Market Failure in Kidney Exchange*†

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

April 20, 2018

Abstract

Kidney exchanges facilitate over 800 transplants per year in the United States. We show that, despite this success, kidney exchange markets suffer from market failures that cause the loss of hundreds of transplants per year but could be addressed with simple policies. Our argument has three parts. First, we document that the market is highly fragmented, with approximately 60% of kidney exchange transplants being arranged by small platforms, often within hospitals, as opposed to in large, national platforms. Moreover, we find direct evidence that small platforms often arrange inefficient exchanges. Second, we propose a simple model to show that inefficiency arises for two reasons: hospitals do not internalize their patients' benefits from participation and current mechanisms do not give hospitals adequate incentives. Third, we estimate a platform's production function to quantify the inefficiency and design practical mechanisms. Our estimates show that hospitals' production scales are too small to match patients efficiently. Eliminating this inefficiency requires a combined approach that uses new mechanisms and incentivizes participation.

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, 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

Kidney exchange markets enable approximately 800 transplants per year. Patients in this market need a kidney transplant, and most come with a living, but incompatible, donor. Transplants are organized using patient-donor swaps and chains initiated by altruistic donors. The importance of this market has grown because of a shortage of organs from deceased donors and because current law prohibits using monetary compensation to acquire organs.[1] The economic value of a transplant is estimated at more than a million dollars: kidney transplants save lives, reduce healthcare costs, and save taxpayer funds.[2]

This paper evaluates the kidney exchange market and shows that, despite significant success in organizing transplants, the market suffers from two serious and fixable market failures that together result in hundreds of lost transplants per year. Indeed, our descriptive evidence suggests that the market is fragmented and operates inefficiently. We use price-theoretic arguments to explain that this inefficiency arises from two standard market failures. Each failure corresponds to a specific problem with current institutions. Both problems yield suboptimal incentives for hospitals, the key decision makers (Roth et al., 2005; Ashlagi and Roth, 2014; Rees et al., 2009). First, large national kidney exchange platforms use inefficient mechanisms to match patients. Due to biological compatibility, there is substantial variation in the additional transplants that result from different patient and donor types being submitted to a national platform. But current mechanisms ignore these differences in social value. This limitation in the design pushes hospitals to match patients with high social value outside the platform, even when it is inefficient. Second, there is scope for agency problems: hospitals face most of the costs of participating in national platforms but receive only a fraction of the benefits. We show that fixing these two problems can result in hundreds more transplants per year.

This argument has three parts. First, we use administrative datasets and institutional detail to diagnose and document the causes and consequences of these market failures. Second, we argue that, although kidney exchange markets look different from other markets, their inefficiency is caused by traditional market failures. To analyze the market, it is useful to think about the role of a kidney exchange platform as a producer of transplants that takes patients and donors supplied by hospitals as inputs. This re-framing, based on neoclassical producer theory, allows us to use classic price theory arguments that have clear policy

[1] 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 (NOTA) 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).

[2] 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 (ESRD). The ESRD 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).

implications. Third, we estimate the key primitive of the model, a transplant production function, using data from the largest kidney exchange platform. The primitive allows us to estimate the magnitude of the inefficiencies, design practical alternative mechanisms, and suggest appropriate policy responses.

In the first part, we document key facts about the market to motivate our model. We merge administrative data on the universe of transplants in the United States and proprietary data from the three largest kidney exchange platforms. The data reveal several signs of inefficiency. Rather than most transactions taking place at a few large platforms, the market is highly fragmented. During our sample period, 62% of kidney exchange transplants involved patients and donors from the same hospital. Crucially, we find direct evidence that many within-hospital exchanges are inefficient from a social perspective. An efficient market would use easy-to-match donors only for hard-to-match patients. We show that small within-hospital exchanges often violate this simple rule, but large national platforms do not.

We also find evidence consistent with inefficient hospital incentives causing these problems. Many hospitals do not participate in national platforms. Consistent with barriers due to fixed costs of participating in kidney exchange, smaller hospitals are less likely to participate. Previous surveys and a study of the payment structure in this market suggests that participation costs are a significant fraction of hospital revenue from kidney exchange. However, these costs are small relative to the social benefit 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 a competing small kidney exchange platform.

These facts motivate our model in the second part. The key insight underlying our approach is that although kidney exchange markets do not directly use monetary incentives to acquire organs, it is useful to analyze them using neoclassical producer theory. The 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 re-framing is based on three key institutional features. First, hospitals are the key decision-makers. They submit patients and donors to a platform voluntarily and are not forced to participate. Instead of submitting a patient-donor pair to a large platform, a hospital can perform a transplant with other pairs that it is treating. Thus, kidney exchange platforms must reward hospitals with transplants to procure necessary inputs (patients and donors). Second, due to biological compatibility constraints, some types of patients and donors enable more transplants than others, i.e. some inputs are more productive than others. For example, blood type O donors can donate to patients of any blood type. This fact makes them both scarce and valuable when organizing kidney exchanges.[3] Because hospitals can perform exchanges within their set of patients and donors, a platform may have to increase the rewards for submitting valuable types in order to procure them. Third, transplants

[3] In a simple model, Roth et al. (2007) show that certain "over-demanded" types enable two additional transplants at a platform while others add no value to the pool. Our calculations will generalize this idea by computing the additional number of transplants generated by a rich classification of types using data from the largest exchange.

are a natural numeraire good for rewarding hospitals. A platform must ration transplants amongst the types of patients and donors that are, due biological compatibility constraints, in abundance.[4] A transplant can therefore be transferred to a hospital by selecting its patient instead of another hospital's. This allows transplants to act as a numeraire good, even though direct monetary transfers for acquiring organs are prohibited.

The key primitive of the model is the transplant production function. Its shape determines important features of the economics of kidney exchange. The returns to scale determine how efficient it is to match patients in large exchanges. The marginal products of different types of patients and donors determine whether some submissions produce more transplants than others. For our empirical analysis in the third-part, we consider a steady-state interpretation of the model where all units are measured in flows.

Our main result, Theorem 1, formally shows that market failure can result from two sources of inefficiency. The first source is based on an inefficient mechanism. Most kidney exchange platforms run optimization algorithms that select which patients are matched. When a hospital submits a patient to the exchange, the hospital is rewarded according to the *probability* with which the patient is matched. In contrast, Theorem 1 shows that, to maximize hospital welfare, hospitals should be rewarded based on the *marginal product* of their submissions. Because existing platforms do not reward hospitals for the social value of their submissions, even a hospital that maximizes the number of its own patients that are transplanted has to perform socially inefficient matches. Hospitals must therefore decide between helping their patients or performing socially efficient matches. This problem can be ameliorated using point mechanisms that reward hospitals according to marginal products.[5]

The second source of inefficiency are due to externalities if hospital welfare differs from social welfare. These externalities arise if hospitals do not maximize the number of their own patients that are transplanted. It creates inefficiency, for example, if a hospital performs too many internal matches to avoid the administrative costs of kidney exchange, even though these matches happen at an inefficiently small scale. We argue that these administrative costs are small relative to the value of a transplant to a patient and his or her health insurer. But, these costs are a significant fraction of hospital revenue from kidney exchange. Therefore, these externalities are likely due to agency problems created by the reimbursement system for hospitals.

Theorem 1 shows that there is no inefficiency if neither of these two sources of market failure is present. This decomposition of market failure sources is consistent with long-standing concerns of surgeons, insurers, platforms, and researchers, and with recent policy changes.[6]

[4] A patient with blood type O can only receive a kidney from a blood type O donor. This feature creates an abundance of O patients with donors of other blood types. Similarly, patients without a related living donor are also in abundance. Many patients and donors in such submissions will inevitably remain unmatched. See Roth et al. (2007) for arguments based on a limit economy.

[5] Because our simplified steady-state model is not dynamic, it does not pin down the game form of an optimal dynamic mechanism. Existing theoretical work shows that optimal dynamic mechanisms in related settings are complicated. However, simple token mechanisms that keep track of point balances are both simple and highly efficient, making these mechanisms good candidates for practical implementation. We discuss these theoretical and practical issues in Section 6.

[6] Roth et al. (2005) and Ashlagi and Roth (2014) recognized that hospitals may have incentives to match

Our analysis pinpoints the sources of market failure to specific wedges, a result that points to practical policy implications.

The third part of our argument uses the data to quantify these issues' importance and to design responses. We estimate the platform's transplant production function using administrative data from the largest kidney exchange platform, the National Kidney Registry (NKR), and detailed information about the matching algorithms and operational procedures used in the platform. This information allows us to build an empirical model of a kidney exchange platform.[7] The model fits the data well and allows us to estimate the flow of transplants that would be produced given a flow of submissions.

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.[8] At the same time, almost all single-hospital hospital platforms are far below the efficient scale. This difference in productivity suggests considerable inefficiency due to fragmentation. We estimate that the gains from moving all the production to the efficient scale is at least 200, and likely closer to 400 transplants. Thus, consistent with the descriptive evidence and the shape of the production function, fragmentation has a large efficiency cost.

Second, we assess the current mechanism and characterize rewards in an optimal mechanism.[9] Theorem 1 shows that optimal mechanisms reward submissions approximately according to marginal products, while current mechanisms reward submissions according to probabilities of matching. Motivated by this result, we calculate the marginal products and probabilities of matching for each type of submission using the estimated production function. We find that marginal products are considerably different from the probabilities of matching, which implies that existing mechanisms are far from optimal. We then develop point mechanisms that give approximately optimal incentives to hospitals.

Third, we study the importance of the two sources of market failure. In our price-theoretic approach, 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,

---

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).

[7] Standard methods of estimating a production function are not suited to our empirical setting because they rely on low-dimensional production functions, typically Cobb-Douglas with a few types of inputs, and data on observed input and output data for many firms (Marschak and Andrews, 1944; Olley and Pakes, 1996).

[8] Because returns to scale are roughly constant at a large enough scale, gains due to scale economies alone from merging the major kidney exchange platforms are likely small. This observation on the scale economies can also explain why, instead of tipping to a single platform, multiple large national exchanges co-exist (Ellison and Fudenberg, 2003).

[9] This is in the spirit of the sufficient statistics approach in the public finance literature (Dixit and Sandmo, 1977; Saez, 2001; Chetty, 2009). In particular, our analysis suffers from the standard caveat that marginal products depend on the composition of submissions, which is endogenous.

the deadweight loss is significant but considerably lower than the inefficiency due to market fragmentation. Therefore, except in extreme cases, both the current mechanism and agency problems cause significant inefficiency in the market. Taken together, these results motivate a combined approach that improves the mechanism design and implements policies that encourage hospital participation in the national platforms.

After presenting our main argument, we discuss important limitations of our analysis and some extensions. First, we discuss unsolved theoretical and practical issues involved in using our results in the real world. Second, we consider kidney exchange mechanisms that attempt to solve both market failures simultaneously. Third, we consider the effect of imperfect competition between oligopolistic platforms on efficiency. Fourth, we discuss recent changes in kidney exchange markets that attempt to address the two market failures that we describe.

Our paper is organized as follows. Section 2 describes institutional details and data sources. The three parts of our argument are developed in Sections 3, 4, and 5, respectively. Section 6 discusses the limitations and extensions, and Section 7 concludes.

2 Background and Data

2.1 Basics of Kidney Exchange

This subsection describes the basics of kidney transplantation and kidney exchange and can be skipped by readers who are familiar with this literature.

End-Stage Renal Disease (ESRD) afflicts more than half a million Americans. The disease is almost universally covered by Medicare, including 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 several hundred thousand dollars 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 suitably screened living donors are small. Taken together, these facts indicate that living donor kidney transplants have large economic value. Held et al. (2016) places the economic value of an average kidney transplant at $1.1 million.[10]

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

[10] Held et al. (2016) conduct a comprehensive cost-benefit analysis with sensitivity to a range of assumptions. They value one year of perfect health at $200,000 and adjust for differences in quality of life between dialysis and a transplant. The costs include differences in expected medical costs incurred over the lifetime of a dialyzed patient and a transplanted patient, discounted at the rate of 3% per year. The cost savings on dialysis alone are 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).

