

We now estimate the gain in welfare from moving to the point mechanism described above. This gain is equal to the deadweight loss that can be avoided by rewarding hospitals optimally as in Theorem 1. We begin by considering the gain in hospital welfare and later consider the gain in social welfare.

The deadweight loss is given by a multidimensional version of the standard Harberger triangle from linear commodity taxation. Figure 9 depicts the current aggregate supply $\boldsymbol{q}_0$, the current rewards $\boldsymbol{p}_0$, the current marginal products $\boldsymbol{\nabla f}_0$, and the optimal aggregate supply $\boldsymbol{q}^*$. The hospital deadweight loss $W^H(\boldsymbol{q}^*) - W^H(\boldsymbol{q}_0)$ equals the area between marginal product curve $\boldsymbol{\nabla f}$ and marginal cost/supply curve $\boldsymbol{\nabla C} = \boldsymbol{P}_S$ between $\boldsymbol{q}_0$ and $\boldsymbol{q}^*$. Therefore, this deadweight loss is the integral of $\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{P}_S(\boldsymbol{q})$ as $\boldsymbol{q}$ goes from $\boldsymbol{q}_0$ to $\boldsymbol{q}^*$. This calculation is the multidimensional version of the Harberger triangle formula, that is, the area between the marginal benefit and marginal cost curves.

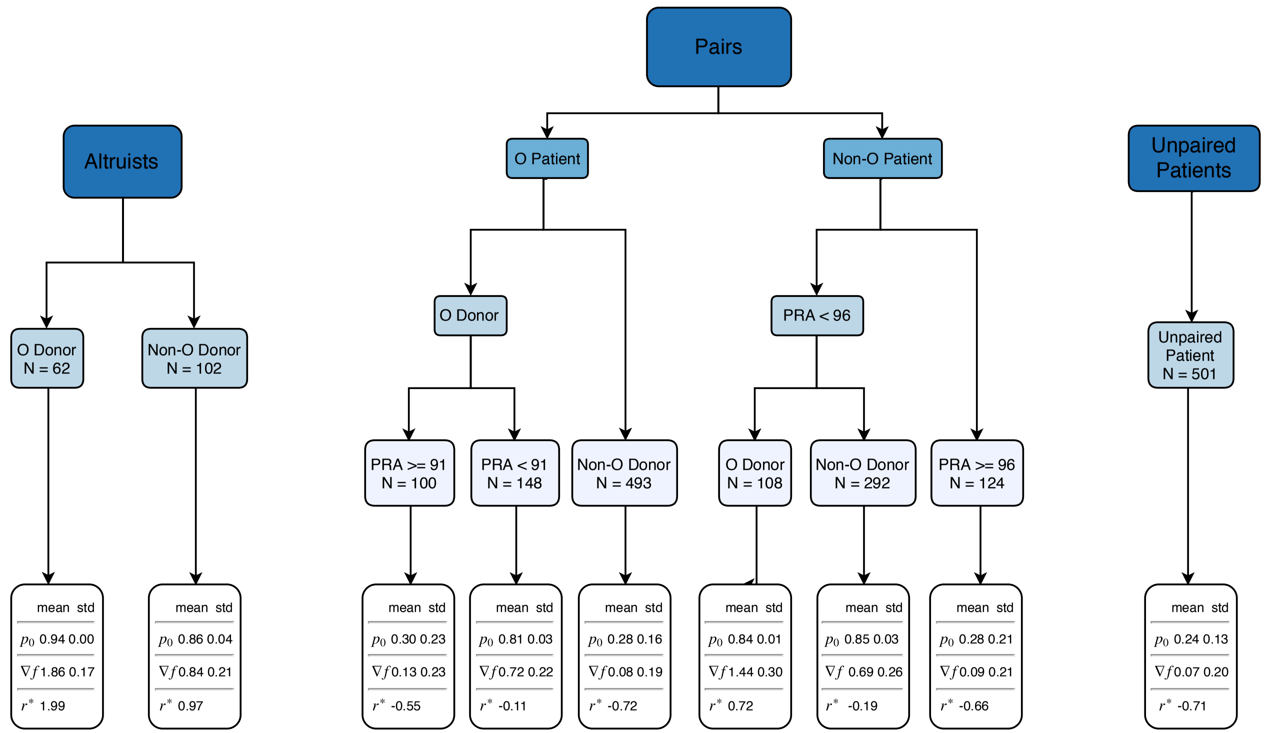


Figure 8: Regression Tree for Marginal Products

Notes: Categories are determined by regression tree analysis to predict marginal products as a function of whether a submission is a pair an altruistic donor, or an unpaired patient; blood types; and the patient's PRA. Our procedure followed standard recommendations in Friedman et al. (2001). Specifically, we used 10-fold cross-validation to pick the penalty parameter on the number of nodes, required each leaf to have at least 20 observations and pruned a leaf if it did not increase the overall fit by at least 2%. Standard errors for the simulations are calculated by following Chapter 12 of Robert and Casella (2004). The within category standard deviation is estimated using shrinkage methods recommended in Morris (1983). $\boldsymbol{p}_0$ denotes the match probabilities in the current mechanism, $\boldsymbol{\nabla f}$ denotes marginal products, and $\boldsymbol{r}^*$ denotes the optimal rewards at transplantation. We calculate $\boldsymbol{r}^*$ by dividing $\boldsymbol{\nabla f}$ by $\boldsymbol{p}_0$ and subtracting 1 for all types except altruistic donors.