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.[11] 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. This form of transplantation 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. Finally, there are some **unpaired patients**, who do not have a willing live donor.

The platform organizes transplants in two ways. The first, called a **cycle**, involves a set of pairs. The kidney from one of the 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) indicate 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 in our data.

There are two types of biological compatibility constraints on kidney transplants: **ABO 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 either A or B if it has one of these antigens, AB with both, and O with neither. A donor is tissue-type incompatible with a patient if the donor has certain cell-surface proteins (antigens) to which the patient has an immune response.[12] Patient sensitization levels differ considerably. The most common measure of sensitization is the **Panel Reactive Antibody (PRA)** score. A patient's PRA

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

[12] 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. Each donor has up to 6 possible human leukocyte antigen (HLA) proteins 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

is between 0 and 100 and denotes the percentage of a representative population of donors with whom a patient is tissue-type incompatible.

2.2 Key Institutional Features and the Economics of Kidney Exchange

There are three key institutional features that drive outcomes and will factor in our analysis.

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.[13] Besides these major platforms, there are small regional platforms and individual hospitals that also organize kidney exchanges. As we will see in the next section, most hospitals that participate in large national platforms also match patients outside those platforms.

Moreover, hospitals are not forced to participate in platforms. Platforms effectively reward hospitals with transplants in order to receive submissions (Roth et al. 2005; Ashlagi and Roth 2014) 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 exposition, assume that there are only two blood types, O and A. These two types together are the vast 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. For this simplified case, Roth et al. (2007) showed that the number $f(\boldsymbol{q})$ of transplants that can be performed 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

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 patient-donor has become possible due to development of desensitization technologies (Orandi et al., 2014).

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

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 such variation in social value of submissions, thereby inducing hospitals to perform socially inefficient internal matches. Consider a hospital with two over-demanded pairs. The hospital could match perform a pairwise exchange to conduct two transplants. However, if the hospital submits the pairs to the platform, the hospital receives twice the probability that a pair is matched. According to our data, the hospital expects only 1.6 transplants from submitting these two pairs to the platform because the platform uses each of these specific pairs in an exchange with probability 0.8. This expectation pushes the hospital to match its patients outside the platform. If the hospital submits the pairs to the platform, the two pairs together generate twice their marginal product. The Roth et al. (2007) model suggests the platform could generate four additional transplants using these 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 to help its patients. Our theoretical analysis will summarize all the ways in which such inefficiencies can happen, and clarify how platform rules can be redesigned to eliminate this problem.

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 by choosing which under-demanded pairs to match, without compromising efficiency.

Third, hospitals do not necessarily maximize a utilitarian measure of patient welfare, causing externalities. We refer to any such divergence as an agency problem between hospitals and patients because they primarily arise from a reimbursement system that does not incentivize efficient actions from the patients' perspective. Hospitals incur most of the costs of kidney exchange but receive only a fraction of the social benefits. While the social value from one transplant is more than $1,000,000, hospital revenues are between $100,000 to $160,000 per transplant.[14] Variable profits are likely much smaller. 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.[15] 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, anecdotal evidence indicates

[14] See Held et al. (2016); USRDS, United States Renal Data System (2013). This amount includes payments for surgery teams, drugs, equipment, and capital.

[15] 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.

considerable heterogeneity regarding hospital sophistication: some hospitals report using 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 National Kidney Registry (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.[16] 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 the same information as the 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 very high degree of certainty.[17]

[16] 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.

[17] In fact, 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.

In order to estimate a platform's transplant production function, we will require more than just information on transplants organized by a platform (the output) – we also require information about which patients and donors are submitted to the platform (the inputs). The OPTN dataset is not well suited for this task because it does not identify a patients' related incompatible donor, if any. Therefore, the data allow us to determine neither the pairs that participated in kidney exchange nor the patients and donors that remain unmatched.

To address this limitation, 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. Therefore, we can 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 use detailed administrative dataset to describe three key facts: the kidney exchange market is highly fragmented; this fragmentation leads to inefficiency; and there is evidence of broadly defined agency problems between hospitals and patients.

3.1 Fragmentation

We document market fragmentation by showing that a large fraction of transplants are facilitated by individual hospitals instead of large national platforms. A market coordinated on a single platform would have all transplants facilitated through that platform. In contrast, a completely fragmented market would have no transplants facilitated through a multi-hospital platform. In this case, all transplants would be **within hospital**, meaning each transplanted patient and donor belong to the same hospital.

Therefore, to measure fragmentation, we first identify within hospital transplants by checking whether the hospital associated with a patient transplanted through a kidney exchange is the same as the hospital that recovered the organ.[18] From this set, we exclude all transplants facilitated by one of three large exchanges: the NKR, the APD and UNOS.

[18]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 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.

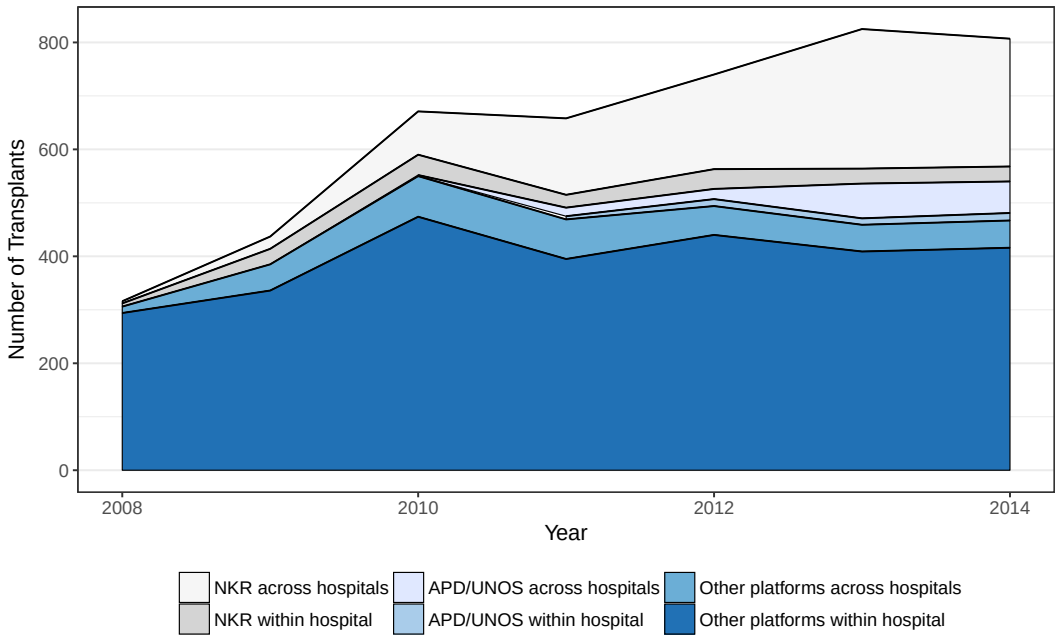


Figure 1: Market fragmentation and trends in kidney exchange

Notes: The figure displays the number of kidney exchange transplants in different categories. NKR, APD/UNOS and Other classifies a transplant into the platform that facilitated the exchange, including single-hospital platforms. Within hospital and across hospital classify a transplant into whether the donor hospital was the same as the patient hospital.

Figure 1 shows that the market is 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, the vast majority of the transplants facilitated by multi-hospital platforms occur **across hospitals**, involving a living donor and a patient from different hospitals. This contrast between the overall market and the platforms is striking because the platforms to 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 co-ordinating across hospitals has potential gains. We will formally analyze and quantify these gains in the subsequent sections.

Figure 1 also shows that the total number of kidney exchange transplants grew from about 400 transplants in 2008 to about 800 in 2014.[19] However, overall market growth seems to have slowed in recent years. The total number remains at around 800,[20] well below some

[19] Our data for the NKR extend until December 4, 2014. This censoring may account for the slight drop in transplants in the last year of this figure.

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

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.[21] 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 points to inefficiency if the returns to scale in matching patients and donors are high. However, it is possible, at least in principle, that the market is highly fragmented but close to efficient. For example, it may be that hospitals operate at an approximately efficient scale, so that the loss due to market fragmentation is small. We now present direct evidence that this is not the case by showing that hospitals often conduct 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.[22] 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 this patient if it is the only way to get this 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.

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

[22] 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.

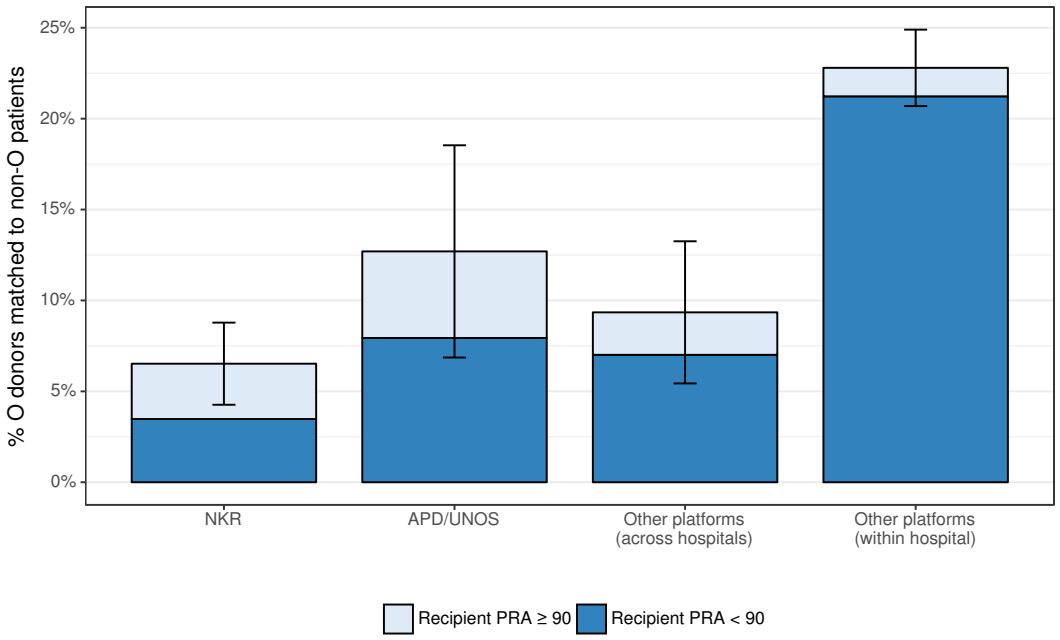


Figure 2: Smoking-gun evidence of hospitals performing inefficient matches

Notes: The bars display the percentage of transplanted O donors whose kidneys were transplanted into non-O patients for different categories of transplants. NKR, APD/UNOS and Other classifies a transplant into the platform that facilitated the exchange, including 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 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 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. 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

We have shown that the kidney exchange market is highly fragmented, with many hospitals performing matches outside the three multi-hospital platforms. Moreover, this fragmentation leads to real efficiency losses because internal matches performed by hospitals are often socially inefficient. These results lead us to ask why hospitals do not participate more in national platforms. To address this issue, we first document key facts about hospital behavior and then discuss different hypotheses that can explain it. Our results indicate 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). The results are qualitatively similar if we consider participation at any of the three largest kidney exchange platforms.

We start by describing the relationship between hospital size and the extensive and intensive margins of participation in the NKR. We measure hospital size as the number of transplants conducted per year, including deceased donor and direct living donor kidney transplants.[23] Figure 3 depicts the extensive margin of participation amongst 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. 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 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

[23] This broad measure of size limits the endogenous effect of participation in the uiNKR 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.

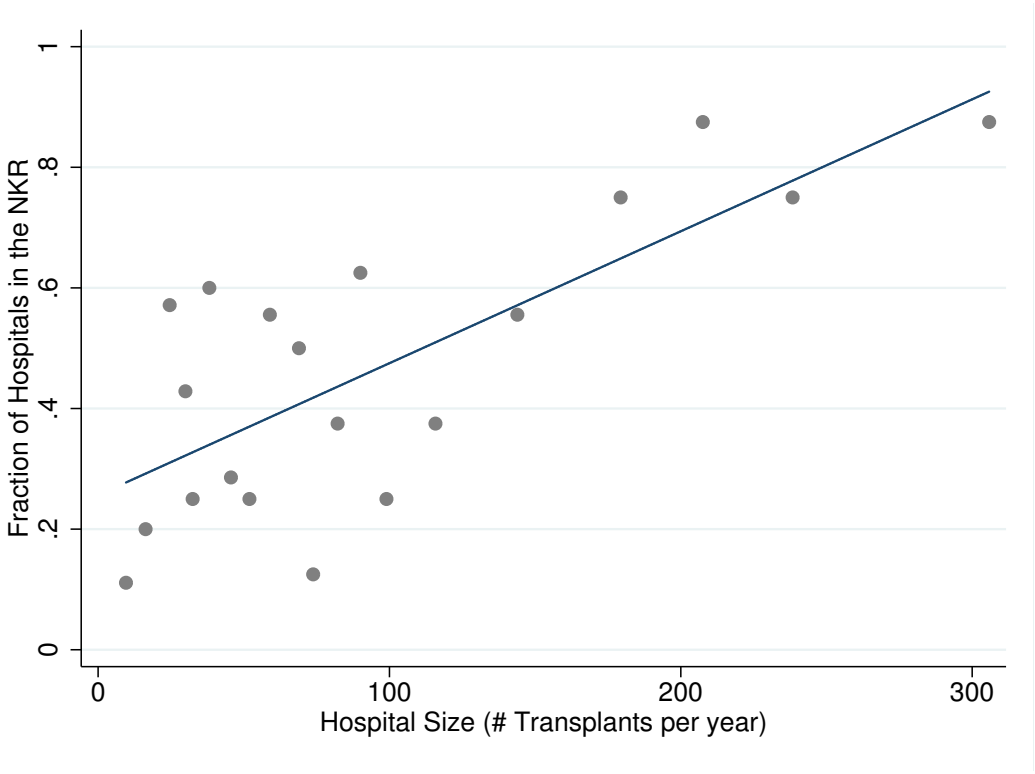


Figure 3: Heterogeneity in participation in the NKR

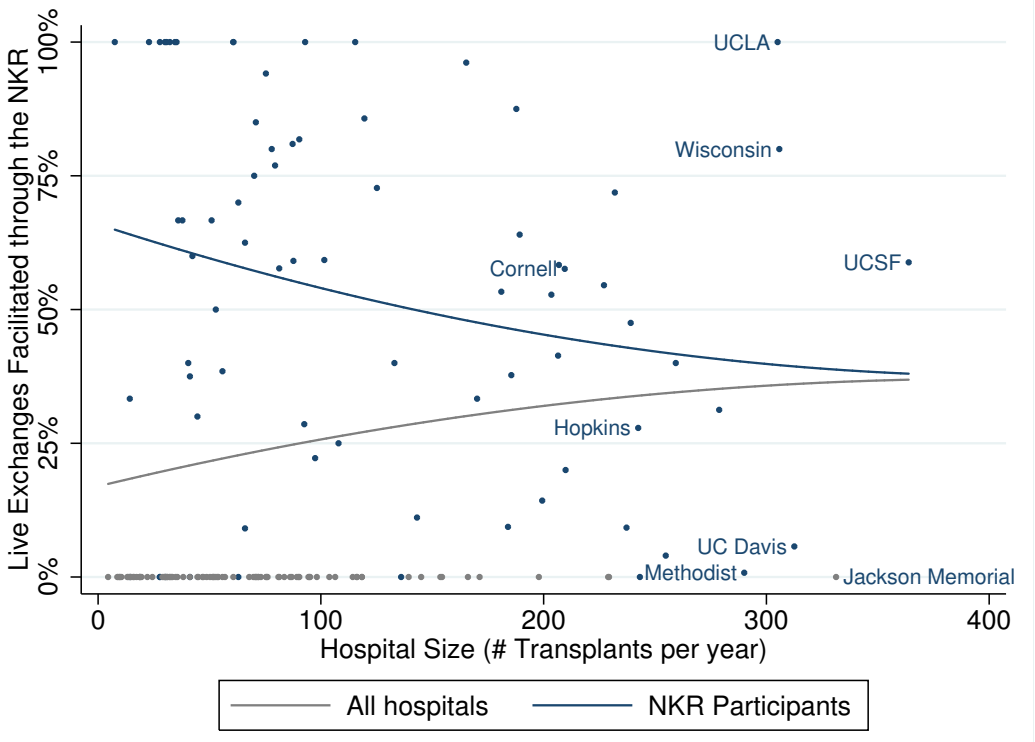


Figure 4: Reliance on the NKR for live-donor exchanges

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 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, amongst 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 transplanted through the NKR blood type O, and 41% of donors were 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 fact supports our equal treatment of all transplants for welfare calculations irrespective of whether they are facilitated though a national platform.

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%	–

Note: A pair is overdemanded if the patient is blood-type compatible with the related donor. Underdemanded pairs either are O-patients without O-donors or are AB-donors without AB-patients. Sample of all patients and donors registered in the NKR between April 4, 2012 and December 1, 2014.

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.

3.3.2 Implications for Hospital Behavior

The facts detailed above have important 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 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). Their main prediction is 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. For example, conditional on participating, larger hospitals perform fewer transplants through the NKR because these hospitals have more opportunities to match patients outside the platform (Figure 4). However, the hypothesis does not explain many other important aspects of the data. For example, many hospitals do not participate in a national platform at all, even though all hospitals likely have patients who cannot be matched.[24] 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. These 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, for example, 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. Its behavior fits neither the prediction of a pure profit maximization motive nor the motive to purely maximize the welfare of the hospital's own patients. 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 both with the anecdotal evidence on kidney exchange reviewed in Section 2 and with typical findings about the behavior of healthcare providers. For example, Clemens and Gottlieb (2014) investigate how providers respond to changes in Medicare reimbursement rates, and Kolstad (2013) investigates how surgeons respond to the introduction of publicly observable report cards. These studies find

[24] 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.

that providers respond to incentives but are not purely motivated by profits and also take patient welfare into account.

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. Under most matching technologies, submitted patients will be harder to match than those not submitted. Unfortunately, we cannot directly test this prediction because we do not have data on the entire pool of patients available to individual hospitals. Nonetheless, as we discussed above, the data show that submitted pairs are much harder to match than the general population. Moreover, the NKR transplants patients and donors that are harder to match. If the NKR and hospitals are equally capable of matching hard-to-match patients, then the data suggest that hospitals select hard-to-match types for submission to the NKR. However, some selection may be driven by the NKR being significantly better at finding kidney exchanges for hard-to-match patients with a given pool. It is reassuring that the results on selection do not falsify the two theories that best fit the participation behavior.

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, there is no simple model that describes the behavior of all hospitals. These two findings represent a theoretical and empirical challenge because we cannot rely on a simple stylized model of hospital behavior to either propose practical market design solutions or perform welfare analyses. But, these results do indicate that agency problems are a potentially important driver of inefficiency and should therefore not just be assumed away. Both implications are important elements in the analysis we develop below.

4 Theory

The evidence above shows that kidney exchange markets are fragmented and that this fragmentation leads to real efficiency losses. We now develop a model to explain how inefficiency arises. The model also shows what statistics must be estimated to measure the inefficiency and to design better mechanisms.

The basic idea is that the kidney exchange market is fundamentally similar to a traditional market. Namely, a kidney exchange clearinghouse is a platform that procures submissions (donors and patients) from hospitals and rewards these hospitals with transplants. The platform's ability to produce transplants from submissions is given by a production function, which describes how many transplants can be performed with a given pool of submissions.

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** indexed by $i = 1, \ldots, I$. A vector of **quantities** $\boldsymbol{q} = (q_i)_{i=1,\ldots,I}$ in $\mathbb{R}_+$ specifies a quantity q_i of each submission type available to the platform. 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 summarizes what matches are possible. Roth et al. (2007) calculated the production function using a simple model that we described in Section 2. Roth et al. (2007) assumed that all submissions are pairs and that only blood type compatibility matters, so their model has $I = 16$ types. Our analysis applies both to simple, theoretically tractable, production functions, such as the one in Roth et al. (2007), and to sophisticated 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.

The platform produces transplants using submissions provided by **hospitals** indexed by $h = 1, \ldots, H$. 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** $\boldsymbol{p} = (p_i)_{i=1,\ldots,I}$ 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.

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 the hospital (Sonmez and Unver, 2013; 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 quasilinear in the number of transplants they receive from the platform and the private cost of their submissions. Given a vector of rewards $\boldsymbol{p}$, hospital h maximizes

$$\boldsymbol{p} \cdot \boldsymbol{q}^h - C^h(\boldsymbol{q}^h),$$

where $C^h(\boldsymbol{q}^h)$ is the **private cost** of submissions in transplant units. One particular case is when hospitals maximize the number of their own patients that are transplanted, in which case $C^h(\boldsymbol{q}^h)$ is the number of within-hospital transplants that the hospital would have to forfeit in order to submit $\boldsymbol{q}^h$. If all transplants that are performed must be allocated to some hospital, we have that $f(\boldsymbol{q}) = \sum_{h=1}^{H} \boldsymbol{p} \cdot \boldsymbol{q}^h$.

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,

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

where $\boldsymbol{q}$ is the aggregate quantity. Hospital welfare measures efficiency according to hospitals' objectives. 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 total 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 **externality** from hospital h's submissions given by

$$E^h(\boldsymbol{q}^h) = C^h(\boldsymbol{q}^h) - SC^h(\boldsymbol{q}^h).$$

For example, E^h is positive 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.

Given an allocation, let $E(\boldsymbol{q}^1,\dots,\boldsymbol{q}^H)$ be the **aggregate externality**, which equals the sum of externalities from all hospitals. The externality represents the benefits to stakeholders other than hospitals. In the particular case where there are no agency problems, we have $E \equiv 0$.

Define **total welfare** to be

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

Thus, the difference between social and hospital welfare is equal to the aggregate externality E. Define the **aggregate cost function** $C(\boldsymbol{q})$ as the minimum of the sum of costs of all hospitals subject to the total quantity submitted being equal to $\boldsymbol{q}$. For simplicity, we assume that the production cost, social cost, and aggregate cost functions are defined over all non-negative real vectors and are smooth. Aggregate costs are strictly convex. 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. Define first-best

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

This basic model finesses important issues, such as efficiency costs of transferring transplants, the choice of a particular welfare function, and the case of multiple competing platforms. We will return to these issues in Section 6.

4.2 Illustrative example

The model clarifies that kidney exchange is, in many ways, similar to a traditional market. Regulating a kidney exchange platform is similar to the Ramsey (1927)-Boiteux (1956) problem of regulating a multi-product firm. The key difference is that, instead of a firm producing many products, a kidney exchange platform procures multiple submission types from hospitals.[25] This connection requires some abstraction because the relevant numeraire are transplants. In this section, we consider a concrete example to clarify this connection. The example is only for illustration and specific assumptions are not used in the subsequent analysis.

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 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.

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 (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 utility equals the number of transplants minus the monetary costs divided by the value per transplant; i.e., hospital h's cost function is

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

Society values each transplant at V dollars. This social value includes the value of transplants to insurers, taxpayers, patients, and other patients who benefit from reductions in the deceased donor waitlist. Held et al. (2016) estimate that health care savings alone from the average transplant are around \$300,000 and the gains in quality and length of life are valued at over \$1,000,000.[26] Thus, it is reasonable to estimate that V is in the ballpark of \$1,300,000. Taken together, these estimates show a substantial wedge between the social and private values of a transplant because V is more than an order of magnitude larger than

[25] 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).