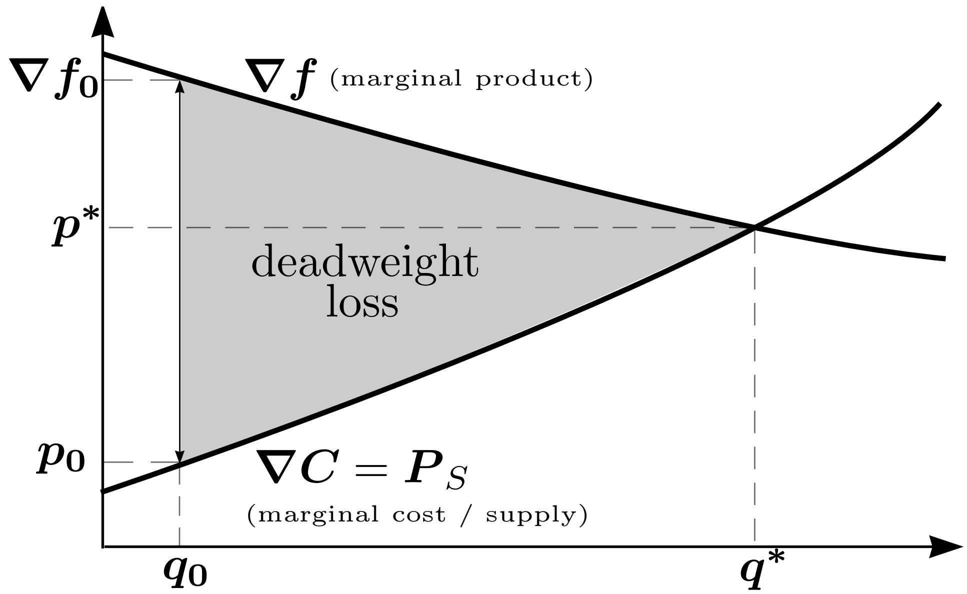


Figure 9: Hospital-Welfare Deadweight Loss from the Current Mechanism

Notes: The horizontal axis represents aggregate quantity and the vertical axis represents rewards vectors, marginal costs and marginal products, so both axes represent I-dimensional vectors. The deadweight loss from the current mechanism is the shaded area between marginal products and the supply curve (mathematically, the area is a path integral going from current rewards $\boldsymbol{p}_0$ to optimal rewards $\boldsymbol{p}^*$). Current rewards are $\boldsymbol{p}_0$, equal to the probability of matching each type of submission, while optimal rewards $\boldsymbol{p}^*$ equal marginal products. Current quantities $\boldsymbol{q}_0$ and rewards $\boldsymbol{p}_0$ are observed. Marginal products $\boldsymbol{\nabla f}$, including the current value $\boldsymbol{\nabla f}(\boldsymbol{q}_0)$, can be calculated from the production function. In contrast, the supply curve $\boldsymbol{P}_S = \boldsymbol{\nabla C}$ and optimal rewards $\boldsymbol{p}^*$ and quantities $\boldsymbol{q}^*$ are not observed and depend on the elasticity of supply.

Proposition 1. *Consider an aggregate supply of pairs, $\boldsymbol{q}_0$, that results when hospitals choose supply optimally given rewards, $\boldsymbol{p}_0$. Further, consider strictly positive aggregate supply, $\boldsymbol{q}^*$, and rewards, $\boldsymbol{p}^*$, that maximize hospital welfare as in Theorem 1. Assume that the matrix $\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)$ is finite and non-singular and that the production function has constant returns to scale at $\boldsymbol{q}^*$. Then, the deadweight loss in hospital welfare at $\boldsymbol{q}_0$ can be approximated by either*

$$\frac{1}{2}\left[\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0\right] \cdot \left(\boldsymbol{q}^* - \boldsymbol{q}_0\right).$$

or

$$\frac{1}{2}\left[\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0\right] \cdot \left[\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)\right]^{-1}\left[\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0\right]'. \tag{3}$$

The error in both approximations is $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$.

These formulas are a multidimensional version of the Harberger triangle formulas in one-dimensional linear commodity taxation. The first formula is the multidimensional version of the one half base times height formula. The second formula is the equivalent of the one half of the tax wedge squared times the inverse of the difference between the derivative of inverse supply and the derivative of inverse demand. The second formula shows that the deadweight loss is one half of a quadratic expression in the wedge $\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0$. The term $\boldsymbol{DP}_S(\boldsymbol{q}_0)$ accounts for the fact that a more elastic supply leads to larger deadweight losses. The term $\boldsymbol{D}^2\boldsymbol{f}$ accounts for the change in marginal products in response to a change in $\boldsymbol{q}$. For example, the deadweight loss is lower if increasing the supply of overdemanded pairs results in these pairs becoming less useful.

The proposition shows that estimating the deadweight loss requires estimates of $\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0$ and either $\boldsymbol{q}^* - \boldsymbol{q}_0$ or $\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)$. We can estimate $\boldsymbol{\nabla f}_0$, $\boldsymbol{p}_0$, and $\boldsymbol{q}_0$ using the estimated production function. Unfortunately, because we do not have a good estimate of the hospital supply curve, we cannot directly estimate $\boldsymbol{q}^*$ or $\boldsymbol{DP}_S(\boldsymbol{q}_0)$. Nevertheless, the large wedge between the current private and social incentives suggests the deadweight loss is significant unless the supply elasticity is extremely small.

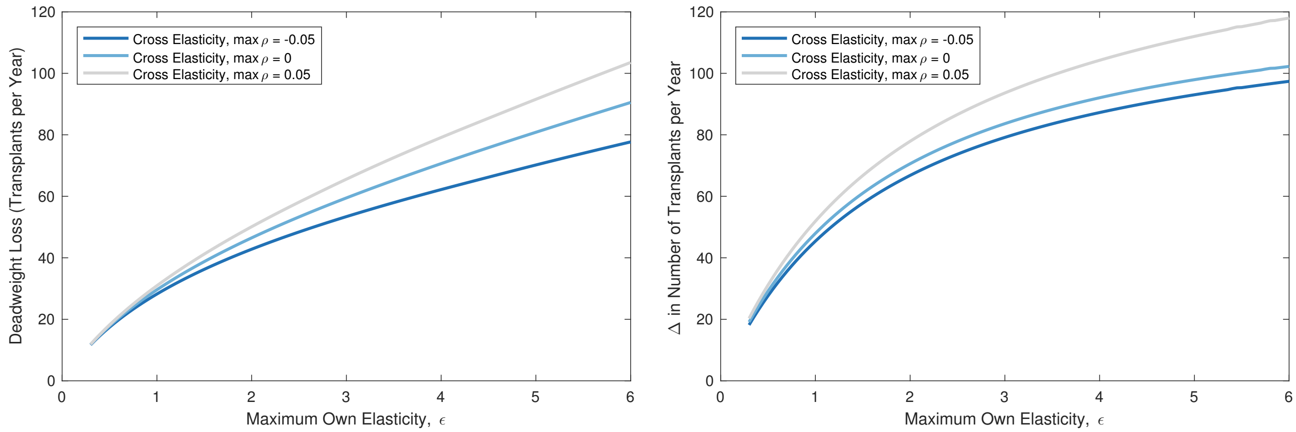
Equation (3) can be used to formalize this point by quantifying the deadweight losses for a range of supply elasticities. We restrict attention to mechanisms that set reward vectors for the categories in the regression tree analysis above (Figure 9). The wedge $\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0$ and the curvature matrix $\boldsymbol{D}^2\boldsymbol{f}$ for these categories are estimated using our production function. To use equation (3), we need to specify supply elasticities through the matrix $\boldsymbol{DP}_S(\boldsymbol{q}_0)$. One challenge in directly specifying this quantity is that different submission types may respond differently to rewards. For example, the submission of hard-to-match types to the system may not substantially decrease when rewards are lowered because there are few other avenues for matching them. Our approach is to calculate the maximum deadweight loss under varying bounds on the maximum elasticity of any type of submission. This method allows us to be agnostic about the supply elasticities of different submission types. The deadweight loss is zero when we assume that the maximum elasticity is zero because the submissions will not respond to the rewards system, resulting in $\boldsymbol{q}^* = \boldsymbol{q}_0$. As we increase the bound on the elasticity, submissions respond and the maximum implied deadweight loss increases. Further, we repeat this exercise for varying assumptions on cross-elasticities.[33]

Figure 10a plots the maximum hospital deadweight loss for bounds on the own-price elasticities ranging from 0 to 6. The curve in the middle describes the results for zero cross-price elasticities, and the other two curves present results for non-zero cross-price elasticities. The hospital deadweight loss is zero if supply is perfectly inelastic and is increasing in elasticity. The deadweight loss is significant for most of this range and above 40 transplants per year if the maximum elasticity is at least 2. For very high elasticities, the deadweight loss in-

[33] Specifically, we solved the problem

$$\begin{aligned}
\max_{\boldsymbol{DP}_S(\boldsymbol{q}_0)} \quad & \tfrac{1}{2}(\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0)[\boldsymbol{DP}_S(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)]^{-1}(\boldsymbol{\nabla f}_0 - \boldsymbol{p}_0)' \\
\text{s.t.} \quad & \left(\frac{\partial P_{S,j}}{\partial q_j}\right)^{-1}\frac{p_{0,j}}{q_{0,j}} \in [0,\varepsilon]\,, \text{ for all } j \in 1\ldots I, \\
& \left(\frac{\partial P_{S,k}}{\partial q_j}\right)^{-1}\frac{p_{0,k}}{q_{0,j}} = \rho\frac{\left(\frac{\partial P_{S,j}}{\partial q_j}\right)^{-1}\frac{p_{0,j}}{q_{0,j}} + \left(\frac{\partial P_{S,k}}{\partial q_k}\right)^{-1}\frac{p_{0,k}}{q_{0,k}}}{2}, \text{ for all } j,k \in 1\ldots I \text{ with } k \neq j.
\end{aligned}$$

for each value of the bound on elasticities, ε.



(a) Hospital-Welfare Deadweight Loss (b) Transplants Lost in the Platform

Figure 10: Losses Due to the Current Mechanism

Notes: Estimated losses from the current mechanism, using the approximation from Proposition 1, as a function of the elasticity matrix of supply. Maximum own-elasticities are in the horizontal axis. The parameter ρ governs the cross-price elasticity of supply as formally described in footnote 33.

creases at a slower rate because of production function curvature. The deadweight loss at an elasticity of 6 is only between 75 and 105 because the marginal products of the productive types that the optimal mechanism attracts decrease with supply. Although the results for large elasticities are subject to greater approximation error, it is unlikely that the deadweight losses come close to the efficiency loss relative to the first-best allocation, even for elasticities of about 6.

The hospital deadweight losses will understate the loss in social welfare if hospitals undervalue transplants. Figure 10b shows the total increase in transplants facilitated by the NKR if it adopts the optimal points system. To do this, we added the area under $\boldsymbol{P}_S = \boldsymbol{\nabla C}$ to the hospital deadweight loss numbers calculated above (see Figure 9). Because a transplant increase at the NKR will come at the cost of fewer transplants at hospitals, this calculation overstates the loss in total welfare from the current mechanism. Not surprisingly, the estimated losses are higher than the previous figure. A little over 40 transplants are lost if the maximum elasticity is 1. This number is between 95 and 120 for an elasticity of 6. Therefore, social deadweight loss is higher than hospital deadweight loss, but the two are qualitatively similar.