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

v. Moreover, the additional social costs of participating in a platform $K^h(\boldsymbol{q}^h)$ are negligible compared to the benefits of performing additional transplants. This fits our model with social costs

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

The externality term equals

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

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

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

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

This 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. For example, if k is \$10,000 and v is \$50,000, then the wedge per submission 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. Therefore, the wedge is large in this example, and precisely because hospitals receive a small share of transplants' social benefits but bear many of the costs. Therefore, costs that are negligible from a social perspective are magnified when divided by v.

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 $\boldsymbol{q}_0$. For simplicity, the figure assumes that the aggregate social cost is a smooth function, $\widetilde{SC}(\boldsymbol{q})$, of the aggregate quantity, as long as each hospital chooses an optimal quantity given a vector of rewards.[27]

The figure shows that the current allocation is inefficient from both the hospital and social perspectives. The hospital-optimal quantity is $\boldsymbol{q}^*$ because it equates $\boldsymbol{\nabla f}$ with marginal private costs. 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 private cost and social cost differ, so that $E \neq 0$. The socially efficient quantity, $\boldsymbol{q}^{**}$, can be achieved by using

[27] Sufficient conditions are that each hospital's optimal quantity is a smooth function of rewards, and the function mapping rewards into the optimal aggregate quantity by all hospitals is a bijection of $[0, \infty)^I$ into itself.

efficient incentives, $\boldsymbol{p} = \boldsymbol{\nabla f}$, and solving agency problems so that $E \equiv 0$. In this example, agency problems can be solved by reimbursing hospitals for the costs of kidney exchange through the platform, i.e. $K^h\left(q^h\right)$.

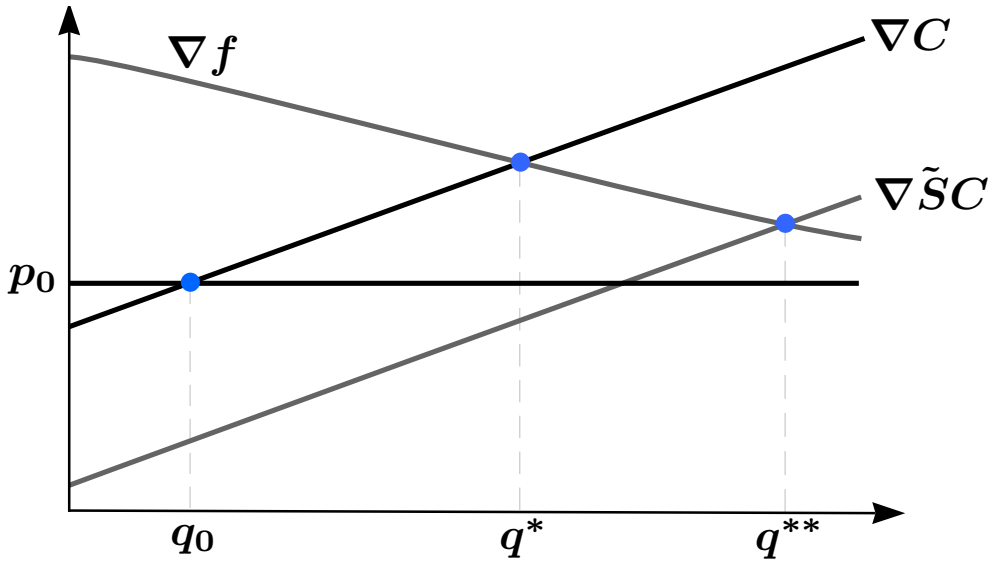


Figure 5: The Two Sources of Market Failure

Notes: The horizontal axis represents aggregate quantity of submissions into the kidney exchange platform. The curves represent the marginal product of submissions $\boldsymbol{\nabla f}(\boldsymbol{q})$, the marginal private cost of submissions $\boldsymbol{\nabla C}(\boldsymbol{q})$ from hospital's perspective, and the marginal social cost of submissions $\boldsymbol{\nabla \tilde{SC}}(\boldsymbol{q})$. Both axes represent I-dimensional vectors. The figure depicts the current quantity $\boldsymbol{q}_0$, with agency problems and a suboptimal mechanism, the quantity $\boldsymbol{q}^*$ from a hospital-optimal mechanism but with agency problems, and the first-best quantity $\boldsymbol{q}^{**}$ with an efficient mechanism, and no agency problems. For simplicity, the figure assumes that aggregate social cost is a uniquely defined function of rewards and that the platform has constant returns to scale.

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. Achieving the first-best requires setting rewards equal to the marginal products plus the externalities. But even with constant returns to scale, there are only enough transplants to implement reward hospitals that equal the marginal product, making it impossible to reward hospitals to correct for externalities. 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 can now describe optimal reward vectors, which can guide the design of optimal mechanisms. The following theorem collects the main insights.

Theorem 1 (Optimal Rewards)**.** *Consider a vector of rewards* $\boldsymbol{p}$ *and allocation* $(\boldsymbol{q}^h)_{h=1,\dots,H}$ *with strictly positive aggregate quantity* $\boldsymbol{q}$ *that maximizes hospital welfare subject to all hospitals choosing supply optimally given* $\boldsymbol{p}$ *and subject allocating the same number of transplants that are produced,* $f(\boldsymbol{q}) = \boldsymbol{p} \cdot \boldsymbol{q}$*. Then:*

1. *The platform rewards each type of submission with its marginal product minus an adjustment term,*

$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{A},$$

   where

$$\boldsymbol{A} = \frac{\boldsymbol{\nabla f} \cdot \boldsymbol{q} - f}{\boldsymbol{q}' \cdot \boldsymbol{D^2 C} \cdot \boldsymbol{q}} \cdot \boldsymbol{q}' \boldsymbol{D^2 C}.$$

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

3. *If, in addition, there are no agency problems, in the sense that* $\boldsymbol{\partial E} = \boldsymbol{0}$*, then this allocation also maximizes total welfare.*

The theorem characterizes rewards in a mechanism that maximize 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 for this result 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 cannot allocate more transplants than it produces. This constraint is binding if f exhibits weakly increasing returns to scale because the number of transplants produced, $f(\boldsymbol{q})$, is less than the sum of marginal products of submissions, $\boldsymbol{\nabla f}(\boldsymbol{q}) \cdot \boldsymbol{q}$. It is easy to visualize this if $\boldsymbol{q}$ is one-dimensional because the tangents to the production function cross the vertical axis below zero. Therefore, a platform has to shade its rewards to hospitals relative to the marginal products. The optimal level of shading for each type of submission is given by the adjustment term $\boldsymbol{A}$. The adjustment term says that the platform should shade more aggressively on submissions with less elastic supply. To see this, note that marginal cost $\boldsymbol{\nabla C}$ equals the inverse supply curve, under some regularity conditions, so that $\boldsymbol{D^2 C}$ is the derivative of the inverse supply curve. In

fact, our formula is identical to the Ramsey (1927)-Boiteux (1956) formula for an optimal linear commodity tax. For example, if the cross-elasticities of supply are zero, we obtain an inverse-elasticity rule for the optimal shading, as in Ramsey's work and in the Lerner index from optimal monopoly pricing.

The theorem suggests 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 because they do not depend on supply elasticities.

Moreover, there is no need to consider nonlinear rewards because we can achieve first-best welfare by rewarding hospitals linearly with their submissions. 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 (Hauser and Hopenhayn, 2008). We will 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 in our model: 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. $E \neq 0$. The market functions efficiently if platform incentives are optimal ($\boldsymbol{p} = \boldsymbol{\nabla f}$) and there are no agency problems ($E \equiv 0$).

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 $\boldsymbol{q}^{**}$. 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).[28] We use these estimates to measure the magnitude of the inefficiencies, and develop simple policy responses. The key results come from three analyses. First, we measure total inefficiency by estimating both the returns to scale and how many more transplants would be performed if production was moved to the efficient scale. Second, we design optimal mechanisms using marginal products. Finally, we use our price-theoretic framework to measure the gains from implementing optimal incentives through the mechanism alone.

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, these methods are best suited for low-dimensional production functions that only depend on a few inputs, such as capital and labor, but 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 are able 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

[28] A similar exercise can be conducted using data from the APD or UNOS. There are two challenges that must be overcome. First, these platforms together have relatively small number of transplanted patients and donors creating difficulties in accurately estimating the parameters of the model. Second, differences in the data collection protocols and logistics across these platforms would require us to write separate code for these two platforms. We therefore focus our efforts on data from the NKR.

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 any set of inputs $\boldsymbol{q}$ and parameters θ. There are four key events 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 chain lengths and transplantation probabilities for various patient and donors. 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.

   We estimate the vector of current submission rates, $\boldsymbol{q}_0$, by resampling from the administrative NKR data. Specifically, we estimate the total submissions per day as an empirical average and then draw specific patient/donor/pair characteristics and minimum acceptance criteria from the dataset with replacement. The time-average of these simulation draws yields $\hat{q}_{0,i}$ for any finite partition of individual patients, donors, or pairs into types, $i = 1, \ldots, I$. An identical resampling process allows us to simulate arrivals for any alternative rate of submissions $\boldsymbol{q} \neq \boldsymbol{q}_0$.

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.[29]

[29]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

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 transplants 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:** Proposed transplants are 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.[30] 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 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, chain lengths, and pool size observed in our data. We match these statistics by the following submission types: altruistic donor,

this last constraint and iteratively add the constraints to prohibit large cycles. Appendix C provides further details on the algorithm.

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

patient-donor pairs, and unpaired patients.[31] These quantities are precisely the rewards that the NKR implicitly sets for hospitals, $\boldsymbol{p}_0$, and are therefore key components of our economic model.

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 the transplant rates and chain lengths observed in the dataset. 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 a transplant.

   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.[32] 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

[31] 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.

[32] 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.

run. The dependence on the initial pool eventually fades away. Let $y_{jk}^{*,t} = 1$ if a kidney from donor k is transplanted to patient j. We compute the time average of the total number of transplants to estimate f for each simulation s:

$$\hat{f}_s(\hat{q}) = \frac{1}{T}\sum_{t=1}^{T} \left|y_s^{*,t}\right|,$$

where T is the total number of days simulated and $|y_s^{*,t}|$ is the total number of transplants in period t in simulation s. In what follows, we report estimates based on an average of 100 simulations.

5.2 Scale, Inefficiency and Optimal Rewards

This section uses the estimated production function to measure the amount of inefficiency, evaluate the importance of the two sources of market failure, and design better mechanisms. We use the total number of kidney exchange transplants (including in both platforms and hospitals) as our measure of social welfare, SW. For simplicity, we begin by reporting the results under our main specifications, and then we discuss detailed robustness analyses.

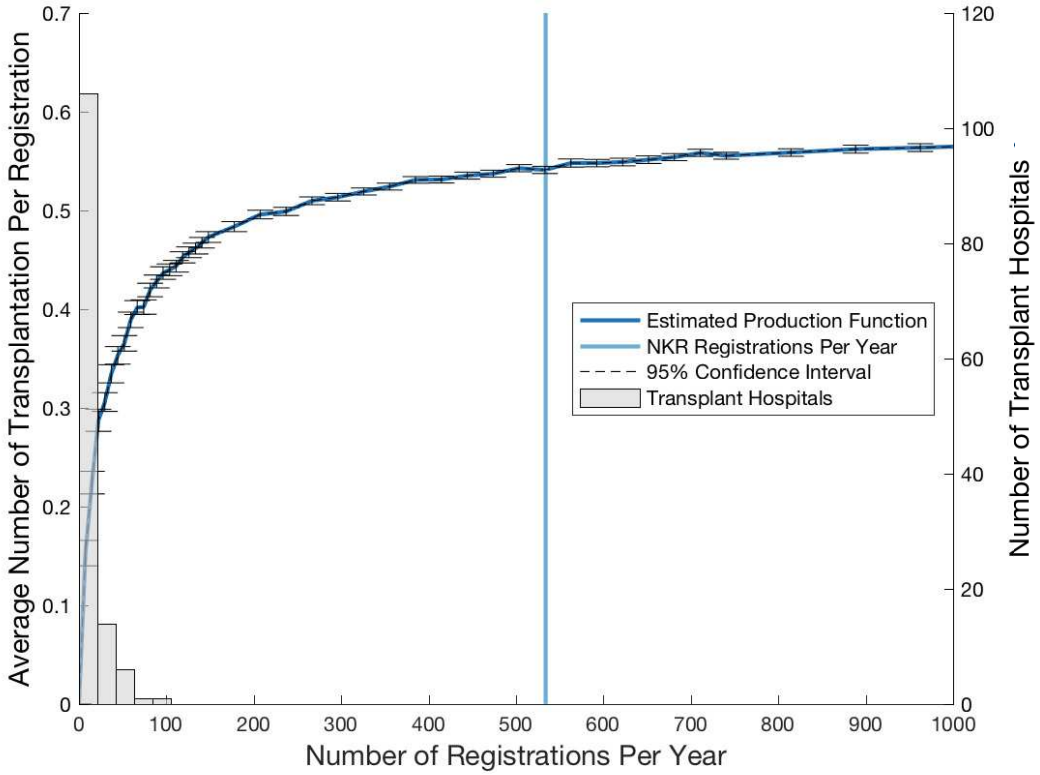
5.2.1 Returns to scale

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 altruists and pairs per year. Figure 6 depicts average products, equal to $f(\boldsymbol{q})/\|\boldsymbol{q}\|_1$, as a function of the total flow, $\|\boldsymbol{q}\|_1$.