Taken together, these results imply that addressing the inefficient platform incentives has a large positive impact unless the elasticity of supply is extremely low. While we do not have quasi-experimental evidence on the magnitude of elasticities, the evidence in Section 3 is typical of markets with elastic supply. Most hospitals only register a subset of their patients with the NKR, and many other hospitals do not participate. These observations are consistent with many hospitals being on the margin, suggesting that hospitals respond to incentives and that supply is at least moderately elastic. Therefore, optimal point mechanisms are not only low-dimensional but also likely to have a substantial effect on the total number of transplants.

5.4 Importance of agency problems and inefficient platform incentives

We now discuss the quantitative importance of the two market failures identified above. While we cannot decompose the effects of each market failure, the results give us useful information on whether these market failures are important.

First, the misallocation analysis yields a conservative lower bound for the deadweight loss of about 200 transplants per year. The true deadweight loss is potentially much larger as most specifications yield numbers approximately twice as large. Therefore, it must be the case that at least one market failure is quantitatively important.

Second, the Harberger triangle analysis shows that inefficient platform incentives significantly reduce hospital welfare if supply is not inelastic. Moreover, if there are agency problems, the gains in social welfare from an optimal mechanism will be even higher because hospitals undervalue transplants. Specifically, hospital welfare deducts the transplant-denominated private cost incurred when hospitals provide more submissions to the platform. When there are agency problems, these private costs are significantly inflated relative to social costs.

Our results imply that agency problems are important unless elasticities are extremely high. Under the hypothesis that there are no agency problems, hospital welfare equals social welfare, and the optimal mechanism reaches first-best welfare (Theorem 1). Thus, the total deadweight loss in the misallocation analysis must be completely accounted for by the deadweight loss in the Harberger triangle analysis. Yet, even for a high elasticity of 6, the Harberger triangle yields a deadweight loss of at most 120 transplants, still below our lower bound result of 200 transplants from the misallocation analysis. The only way these estimates can overlap is if we have high elasticities and the approximation in Proposition 1 is significantly downward biased. The bias in the approximation depends on the deviation of the production function minus the aggregate cost function from the quadratic Taylor approximation, so that the bias is high if $\boldsymbol{\nabla f} - \boldsymbol{\nabla C}$ is extremely convex. Thus, attributing all the deadweight loss to inefficient platform incentives requires that elasticities are high, $\boldsymbol{\nabla f} - \boldsymbol{\nabla C}$ is sufficiently convex, *and* the downward bias in the estimated lower bound on inefficiency is small.

The upshot is that a policy that addresses either market failure is likely to be valuable and generate gains in the order of hundreds of transplants per year. Except under extreme assumptions about the supply function, there are significant gains both from implementing more efficient mechanisms and from solving agency problems.

6 Theoretical Extensions and Discussion

6.1 Maximizing social welfare

Theorem 1 describes mechanisms that maximize hospital welfare. A natural alternative would be to use mechanisms that directly maximize social welfare. However, since hospitals consider private rather than social cost in response to a rewards vector, they won't necessarily choose submissions vectors that minimize the aggregate social cost. To account for this subtlety, define the **reward-moderated social cost** by

$$SC(\boldsymbol{q}) = \sum_{h=1}^{H} SC^h\left(\boldsymbol{S}^h\left(\boldsymbol{P}_S(\boldsymbol{q})\right)\right).$$

We assume that each hospital's supply is single-valued to ensure that this function is well-defined.

The reward-moderated social cost represents the aggregate social cost of inducing an aggregate supply $\boldsymbol{q}$ by using a linear and anonymous rewards scheme. Our next result describes the rewards in mechanisms that maximize social welfare.

Proposition 2 (Optimal Rewards for Maximizing Social Welfare)**.** *Consider a vector of rewards* $\boldsymbol{p}$ *and an allocation* $\left(\boldsymbol{q}^h\right)_{h=1}^{H}$ *with strictly positive aggregate quantity* $\boldsymbol{q} = \sum_h \boldsymbol{q}^h$ *that maximize social welfare subject to all hospitals choosing* $\boldsymbol{q}^h \in \boldsymbol{S}^h(\boldsymbol{p})$ *and subject to the total rewards allocated being the same as the number of transplants produced, that is,* $f(\boldsymbol{q}) = \boldsymbol{p}\cdot\boldsymbol{q}$*. Assume the production function has constant returns to scale at the optimal* $\boldsymbol{q}$*, and* $SC(\boldsymbol{q})$ *is differentiable. Then:*

1. *The platform rewards each type of submission with its marginal product plus an adjustment term.*
$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) + \boldsymbol{A}^{SW}(\boldsymbol{q}),$$
where
$$\boldsymbol{A}^{SW}(\boldsymbol{q}) = \frac{1}{1+\lambda^{SW}}\left[\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})\right] - \frac{\lambda^{SW}}{1+\lambda^{SW}}\boldsymbol{q}'\boldsymbol{DP}_S(\boldsymbol{q}).$$
and
$$\lambda^{SW} = \frac{\left[\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})\right]\cdot\boldsymbol{q}}{\boldsymbol{q}'\boldsymbol{DP}(\boldsymbol{q})\cdot\boldsymbol{q}}.$$
2. *The adjustment term* $\boldsymbol{A}^{SW}(\boldsymbol{q})$ *can be non-zero even with constant returns to scale at* $\boldsymbol{q}$*.*
3. *The allocation* $(\boldsymbol{q}^h)_{h=1}^{H}$ *maximizes* $f(\boldsymbol{q}) - SC(\boldsymbol{q})$ *if and only if, at the optimum, the average wedge between marginal cost and marginal reward-moderated social cost,* $\left[\boldsymbol{\nabla C}(\boldsymbol{q}) - \boldsymbol{\nabla SC}(\boldsymbol{q})\right]\cdot\boldsymbol{q}$*, is zero.*

Part 1 shows the optimal mechanism rewards submissions by their marginal products plus an adjustment term. The adjustment equals an externality term, which is greater for submissions whose marginal social costs are less than their marginal private costs, minus a shading term that depends on elasticities. At the optimum, hospitals are rewarded for their marginal contributions to the platform as well as to compensate them for any components of marginal private cost that aren't a part of marginal social cost. However, if there are not enough transplants to pay for these differences, the planner has to shade rewards. As in optimal linear commodity taxation, it is also better to shade rewards for submissions with more inelastic supply.

Part 2 shows that the key difference in this case, relative to Theorem 1, is that the adjustment term is not zero, even for constant returns to scale. Therefore, the optimal rewards depend on more information. To set optimal rewards, one must know, for each type of submission, the wedge between marginal private and social costs. Such knowledge requires identifying the submission types for which hospital objectives deviate most from social objectives. Moreover, one needs to know the elasticity matrix in order to measure how much shading must be done for each submission type. Elasticities matter so long as the average wedge between marginal private and social cost is non-zero because it results in the multiplier λ^{SW} being non-zero and an adjustment term that depends on elasticities. Finally, part 3 shows that the optimal reward vector does not attain first-best social welfare. Therefore, allocations that achieve first-best social welfare require non-linear and complex incentives for hospitals.

Taken together, using only the kidney exchange mechanism to maximize social welfare, as opposed to hospital welfare, runs into important challenges. Optimal rewards are more complex, depend on more information, and are sensitive to changes in the incentives facing hospitals that can affect overall externalities. These results suggest that solving agency problems is an important complement to improving the design of the kidney exchange mechanism.

6.2 Competing platforms

Two natural policy responses to the fragmentation are to mandate participation in a single platform or to merge platforms. These recommendations raise questions about the optimal strategy for competing platforms and the efficiency costs of imperfect competition. We now consider a platform that maximizes the number of transplants it facilitates.

Proposition 3 (Oligopolistic Platforms)**.** *Consider a platform facing a smooth inverse residual supply curve of submissions $\boldsymbol{P}_{RS}(\cdot)$. Consider a vector of rewards $\boldsymbol{p}$ and a strictly positive aggregate quantity $\boldsymbol{q}$ that maximize the number of transplants in the platform subject to $\boldsymbol{p} = \boldsymbol{P}_{RS}(\boldsymbol{q})$ and subject to allocating the same number of transplants that are produced, that is, $\boldsymbol{p} \cdot \boldsymbol{q} = f(\boldsymbol{q})$. Assume the production function has constant returns to scale at the optimal $\boldsymbol{q}$. Then:*

1. *The platform rewards each type of submission with its marginal product, plus an adjustment term,*
$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) + \boldsymbol{A^C}(\boldsymbol{q}),$$
where
$$\boldsymbol{A^C}(\boldsymbol{q}) = \left(\frac{\boldsymbol{q'DP}_{RS}(\boldsymbol{q}) \cdot \boldsymbol{q}}{f(\boldsymbol{q})}\right) \boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{q'DP}_{RS}(\boldsymbol{q}).$$

2. *The adjustment term $\boldsymbol{A^C}(\boldsymbol{q})$ can be non-zero, even with constant returns to scale at $\boldsymbol{q}$. In particular, rewards are different from the rewards in Theorem 1.*

3. *If residual supply is perfectly elastic, so that the matrix* $\boldsymbol{DP_{RS}}$ *is zero, then rewards equal marginal products, as in Theorem 1.*

The proposition shows that a platform that maximizes the number of facilitated transplants does not set socially efficient rewards. Instead of setting rewards equal to marginal products, the platform subsidizes submissions that have elastic supply and are very productive. To see this, consider the simplest case, where $\boldsymbol{DP}_{RS}$ is a diagonal matrix (i.e. all cross-elasticities of residual supply are zero). Then, for each type i, the reward is marked down from marginal product according to an analogue of the Lerner index formula,

$$\frac{\frac{\partial f}{\partial q_i}(\boldsymbol{q}) - p_i}{p_i} = \frac{1}{\varepsilon_i^{RS}} - \frac{1}{\lambda} \cdot \frac{\frac{\partial f}{\partial q_i}(\boldsymbol{q})}{p_i},$$

where ε_i^{RS} is the own-price residual supply elasticity and λ is the Lagrange multiplier on the constraint $f(\boldsymbol{q}) = \boldsymbol{P}_{RS}(\boldsymbol{q}) \cdot \boldsymbol{q}$. The expression shows the platform has incentives to skew the rewards: optimal markdowns are larger for submissions with low elasticities and submission categories that are less productive on the margin.

The proposition implies that competing, empire-building platforms exploit their market power and set rewards inefficiently. Additionally, the proposition implies that platforms set efficient rewards if the market is very competitive. Optimal rewards are close to marginal products if residual supply is very elastic, i.e. if ε_i^{RS} is close to infinity or, more generally, $\boldsymbol{DP}_{RS}$ is close to zero.