The estimates show a remarkable pattern. Although the returns to scale always increase, they reach a plateau fairly quickly. With a scale of 534 donor arrivals per year, the NKR is well within the region of approximately constant returns to scale. It has an average product of 0.54, which varies marginally 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, which collectively account for the majority of kidney exchange transplants, operate at an efficient scale. 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

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 plot uses the baseline parameters and the pool composition from the NKR.

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. The scale of hospital h is then $q^h = \hat{f}^{-1}\left(y^h\right)$ where $\hat{f}$ is the estimated production function.

This exercise suggests that individual hospitals as well as UNOS and APD 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. The 90th percentile is 26. The largest, Methodist Hospital in San Antonio, has a scale of 104. The average product at these efficient scales is 0.16, 0.30 and 0.44, respectively. Thus, at our estimated production function, even the largest single-hospital platforms does not operate at an efficient scale. UNOS and APD have estimated average products of 0.41 and 0.42 respectively. 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, $q^h = \hat{f}^{-1}\left(y^h\right)$, 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 that are conducted through kidney exchange each 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 participation rate, and hospitals in the bottom quartile.[33] Because hospitals in the top quartile of participation are conducting a larger fraction of kidney exchanges through the NKR, they should be submitting a less selected pool of patients and donors. 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 the hospital. Comparing estimates suggests that overall inefficiency is not particularly sensitive to these compositional 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.[34] However, this extremely conservative calculation is likely slack for two reasons. First, even within the non-NKR hospitals that

[33] 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.

[34] 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 respective platform.

participate in either UNOS or APD, two-thirds of kidney exchange transplants are performed within the hospitals (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. Within 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.

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.[35] On the other hand, single-hospital programs face simpler logistical constraints, which may

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

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.[36]

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 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, are 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 sensitive to potential compositional differences in the kidney exchange pool. Table D5 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.

5.2.3 Marginal products and inefficiency of current mechanisms

We now test whether current platform rules are efficient. Theorem 1 shows that, if the rules are optimal, then rewards for different submission types, $\boldsymbol{p}_0$, should approximate marginal products, $\boldsymbol{p}^* = \boldsymbol{\nabla f} - \boldsymbol{A}$. We will test this equality at the composition and rate of submissions, $\boldsymbol{q_0}$, for the NKR during our sample period.

[36] 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.

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 $\boldsymbol{\nabla f}$ can be estimated by numerically differentiating the production function.

In principle, calculating the adjustment term requires estimates of the supply elasticity matrix. It is difficult to estimate this matrix using observed submissions to the NKR because it is high-dimensional. 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

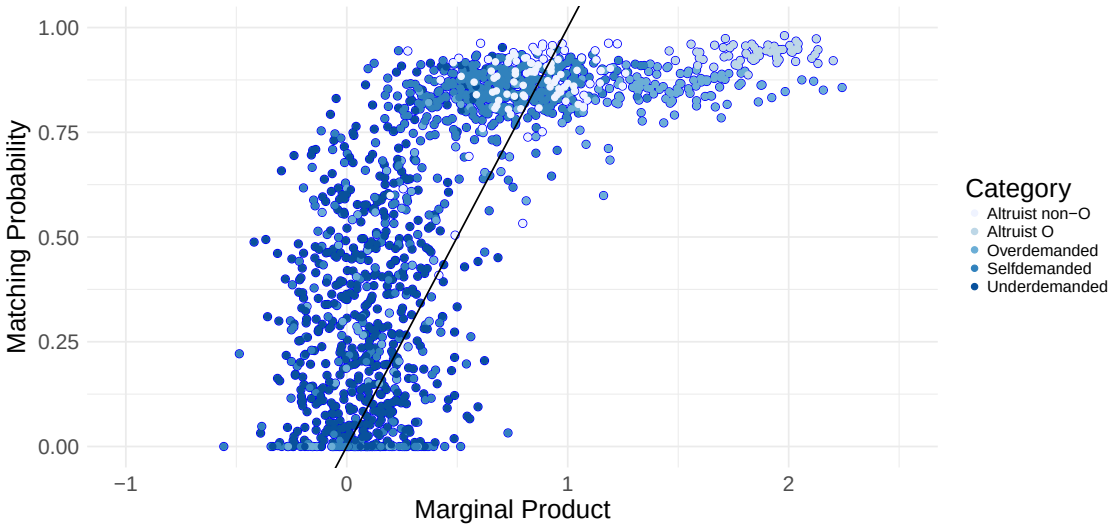
$$\boldsymbol{A} \cdot \frac{\boldsymbol{q}}{\|\boldsymbol{q}\|_1} = \boldsymbol{\nabla f} \cdot \frac{\boldsymbol{q}}{\|\boldsymbol{q}\|_1} - \frac{f}{\|\boldsymbol{q}\|_1}.$$

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 derivates for each of the 1429 patient/donor types yields an average shading of only 2.16×10^{-4}. In what follows, we simply approximate optimal rewards with marginal products because shading is not a major concern.

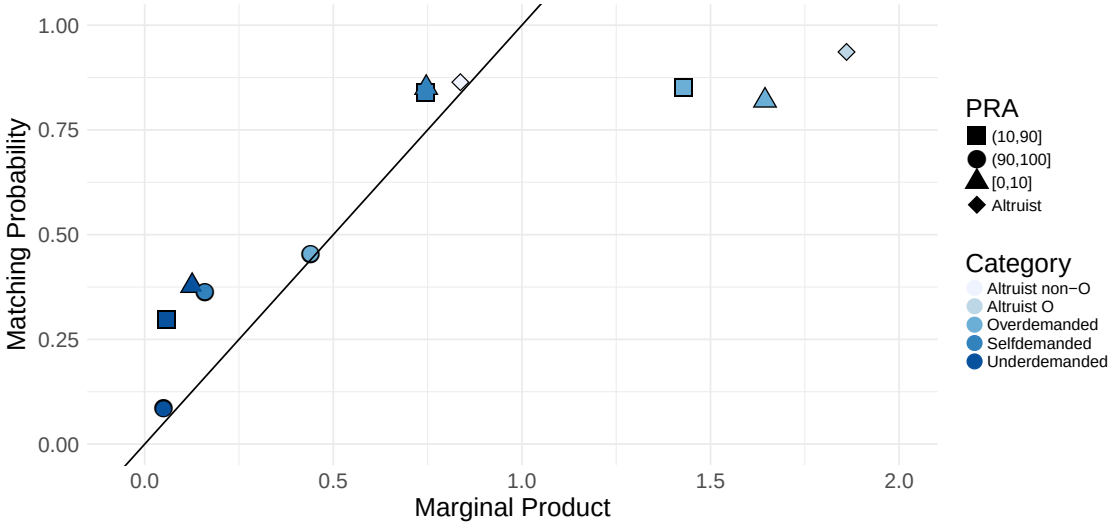
Figure 7a plots current rewards (the probabilities of matching $\boldsymbol{p}_0$) versus optimal rewards (marginal products $\boldsymbol{\nabla f}$). Some of the 1429 types have negative estimated marginal products because the matching algorithms are myopic, which can result in crowding out of future 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 underdemanded, overdemanded and selfdemanded 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. Consistent with our estimates, Roth et al. (2007) deliver a marginal product of 0 for underdemanded pairs, but our estimates differ for some other types. For example, the model in Roth et al. (2007) delivers a marginal product of 2 for overdemanded pairs. We estimate that an overdemanded pair with low sensitization has a marginal product of 1.64. A reason for this difference is that these pairs only are matched with probability 0.82 according to our data. 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-degree line. Altruistic donors and overdemanded pairs with low PRA are far below this line. Overdemanded 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



(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 correspond to a submission in the data. Probabilities of matching 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 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).

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. Underdemanded pairs with low sensitization have marginal products of approximately 0.13 but have a probability of being matched of around 0.38, which is substantial. 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.

5.2.4 Optimal point mechanisms

The marginal products and the small adjustment term $\boldsymbol{A}$ together suggest that current platform rules are far from optimal. We now show that marginal products are highly predictable using a small number of patient and donor categories, and use these categories to suggest a practical point mechanism that is close to optimal.

We use a regression tree to construct categories that best predict marginal products. The approach allows for a data-driven categorization of patient and donor types. 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 with category mean marginal products $\boldsymbol{\nabla f}$ and probabilities of matching $\boldsymbol{p}_0$ that 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 C4).[37] Points could be awarded when the hospital submits a patient and/or donor to the NKR, or at the time of transplantation. Our theory suggests that points, if awarded upon submission, should be only slightly shaded relative to marginal products $\boldsymbol{\nabla f}$. 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 motives awarding points at 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.

The platform could impose a lower bound on the point balance a hospital must have before one of its submissions is transplanted. A hospital that has hit this bound would only be able to transplant either an altruistic donor or a overdemanded pair that does not have a highly sensitized patient (Non-O Patient, O Donor, PRA$<$96%). We postpone a more detailed discussion of these implementation details about the point system to Section 6.

[37] Evidence from the psychology literature suggests that professionals are likely to use data-driven advice when it is presented with simple decision trees (Gigerenzer and Goldstein, 1996; Gigerenzer and Kurzenhaeuser, 2005).