6.3 Implementing a point mechanism

Our steady-state model shows that a mechanism that rewards hospitals with marginal products is efficient. Moreover, our empirical results suggest that a low-dimensional point mechanism would likely achieve sizable efficiency gains. Unfortunately, our simplified steady-state model does not specify an extensive form game. Therefore, our model cannot be used to fully specify optimal mechanisms or game forms, or to evaluate them. This raises practical and theoretical questions about how to design and implement a dynamic points mechanism, a task that requires detailed specification of rules and an analysis of resulting incentives. While resolving all these details is beyond the scope of our paper, we discuss some key theoretical and practical issues.

In both theory and practice, a natural mechanism for solving this problem is the point system described in Section 5.3.2. A motivation for this kind of mechanism comes from the dynamic mechanism design literature. Möbius (2001), Hauser and Hopenhayn (2008), Friedman et al. (2006), and Guo and Hörner (2015) call this kind of mechanism a chips, scrips, or token mechanism. Möbius (2001), Hauser and Hopenhayn (2008), and Abdulkadiroğlu and Bagwell (2013) consider dynamic favor exchange, and Guo and Hörner (2015) presents provision of goods to a consumer with stochastic valuations. The general finding of this literature is that token mechanisms, as proposed in Möbius (2001), do better than autarky but not as well as an optimal dynamic mechanism. In fact, token mechanisms are close to first-best if players are patient and there are many time periods. Results in Jackson and Sonnenschein (2007) imply that token mechanisms' inefficiency declines as square root of the number of periods (see Guo and Hörner, 2015). Thus, the theoretical literature suggests that point systems, while not exactly optimal, are simple and achieve a high level of efficiency.[34]

Another motivation for using a point mechanism is practicality. The simplicity of the mechanism and similarity to fiat money makes it promising. Similar mechanisms have been previously used in market design applications. For example, Prendergast (2017) describes how a similar mechanism was used to increase the efficiency of food distribution across food banks.

An important issue with applying point systems is that they require several "plumbing" decisions (Duflo, 2017). Should the matching algorithm impose a strict bound on negative balances? If so, what is the optimal minimum balance constraint? A tight constraint provides stronger incentives to hospitals but may reduce efficiency. Should points be credited when patients and donors are submitted, or should points be credited when transplants are conducted? How often should marginal products be recalculated as the composition of patients and donors in the platform changes? Recalculating them often is complex and reduces transparency, but recalculating infrequently can reduce efficiency.

7 Conclusion

Kidney exchange improves a patient's quality of life and extends life expectancy while reducing costs. We demonstrate that fragmentation in the US kidney exchange market results in an efficiency loss of between 25 to 55 percent of the about 800 kidney exchange transplants performed per year, implying a waste of hundreds of transplants per year.

The inefficiency arises due to two standard market failures. First, platforms use inefficient mechanisms that do not reward hospitals according to marginal products of their contributions. This induces hospitals to perform inefficient within hospital matches, even if hospitals solely maximize the welfare of their own patients. Second, there are agency problems that make hospitals too sensitive to the costs of participating in kidney exchange platforms.[35] Our analysis shows that both market failures are likely important and that platforms could use simple alternative mechanisms to substantially increase efficiency.

These findings have both short-term policy implications and broader implications for the design of kidney exchange markets. There are two clear short-term policy implications. First, there could be returns to existing platforms experimenting with point systems. Such systems can be implemented by individual platforms and it is will likely help them expand. Second,

[34] This message is consistent with the literature on monetary economics. Although optimal dynamic mechanisms can often improve on money (Kocherlakota, 1998), models in the tradition of Kiyotaki and Wright (1989) show that money can achieve high levels of efficiency even with simple institutions.

[35] This decomposition of market failure sources is consistent with long-standing concerns of surgeons, insurers, platforms, and researchers, and with recent policy changes. Roth et al. (2005) and Ashlagi and Roth (2014) recognized that hospitals may have incentives to match patients internally in static models. Surgeons and insurers have noted that it may be in the interest of insurers to subsidize exchanges, and have proposed that they do so (Rees et al., 2012).

third-party payers should consider subsidizing kidney exchange at platforms. We argued that hospitals are likely responsive to costs of participating in kidney exchange platforms, a behavior that leads to significant welfare losses. Subsidies by Medicare and private payers could alleviate this problem. Moreover, our analysis suggests that this two-pronged approach, which addresses the two market failures separately, is likely to be more robust and to have lower data requirements than mechanisms that address both market failures simultaneously.

Consistent with these results, there are initiatives moving in the direction of these policy changes. The NKR recently started experimenting with a points system through their "Center Liquidity Contribution Program." Some private insurers have started covering the costs of participating in kidney exchange platforms. Our results indicate that there could be large gains from continuing to move in this direction. Further, all platforms could use data-driven rewards system. Future research can contribute to the design and evaluation of these policies.

More broadly, our results raise the question of whether or not to use heavy-handed regulation, such as mandating participation in a single platform. For example, the U.K., Netherlands, and Canada (De Klerk et al., 2005; Johnson et al., 2008; Malik and Cole, 2014) mandate participation at a single national program. At a first glance, this approach seems reasonable because of the increasing returns to scale in kidney exchange. However, mandated single-platform participation can reduce competitive incentives for platforms that have arguably contributed to various innovations in kidney exchange.[36] Moreover, our estimates of returns to scale suggest that it would not be inefficient to have a few large platforms in the US because most of the potential efficiency gain comes from moving the market from individual hospitals to national platforms.

References

Abdulkadiroğlu, Atila and Kyle Bagwell, "Trust, Reciprocity, and Favors in Cooperative Relationships," *American Economic Journal: Microeconomics*, 2013, *5* (2), 213–259.

Abraham, David J, Avrim Blum, and Tuomas Sandholm, "Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges," in "Proceedings of the 8th ACM conference on Electronic commerce" ACM 2007, pp. 295–304.

Agarwal, Nikhil, Itai Ashlagi, Eduardo Azevedo, Clayton Featherstone, and Omer Karaduman, "What Matters for the Productivity of Kidney Exchange?," *Mimeo, MIT*, 2018.

American Society of Transplant Surgeons, "Kidney Paired Donation: Community Perspectives and Best Practices," 2016.

[36]Examples include the introduction of non-simultaneous chains (Rees et al., 2009), the developing Global Kidney Exchange program which allows pairs from development countries to overcome financial barriers (Rees et al., 2017), voucher programs to increase donation for future priority (Veale et al., 2017; Wall et al., 2017), and other operational innovations that reduce frictions and improve matching algorithms.

Anderson, Ross, Itai Ashlagi, David Gamarnik, and Yash Kanoria, "A dynamic model of barter exchange," in "Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms" SIAM 2014, pp. 1925–1933.

_ , _ , _ , Michael Rees, Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver, "Kidney Exchange and the Alliance for Paired Donation: Operations Research Changes the Way Kidneys Are Transplanted," *Interfaces*, 2015, *45* (1), 26–42.

Armstrong, Mark, "Competition in two-sided markets," *The RAND Journal of Economics*, 2006, *37* (3), 668–691.

Arrow, Kenneth, "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 1963, *53* (5), 941–973.

Ashlagi, Itai and Alvin E Roth, "Free riding and participation in large scale, multi-hospital kidney exchange," *Theoretical Economics*, 2014, *9* (3), 817–863.

_ , Maximillien Burq, Patrick Jaillet, and Vahideh Manshadi, "On matching and thickness in heterogeneous dynamic markets," *arXiv preprint arXiv:1606.03626*, 2016.

Becker, Gary S and Julio Jorge Elias, "Introducing incentives in the market for live and cadaveric organ donations," *The Journal of Economic Perspectives*, 2007, *21* (3), 3–24.

Bingaman, A W, F H Wright Jr, M Kapturczak, L Shen, S Vick, and C L Murphey, "Single-Center Kidney Paired Donation: The Methodist San Antonio Experience," *American Journal of Transplantation*, 2012, *12* (8), 2125–2132.

Boiteux, Marcel, "Sur la gestion des Monopoles Publics astreints a l'equilibre budgetaire," *Econometrica*, 1956, *24* (1), 22–40.

Chandra, Amitabh and Jonathan Skinner, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic literature*, 2012, *50* (3), 645–680.

Clemens, Jeffrey and Joshua D Gottlieb, "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?," *American Economic Review*, 2014, *104* (4).

Cowan, Nigel C., H A Gritsch, Nima Nassiri, Joseph Sinacore, and Jeffrey L Veale, "Broken Chains and Reneging: A Review of 1748 Kidney Paired Donation Transplants.," *American Journal of Transplantation*, 2017, *17 9*, 2451–2457.

Danovitch, Gabriel M, *Handbook of Kidney Transplantation*, Lippincott Williams & Wilkins, 2009.

Dickerson, John P, Ariel D Procaccia, and Tuomas Sandholm, "Dynamic Matching via Weighted Myopia with Application to Kidney Exchange," in "AAAI" 2012.

Duflo, Esther, "Richard T. Ely Lecture: The Economist as Plumber," *American Economic Review*, 2017, *107* (5), 1–26.

Ellison, Blake, "A Systematic Review of Kidney Paired Donation: Applying Lessons from Historic and Contemporary Case Studies to Improve the US Model," *Mimeo, University of Pennsylvania*, 2014.

Friedman, Eric J, Joseph Y Halpern, and Ian Kash, "Efficiency and Nash equilibria in a scrip system for P2P networks," in "Proceedings of the 7th ACM conference on Electronic commerce" ACM 2006, pp. 140–149.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani, *The Elements of Statistical Learning*, Vol. 1, Springer series in statistics New York, 2001.

Guo, Y and J Hörner, "Dynamic mechanisms with money," Technical Report, Mimeo, Yale University 2015.

Hauser, Christine and Hugo Hopenhayn, "Trading favors: Optimal exchange and forgiveness," *Collegio Carlo Alberto Carlo Alberto Notebooks*, 2008, *88.*

Held, Philip J, F McCormick, A Ojo, and John P Roberts, "A cost-benefit analysis of government compensation of kidney donors," *American Journal of Transplantation*, 2016, *16* (3), 877–885.

Irwin, F D, A F Bonagura, S W Crawford, and M Foote, "Kidney paired donation: a payer perspective," *American Journal of Transplantation*, 2012, *12* (6), 1388–1391.

Jackson, Matthew O and Hugo F Sonnenschein, "Overcoming incentive constraints by linking decisions," *Econometrica*, 2007, *75* (1), 241–257.

Johnson, Rachel J, Joanne E Allen, Susan V Fuggle, J Andrew Bradley, Chris Rudge et al., "Early experience of paired living kidney donation in the United Kingdom," *Transplantation*, 2008, *86* (12), 1672–1677.

Kiyotaki, Nobuhiro and Randall Wright, "On Money as a Medium of Exchange," *Journal of Political Economy*, 1989, *97* (4), 927–954.

Klerk, Marry De, Karin M Keizer, Frans HJ Claas, Marian Witvliet, Bernadette JJM Haase-Kromwijk, and Willem Weimar, "The Dutch national living donor kidney exchange program," *American Journal of Transplantation*, 2005, *5* (9), 2302–2305.