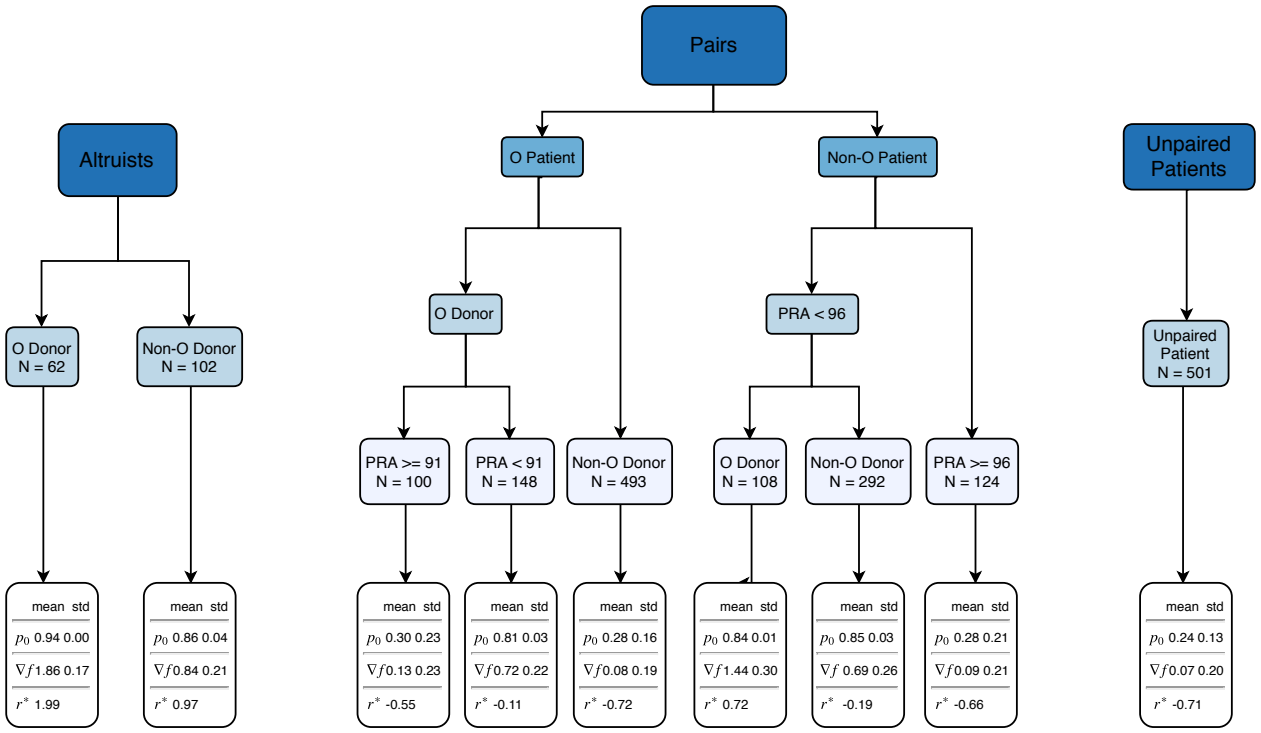


Figure 8: Regression Tree for Marginal Products

Notes: Categories are determined by regression tree analysis to predict marginal products as a function of whether a submission is a pair or altruist, 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).

5.2.5 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.

Calculating deadweight loss is similar to a multi-dimensional version of calculating deadweight loss from linear commodity taxation. Figure 9 illustrates 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 curve $\boldsymbol{\nabla C}$ between $\boldsymbol{q}_0$ and $\boldsymbol{q}^*$. Therefore, this deadweight loss is the integral of $\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{\nabla C}(\boldsymbol{q})$ 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.

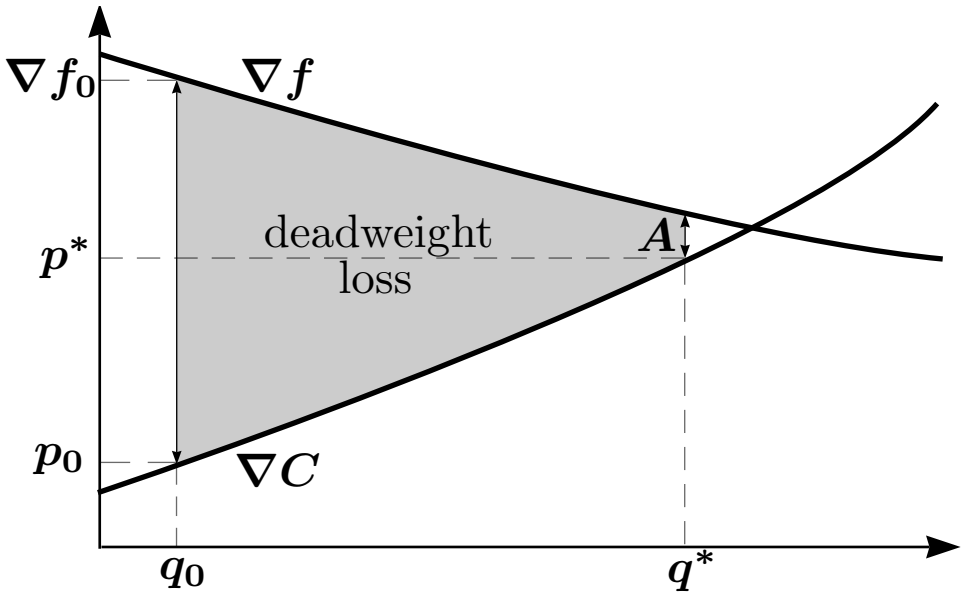


Figure 9: Hospital-Welfare Deadweight Loss from the Current Mechanism

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 $\boldsymbol{\nabla f}_0$ can be calculated from the production function. In contrast, the supply curve $\boldsymbol{\nabla C}$ and optimal rewards $\boldsymbol{p}^*$ and quantities $\boldsymbol{q}^*$ are not observed, and depend on the elasticity of supply.

The following proposition shows that a similar approximation to the hospital deadweight loss holds in our setting.

Proposition 1. *Consider a strictly positive aggregate supply of pairs, $\boldsymbol{q}_0$, that results when hospitals choose supply optimally given rewards $\boldsymbol{p}_0$. Consider aggregate supply $\boldsymbol{q}^*$ and rewards $\boldsymbol{p}^*$ that maximize hospital welfare as in Theorem 1. Assume that the matrix $\boldsymbol{D}^2\boldsymbol{C}(\boldsymbol{q}^*) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}^*)$ is non-singular. Then the deadweight loss in hospital welfare at $\boldsymbol{q}_0$ is approximated by*

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

Alternatively, the deadweight loss is approximated by

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

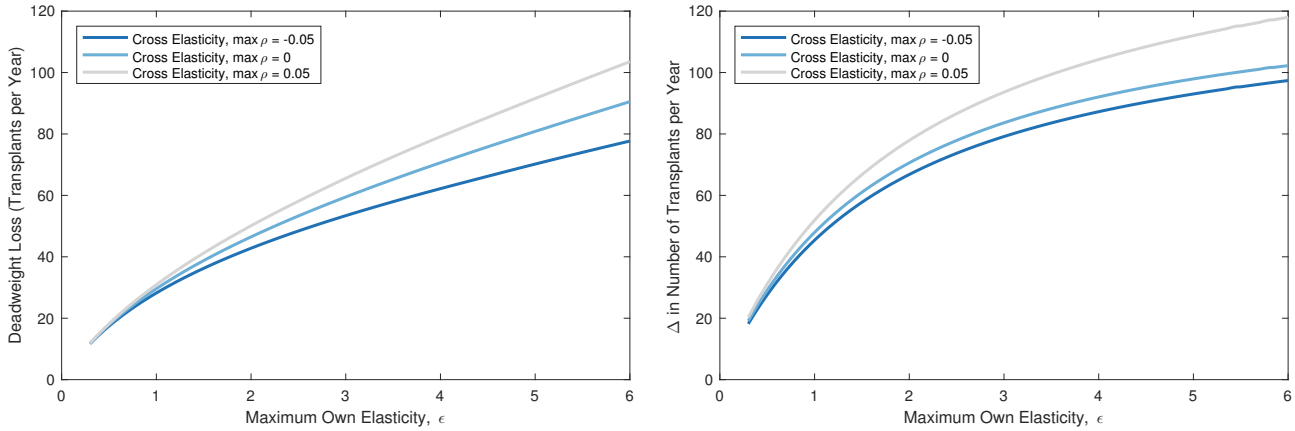
The error in both approximations is bounded by an $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$ term plus a constant multiple of $|\boldsymbol{A} \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0)|$.

These formulas are a multidimensional version of the standard approximation for the Harberger triangle 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 derivative of inverse supply minus the derivative of inverse demand. The second formula shows that the deadweight loss is one half of a quadratic expression in the wedge $\boldsymbol{\nabla}\boldsymbol{f}_0 - \boldsymbol{p}_0$. The term $\boldsymbol{D}^2\boldsymbol{C}(\boldsymbol{q}_0)$ 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 $\boldsymbol{\nabla}\boldsymbol{f}_0 - \boldsymbol{p}_0$ and either $\boldsymbol{q}^* - \boldsymbol{q}_0$ or $\boldsymbol{D}^2\boldsymbol{C}(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)$. We can estimate $\boldsymbol{\nabla}\boldsymbol{f}_0$, $\boldsymbol{p}_0$, 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 $\boldsymbol{D}^2\boldsymbol{C}(\boldsymbol{q}_0)$. Nevertheless, the large wedge between the current private and social incentives suggests the deadweight loss is significant unless the supply elasticity is extremely small.

We can use equation (3) 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). We estimate the wedge $\boldsymbol{\nabla}\boldsymbol{f}_0 - \boldsymbol{p}_0$ and the curvature matrix $\boldsymbol{D}^2\boldsymbol{f}$ for these categories using our production function. To use equation (3), we need to specify supply elasticities through the matrix $\boldsymbol{D}^2\boldsymbol{C}(\boldsymbol{q}_0)$. 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 can repeat this exercise for varying assumptions on cross-elasticities.[38]



(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.

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 increases 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

[38] Specifically, we solved the problem

$$\max_{D^2C(q_0)} \frac{1}{2}(\nabla \boldsymbol{f}_0 - \boldsymbol{p}_0)[\boldsymbol{D}^2\boldsymbol{C}(\boldsymbol{q}_0) - \boldsymbol{D}^2\boldsymbol{f}(\boldsymbol{q}_0)]^{-1}(\nabla \boldsymbol{f}_0 - \boldsymbol{p}_0)'$$

$$\text{s.t.} \left(\frac{\partial^2 C}{\partial^2 q_j}\right)^{-1} \frac{p_{0,j}}{q_{0,j}} \leq \varepsilon \text{ and } \left(\frac{\partial^2 C}{\partial^2 q_j}\right)^{-1} \frac{p_{0,j}}{q_{0,j}} \geq 0$$

$$\left(\frac{\partial^2 C}{\partial q_j \partial q_k}\right)^{-1} \frac{p_{0,k}}{q_{0,j}} = \rho \frac{\left(\frac{\partial^2 C}{\partial^2 q_j}\right)^{-1} \frac{p_{0,j}}{q_{0,j}} + \left(\frac{\partial^2 C}{\partial^2 q_k}\right)^{-1} \frac{p_{0,k}}{q_{0,k}}}{2}.$$

for each value of the bound on elasticities, ε.

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 ∇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.

6 Discussion and Robustness

6.1 Importance of agency problems and inefficient platform incentives

In order to better guide policy, 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 actual deadweight loss is potentially much larger as most specifications yield numbers is about 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.

Taken together, these results imply that agency problems are important unless elasticities are extremely high. Under the hypothesis that there are no agency problems, hospital welfare equals total 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, still below our lower bound result

of 200 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 from the quadratic Taylor approximation, so that the bias is high if $\nabla \boldsymbol{f}$ is extremely convex. Thus, attributing all the deadweight loss to inefficient platform incentives requires that elasticities are high, $\nabla \boldsymbol{f}$ 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.

6.2 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, this model does not specify an extensive form game and abstracts away from dynamic considerations. 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 5.2.4. 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 by Möbius, 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.[39]

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.

[39]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.

An important issue with applying point systems is that they require several "plumbing" decisions. 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 pairs are submitted, or should points be credited when pairs are transplanted? 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.

6.3 Maximizing social welfare

Theorem 1 describes mechanisms that maximize hospital welfare. A natural alternative would be to use mechanisms that maximize social welfare. Rewards in these mechanisms are described in the following proposition.

Proposition 2 (Optimal Rewards for Maximizing Total Welfare)**.** *Consider a vector of rewards* $\boldsymbol{p}$ *and strictly positive aggregate quantity vector* $\boldsymbol{q}$ *that maximize social welfare subject to all hospitals choosing supply optimally given* $\boldsymbol{p}$ *and subject to allocating the same number of transplants that are produced. Assume that the production function has constant returns to scale and that private costs are strictly convex. Define the aggregate externality as a function of aggregate quantity* $\tilde{\boldsymbol{E}}$ *as in equation (A12), and assume that it is smooth at the aggregate quantity* $\boldsymbol{q}$*. Then:*

1. *The platform rewards each type of submission with its marginal product plus an adjustment term.*

$$\boldsymbol{p} = \boldsymbol{\nabla f(q)} + \boldsymbol{A^{SW}},$$

where

$$\boldsymbol{A^{SW}} = \frac{1}{1+\lambda^{SW}} \boldsymbol{\nabla \tilde{E}(q)} - \frac{\lambda^{SW}}{1+\lambda^{SW}} \boldsymbol{q' D^2 C(q)}.$$

and

$$\lambda^{SW} = \frac{\boldsymbol{\nabla \tilde{E}(q) \cdot q}}{\boldsymbol{q' D^2 C(q) q}}.$$

2. *The adjustment term* $\boldsymbol{A^{SW}}$ *can be non-zero even with constant returns to scale.*

3. *The optimal rewards attain first-best social welfare if and only if the average externality at the optimum,* $\boldsymbol{\nabla \tilde{E}(q) \cdot q}$*, is zero.*

Part 1 shows the optimal mechanism rewards submissions by their marginal products plus an adjustment. The adjustment equals an externality term, which is greater for submissions that generate more externalities, minus a shading term that depends on elasticities. In the first-best, hospitals are rewarded for their marginal contributions to the platform as well as any externalities. However, if there are not enough transplants to pay for the externalities,

the planner has to shade rewards. As in optimal linear commodity taxation, it is 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 the externalities generated by each type of submission. 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 externality 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.

6.4 Competing platforms

Two natural policy responses to the fragmentation and increasing returns to scale 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. To address these issues, consider a platform that faces an inverse supply of submissions $\boldsymbol{P_S}(\boldsymbol{q})$. Assume that the platform maximizes the number of transplants $f(\boldsymbol{q})$ that it facilitates. The following proposition describes the optimal rewards.

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

1. *The platform rewards each type of submission with its marginal product, plus an adjustment term,*

$$\boldsymbol{p} = \boldsymbol{\nabla f}(\boldsymbol{q}) + \boldsymbol{A^C},$$

where

$$\boldsymbol{A^C} = \frac{\boldsymbol{q'DP_S}(\boldsymbol{q})\boldsymbol{q}}{f(\boldsymbol{q})}\boldsymbol{\nabla f}(\boldsymbol{q}) - \boldsymbol{q'DP_S}(\boldsymbol{q}).$$

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

3. *If supply is perfectly elastic, so that the matrix* $\boldsymbol{DP_S}$ *is zero, then rewards equal marginal products, as in the optimal rewards 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 are very productive. To understand the intuition for this, consider the case where supply has zero cross elasticities and own-elasticities denoted ϵ_i. Then the optimal rewards formula simplifies to an analogue of the Lerner index formula:

$$\frac{\partial_i f - p_i}{p_i} = \frac{1}{\epsilon_i}\left(\frac{f - \partial_i f \cdot q_i}{f}\right).$$

This formula describes the optimal mark-down in rewards relative to marginal products. If there is only one type of submission, then the right hand side is trivially equal to zero because f exhibits constant returns to scale. In this case, the optimal rewards equal the marginal product. When there are multiple submission types, then the quantity weighted average of $\partial_i f - p_i$ across submissions must be zero because the platform cannot promise rewards that exceed its product. 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 supply is very elastic, i.e. if ϵ_i is close to infinity or, more generally, $\boldsymbol{DP_S}$ is close to zero.

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 kidney exchange 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 problem induces hospitals to perform inefficient matches despite seeking to help their patients. Second, there are agency problems that make hospitals too sensitive to the costs of participating in kidney exchange platforms. 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, existing platforms should experiment with point systems. This recommendation is particularly actionable because it can be implemented by single platforms and it is will likely

help them expand. Second, third-party payers should subsidize 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 can likely alleviate this problem. Moreover, our analysis suggests that this two-pronged approach, which addresses the two market failures separately, is more robust and has lower data requirements than mechanisms that address both market failures simultaneously.

Consistent with our 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 can be large gains from continuing to move in this direction. Further, all platforms can 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.[40] Moreover, our estimates of returns to scale suggest that it would not be inefficient to have a few large exchange in the US because most of the potential efficiency gain comes from moving the market from individual hospitals to national exchanges.

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.

[40] Examples include the introduction of non-simultaneous chains (Rees et al., 2009), the development Global Kidney Exchange 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.

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.

Chetty, Raj, "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 2009, *1* (1), 451–488.

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 : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons*, 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.

Dixit, Avinash and Agnar Sandmo, "Some simplified formulae for optimal income taxation," *The Scandinavian Journal of Economics*, 1977, pp. 417–423.

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.

Ellison, Glenn and Drew Fudenberg, "Knife-Edge or Plateau: When Do Market Models Tip?," *The Quarterly Journal of Economics*, 2003, *118* (4), 1249–1278.

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.

Gigerenzer, Gerd and Daniel G Goldstein, "Reasoning the fast and frugal way: models of bounded rationality.," *Psychological review*, 1996, *103* (4), 650.

_ **and Stephanie Kurzenhaeuser**, "Fast and frugal heuristics in medical decision making," *Science and medicine in dialogue: Thinking through particulars and universals*, 2005, pp. 3–15.

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.

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**, "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.

Saez, Emmanuel, "Using elasticities to derive optimal income tax rates," *The review of economic studies*, 2001, *68* (1), 205–229.

Sonmez, Tayfun and M Utku Unver, "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.

Appendix to "Market Failure in Kidney Exchange"

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

April 20, 2018

A Proofs

A.1 Proof of the Main Theorem

Let $\mathbb{R}_+$ be the set of non-negative real numbers agind $\mathbb{R}_{++}$ be the set of strictly positive real numbers. Let the supply correspondence of hospital h be

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

That is $\boldsymbol{S}^h(\boldsymbol{p})$ is the set of quantities that are optimal for hospital h given the vector of rewards $\boldsymbol{p}$. 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 begin with a lemma showing that hospital supply can be aggregated, so that hospitals behave as a single hospital that takes the aggregate cost curve into account. This is similar to standard aggregation results in neoclassical firm theory (Mas-Colell et al., 1995 p. 148).

Lemma 1. *Fix a vector of rewards $\boldsymbol{p}$. Aggregate supply $\boldsymbol{S}(\boldsymbol{p})$ is the Minkowski sum of individual supply $\boldsymbol{S}^h(\boldsymbol{p})$ for all hospitals. Moreover, if there is a set of individual supply vectors $(\boldsymbol{q}^h)_{h=1}^H$ with each $\boldsymbol{q}^h \in \boldsymbol{S}^h(\boldsymbol{p})$, then*

$$\sum_{h=1}^H C^h(\boldsymbol{q}^h) = C\left(\sum_{h=1}^H \boldsymbol{q}^h\right). \tag{A1}$$

Proof. Note that

$$\begin{aligned}\max_{(\boldsymbol{q}^h)_{h=1}^H} \boldsymbol{p}\cdot\left(\sum_{h=1}^H \boldsymbol{q}^h\right) - \sum_{h=1}^H C^h(\boldsymbol{q}^h) &= \max_{\boldsymbol{q}\in\mathbb{R}^I_+} \max_{(\boldsymbol{q}^h)_{h=1}^H : \sum_{h=1}^H \boldsymbol{q}^h = \boldsymbol{q}} \boldsymbol{p}\cdot\left(\sum_{h=1}^H \boldsymbol{q}^h\right) - \sum_{h=1}^H C^h(\boldsymbol{q}^h) \\ &= \max_{\boldsymbol{q}\in\mathbb{R}^I_+} \boldsymbol{p}\cdot\boldsymbol{q} - C(\boldsymbol{q}).\end{aligned} \tag{A2}$$

Consider $\boldsymbol{q}_0 = \sum_{h=1}^H \boldsymbol{q}_0^h$ with each $\boldsymbol{q}_0^h$ is in $\boldsymbol{S}^h(\boldsymbol{p})$. We have that

$$\boldsymbol{p}\cdot\boldsymbol{q}_0 - C(\boldsymbol{q}_0) \geq \boldsymbol{p}\cdot\left(\sum_{h=1}^H \boldsymbol{q}_0^h\right) - \sum_{h=1}^H C^h(\boldsymbol{q}_0^h). \tag{A3}$$

Optimality of the $\boldsymbol{q}_0^h$ implies that the right-hand side of this inequality attains the maximum in the left-hand side of equation (A2). Hence,

$$\boldsymbol{p}\cdot\boldsymbol{q}_0 - C(\boldsymbol{q}_0) \geq \max_{\boldsymbol{q}\in\mathbb{R}^I_+} \boldsymbol{p}\cdot\boldsymbol{q} - C(\boldsymbol{q}),$$

so that $\boldsymbol{q}_0$ is in $\boldsymbol{S}(\boldsymbol{p})$. Inequality (A3) holds as an equality, which implies (A1) as desired.

Conversely, consider $\boldsymbol{q}_0$ in $\boldsymbol{S}(\boldsymbol{p})$. Equation (A2) implies that

$$\boldsymbol{p}\cdot\boldsymbol{q}_0 - C(\boldsymbol{q}_0) = \max_{(\boldsymbol{q}^h)_{h=1}^H} \boldsymbol{p}\cdot\left(\sum_{h=1}^H \boldsymbol{q}^h\right) - \sum_{h=1}^H C^h(\boldsymbol{q}^h).$$

Let $(\boldsymbol{q}_0^h)$ be supply vectors that minimize the sum of costs conditional on total supply being $\boldsymbol{q}_0$. Then

$$\boldsymbol{p}\cdot\boldsymbol{q}_0 - C(\boldsymbol{q}_0) = \boldsymbol{p}\cdot\left(\sum_{h=1}^H \boldsymbol{q}_0^h\right) - \sum_{h=1}^H C^h(\boldsymbol{q}_0^h).$$

Therefore, each $\boldsymbol{q}_0^h$ is in $\boldsymbol{S}^h(\boldsymbol{q}^h)$.

□

Proof of Theorem 1. Let $(\boldsymbol{p}^*, (\boldsymbol{q}^{h*})_{h=1,\dots,H})$ maximize hospital welfare subject to all hospitals choosing supply optimally given $\boldsymbol{p}$ and subject to not promising more transplants than are produced. Mathematically, the tuple $(\boldsymbol{p}^*, \boldsymbol{q}^*, (\boldsymbol{q}^{h*})_{h=1,\dots,H})$ in $\mathbb{R}^I \times \mathbb{R}^I_+ \times \mathbb{R}^{IH}_+$ maximizes

$$f(\boldsymbol{q}) - \sum_{h=1}^h C^h(\boldsymbol{q}^h)$$

subject to

$$\boldsymbol{q} = \sum_{h=1}^H \boldsymbol{q}^h, \tag{A4}$$

to each

$$\boldsymbol{q}^h \in \boldsymbol{S}^h(\boldsymbol{p}), \tag{A5}$$

and to

$$\boldsymbol{p}\cdot\boldsymbol{q} = f(\boldsymbol{q}). \tag{A6}$$

Lemma 1 implies that this maximization problem is equivalent to finding a pair $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ in $\mathbb{R}^I \times \mathbb{R}^I_+$ that maximizes

$$f(\boldsymbol{q}) - C(\boldsymbol{q}) \tag{A7}$$

subject to

$$\boldsymbol{q} \in \boldsymbol{S}(\boldsymbol{p}) \tag{A8}$$

and to (A6).