Kocherlakota, Narayana R, "Money is memory," *Journal of Economic Theory*, 1998, *81* (2), 232–251.

Kolstad, Jonathan T, "Information and quality when motivation is intrinsic: Evidence from surgeon report cards," *The American Economic Review*, 2013, *103* (7), 2875–2910.

Lerner, AP, "The Concept of Monopoly and the Measurement of Monopoly Power," *Review of Economic Studies*, 1934, *1* (3), 157–175.

Malik, Shafi and Edward Cole, "Foundations and principles of the Canadian living donor paired exchange program," *Canadian journal of kidney health and disease*, 2014, *1* (1), 6.

Marschak, Jacob and William H Andrews, "Random simultaneous equations and the theory of production," *Econometrica*, 1944, pp. 143–205.

Massie, Allan B, Sommer E Gentry, Robert A Montgomery, Adam A Bingaman, and Dorry L Segev, "Center-Level Utilization of Kidney Paired Donation," *American Journal of Transplantation*, 2013, *13* (5), 1317–1322.

Möbius, Markus, "Trading favors," *Mimeo, Microsoft Research*, 2001.

Morris, Carl N, "Parametric empirical Bayes inference: theory and applications," *Journal of the American Statistical Association*, 1983, *78* (381), 47–55.

National Kidney Registry, "Member Center Terms& Conditions," http://www.kidneyregistry.org/docs/NKR_MC_Terms_Conditions.pdf 2016.

Olley, G Steven and Ariel Pakes, "The dynamics of productivity in the telecommunications equipment industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Orandi, BJ, JM Garonzik-Wang, Allan B Massie, Andrea A Zachary, JR Montgomery, KJ Van Arendonk, Mark D Stegall, SC Jordan, J Oberholzer, TB Dunn et al., "Quantifying the risk of incompatible kidney transplantation: a multicenter study," *American Journal of Transplantation*, 2014, *14* (7), 1573–1580.

Prendergast, Canice, "The Allocation of Food to Food Banks," *University of Chicago, mimeo*, 2017.

Ramsey, Frank P, "A Contribution to the Theory of Taxation," *The Economic Journal*, 1927, *37* (145), 47–61.

Rees, Michael A, Jonathan E Kopke, Ronald P Pelletier, Dorry L Segev, Matthew E Rutter, Alfredo J Fabrega, Jeffrey Rogers, Oleh G Pankewycz, Janet Hiller, Alvin E Roth, and others, "A nonsimultaneous, extended, altruistic-donor chain," *New England Journal of Medicine*, 2009, *360* (11), 1096–1101.

_ **, Mark A Schnitzler, EY Zavala, James A Cutler, Alvin E Roth, Frank D Irwin, Stephen W Crawford, and Alan B Leichtman**, "Call to develop a standard acquisition charge model for kidney paired donation," *American Journal of Transplantation*, 2012, *12* (6), 1392–1397.

_ **, Ty B Dunn, Christian S Kuhr, Christopher L Marsh, Jeffrey Rogers, Susan E Rees, Alejandra Cicero, Laurie J Reece, Alvin E Roth, Obi Ekwenna et al.**, "Kidney exchange to overcome financial barriers to kidney transplantation," *American Journal of Transplantation*, 2017, *17* (3), 782–790.

Robert, Christian and George Casella, "Monte Carlo Statistical Methods Springer-Verlag," *New York*, 2004.

Rochet, Jean-Charles and Jean Tirole, "Platform competition in two-sided markets," *Journal of the European Economic Association*, 2003, *1* (4), 990–1029.

Roth, Alvin E, "Repugnance as a Constraint on Markets," *The Journal of Economic Perspectives*, 2007, *21* (3), 37–58.

_ , Tayfun Sönmez, and M Utku Ünver, "Kidney exchange," *The Quarterly Journal of Economics*, 2004, *119* (2), 457–488.

_ , _ , and _ , "Transplant Center Incentives in Kidney Exchange," *Mimeo, Boston College*, 2005.

_ , _ , and _ , "Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences," *The American Economic Review*, 2007, pp. 828–851.

Sönmez, Tayfun and M Utku Ünver, "Market design for kidney exchange," *The Handbook of Market Design*, 2013, pp. 93–137.

Sönmez, Tayfun and M Utku Ünver, "Market design for kidney exchange," *The Handbook of Market Design*, 2013, p. 93.

USRDS, United States Renal Data System, *2013 USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States*, Vol. 2, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2013.

_ , *2016 USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States*, Vol. 2, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2016.

Veale, Jeffrey L, Alexander M Capron, Nima Nassiri, Gabriel Danovitch, H Albin Gritsch, Amy Waterman, Joseph Del Pizzo, Jim C Hu, Marek Pycia, Suzanne McGuire et al., "Vouchers for Future Kidney Transplants to Overcome'Chronological Incompatibility'Between Living Donors and Recipients.," *Transplantation*, 2017.

Wall, Anji E, Jeffrey L Veale, and Marc L Melcher, "Advanced Donation Programs and Deceased Donor Initiated Chains-2 Innovations in Kidney Paired Donation.," *Transplantation*, 2017.

Weyl, E Glen, "A Price Theory of Multi-Sided Platforms," *The American Economic Review*, 2010, *100* (4), 1642–1672.

Wolfe, Robert A, Valarie B Ashby, Edgar L Milford, Akinlolu O Ojo, Robert E Ettenger, Lawrence YC Agodoa, Philip J Held, and Friedrich K Port, "Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant," *New England Journal of Medicine*, 1999, *341* (23), 1725–1730.