Aggregate cost is smooth and strictly convex, and the solution is interior. So constraint (A8) is equivalent to $\boldsymbol{p} = \boldsymbol{\nabla C}(\boldsymbol{q})$. Thus, $\boldsymbol{q}^*$ maximizes (A7) in $\mathbb{R}^I_{++}$ subject to

$$\boldsymbol{\nabla C}(\boldsymbol{q})\cdot\boldsymbol{q} = f(\boldsymbol{q}).$$

The production function and aggregate cost are smooth, and $\mathbb{R}^I_{++}$ is an open set. Therefore, the Lagrange multiplier theorem implies that there exists λ such that $\boldsymbol{q}^*$ maximizes

$$f(\boldsymbol{q}) - C(\boldsymbol{q}) + \lambda \cdot \{f(\boldsymbol{q}) - \boldsymbol{\nabla C}(\boldsymbol{q}) \cdot \boldsymbol{q}\}.$$

Setting the derivative equal to zero, we have

$$\boldsymbol{\nabla f} - \boldsymbol{\nabla C} = \frac{\lambda}{1+\lambda} \boldsymbol{q'D^2C}.$$

To solve for the Lagrange multiplier, we multiply by $\boldsymbol{q}$ on the right, and use the equality $\boldsymbol{\nabla C} \cdot \boldsymbol{q} = f$. We have

$$\boldsymbol{\nabla f} \cdot \boldsymbol{q} - f = \frac{\lambda}{1+\lambda} \boldsymbol{q'D^2Cq}.$$

Therefore,

$$(\boldsymbol{\nabla f} \cdot \boldsymbol{q} - f)(\boldsymbol{q'D^2Cq})^{-1} = \frac{\lambda}{1+\lambda}.$$

Substituting λ we get the final formula,

$$\boldsymbol{\nabla C} = \boldsymbol{\nabla f} - (\boldsymbol{\nabla f} \cdot \boldsymbol{q} - f) \cdot (\boldsymbol{q'D^2Cq})^{-1} \cdot (\boldsymbol{q'D^2C}).$$

Part 1 of the theorem follows by substituting $\boldsymbol{p} = \boldsymbol{\nabla C}$. Part 2 follows because, with constant returns to scale, the constraint (A6) is automatically satisfied. Part 3 follows because welfare and social welfare coincide.

□

A.2 Additional Proofs

Proof of Proposition 1. Note that, by Lemma (1) and the fact that $\boldsymbol{q}_0$ and $\boldsymbol{q}^*$ are interior, $\boldsymbol{p}_0 = \boldsymbol{\nabla C}(\boldsymbol{q}_0)$, and the optimal reward vector $\boldsymbol{p}^* = \boldsymbol{\nabla C}(\boldsymbol{q}^*)$. In what follows, we take $\boldsymbol{q}_0$ sufficiently close to $\boldsymbol{q}^*$ such that the matrix $\boldsymbol{D^2C}(\boldsymbol{q}_0) - \boldsymbol{D^2f}(\boldsymbol{q}_0)$ is non-singular.

Part 1: Approximation of $\boldsymbol{q}^* - \boldsymbol{q}_0$.

Theorem (1) implies that

$$\boldsymbol{\nabla C}(\boldsymbol{q}^*) - \boldsymbol{\nabla f}(\boldsymbol{q}^*) = -\boldsymbol{A}.$$

Taking derivatives, we have

$$\boldsymbol{\nabla C}(\boldsymbol{q}_0) - \boldsymbol{\nabla f}(\boldsymbol{q}_0) + (\boldsymbol{q}^* - \boldsymbol{q}_0)'(\boldsymbol{D^2C}(\boldsymbol{q}_0) - \boldsymbol{D^2f}(\boldsymbol{q}_0)) + \boldsymbol{\epsilon_1}(\boldsymbol{q}_0) = -\boldsymbol{A},$$

where $\boldsymbol{\epsilon_1}(\boldsymbol{q}_0)$ is in $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|)$. Therefore,

$$\boldsymbol{q}^* - \boldsymbol{q}_0 = (\boldsymbol{D^2C}(\boldsymbol{q}_0) - \boldsymbol{D^2f}(\boldsymbol{q}_0))^{-1} \cdot (\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0)' + \boldsymbol{\epsilon}_2(\boldsymbol{q}_0), \tag{A9}$$

where the error term

$$\boldsymbol{\epsilon}_2(\boldsymbol{q}_0) = -(\boldsymbol{D^2C}(\boldsymbol{q}_0) - \boldsymbol{D^2f}(\boldsymbol{q}_0))^{-1}(\boldsymbol{\epsilon_1}(\boldsymbol{q}_0) + \boldsymbol{A})'.$$

Part 2: Approximation of the deadweight loss.

Define the function

$$g(t) = f(\boldsymbol{q}^*) - C(\boldsymbol{q}^*) - [f(\boldsymbol{q}^* + t \cdot (\boldsymbol{q}_0 - \boldsymbol{q}^*)) - C(\boldsymbol{q}^* + t \cdot (\boldsymbol{q}_0 - \boldsymbol{q}^*))].$$

We have that $g(0) = 0$, and $g(1)$ is the deadweight loss. The derivative of g equals

$$g'(t) = -[\boldsymbol{\nabla f}(\boldsymbol{q}^* + t \cdot (\boldsymbol{q}_0 - \boldsymbol{q}^*)) - \boldsymbol{\nabla C}(\boldsymbol{q}^* + t \cdot (\boldsymbol{q}_0 - \boldsymbol{q}^*))] \cdot (\boldsymbol{q}_0 - \boldsymbol{q}^*).$$

Taking derivatives we have

$$g'(t) = -[\boldsymbol{\nabla f}(\boldsymbol{q}^*) - \boldsymbol{\nabla C}(\boldsymbol{q}^*) + t(\boldsymbol{q}_0 - \boldsymbol{q}^*)'(\boldsymbol{D^2 f}(\boldsymbol{q}_0) - \boldsymbol{D^2 C}(\boldsymbol{q}_0)) + \boldsymbol{\epsilon}_3(\boldsymbol{q}_0)] \cdot (\boldsymbol{q}_0 - \boldsymbol{q}^*),$$

where the error $\boldsymbol{\epsilon}_3(\boldsymbol{q}_0)$ is in $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|)$. By Theorem (1),

$$\begin{aligned} g'(t) =& t(\boldsymbol{q}^* - \boldsymbol{q}_0)'(\boldsymbol{D^2 C}(\boldsymbol{q}_0) - \boldsymbol{D^2 f}(\boldsymbol{q}_0)(\boldsymbol{q}^* - \boldsymbol{q}_0) \\ &+ [\boldsymbol{A} + \boldsymbol{\epsilon}_3(\boldsymbol{q}_0)] \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0). \end{aligned}$$

We can substitute $(\boldsymbol{D^2 C}(\boldsymbol{q}_0) - \boldsymbol{D^2 f}(\boldsymbol{q}_0)(\boldsymbol{q}^* - \boldsymbol{q}_0)$ using equation (A9). This yields

$$\begin{aligned} g'(t) =& t(\boldsymbol{q}^* - \boldsymbol{q}_0)' \cdot (\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0)' \\ &- t(\boldsymbol{q}^* - \boldsymbol{q}_0)' \cdot (\boldsymbol{\epsilon_1}(\boldsymbol{q}_0) + \boldsymbol{A})' \\ &+ [\boldsymbol{A} + \boldsymbol{\epsilon}_3(\boldsymbol{q}_0)] \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0). \end{aligned}$$

Integrating in t from 0 to 1, and noting that the dot products are the same if we take transposes, the deadweight loss equals

$$\begin{aligned} g(1) =& \frac{1}{2}(\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0) \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) \\ &+ \frac{1}{2}\boldsymbol{A} \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) \\ &+ \left(\boldsymbol{\epsilon}_3(\boldsymbol{q}_0) - \frac{1}{2}\boldsymbol{\epsilon}_1(\boldsymbol{q}_0)\right) \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0). \end{aligned} \tag{A10}$$

This establishes the first approximation formula.

To establish the second approximation formula we substitute $\boldsymbol{q}^* - \boldsymbol{q}_0$ from equation (A9), yielding

$$\begin{aligned} g(1) =& \frac{1}{2}(\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0)(\boldsymbol{D^2 C}(\boldsymbol{q}_0) - \boldsymbol{D^2 f}(\boldsymbol{q}_0))^{-1} \cdot (\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0)' \\ &- (\boldsymbol{\nabla f}(\boldsymbol{q}_0) - \boldsymbol{p}_0)(\boldsymbol{D^2 C}(\boldsymbol{q}_0) - \boldsymbol{D^2 f}(\boldsymbol{q}_0))^{-1}\boldsymbol{A}' + \frac{1}{2}\boldsymbol{A} \cdot (\boldsymbol{q}^* - \boldsymbol{q}_0) \\ &+ \boldsymbol{\epsilon}_4(\boldsymbol{q}_0), \end{aligned} \tag{A11}$$

where $\epsilon_4(\boldsymbol{q}_0)$ is in $o(\|\boldsymbol{q}^* - \boldsymbol{q}_0\|^2)$.

□

Proposition 2. The assumption that C^h and C are strictly convex implies that the supply and aggregate supply correspondences are single-valued. Define the aggregate externality given quantity $\boldsymbol{q}$ as

$$\tilde{\boldsymbol{E}}(\boldsymbol{q}) = E(\boldsymbol{S}^1(\boldsymbol{\nabla C}(\boldsymbol{q})), \ldots, \boldsymbol{S}^1(\boldsymbol{\nabla C}(\boldsymbol{q}))). \tag{A12}$$

Let $(\boldsymbol{p}^*, (\boldsymbol{q}^{h*})_{h=1,\ldots,H})$ maximize social welfare subject to all hospitals choosing supply optimally given $\boldsymbol{p}$ and subject to not promising more transplants than are produced. Mathematically, the pair $(\boldsymbol{p}^*, (\boldsymbol{q}^{h*})_{h=1,\ldots,H})$, in $\mathbb{R}^I \times \mathbb{R}^{IH}_+$ maximizes

$$f(\boldsymbol{q}) - \sum_{h=1}^{H} C^h(\boldsymbol{q}^h) + \sum_{h=1}^{H} E^h(\boldsymbol{q}^h)$$

subject to constraints (A4), (A5), and (A6). By Lemma 1, the maximization problem is equivalent to finding $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ in $\mathbb{R}^I \times \mathbb{R}^I_+$ that maximize

$$f(\boldsymbol{q}) - C(\boldsymbol{q}) + E(\boldsymbol{S}^1(\boldsymbol{p}), \ldots, \boldsymbol{S}^h(\boldsymbol{p}))$$

subject to constraints (A8) and (A6). Because the maximum is assumed to be strictly positive, we can restrict attention to $\boldsymbol{q}$ in $\mathbb{R}^I_{++}$. In this case, because C is smooth and convex, constraint (A8) is equivalent to $\boldsymbol{p} = \boldsymbol{\nabla C}(\boldsymbol{q})$. Hence, the maximization problem is equivalent to finding $\boldsymbol{q}^*$ in $\mathbb{R}^I_{++}$ that maximizes

$$f(\boldsymbol{q}) - C(\boldsymbol{q}) + \tilde{\boldsymbol{E}}(\boldsymbol{q})$$

subject to

$$\boldsymbol{\nabla C}(\boldsymbol{q}) \cdot \boldsymbol{q} = f(\boldsymbol{q}).$$

By the Lagrange multiplier theorem, there exists a multiplier λ^{SW} such that, at the optimum,

$$\boldsymbol{\nabla f} - \boldsymbol{\nabla C} + \boldsymbol{\nabla \tilde{E}} + \lambda^{SW}(\boldsymbol{\nabla f} - \boldsymbol{\nabla C} - \boldsymbol{q'D^2C}) = \boldsymbol{0}.$$

Substituting $\boldsymbol{p}^* = \boldsymbol{\nabla C}(\boldsymbol{q}^*)$ we obtain the formula for the adjustment term. To obtain the formula for the Lagrange multiplier we multiply on the right by $\boldsymbol{q}^*$ and use the fact that $\boldsymbol{\nabla f}(\boldsymbol{q}^*) \cdot \boldsymbol{q}^* = \boldsymbol{p}^* \cdot \boldsymbol{q}^* = f(\boldsymbol{q}^*)$. The second and third parts of the proposition follow from the formula for the Lagrange multiplier.

□

Proposition 3. The platform chooses $\boldsymbol{q}$ in $\mathbb{R}^I_+$ to maximize

$$f(\boldsymbol{q})$$

subject to

$$f(\boldsymbol{q}) = \boldsymbol{P_S}(\boldsymbol{q}) \cdot \boldsymbol{q}.$$

Because the solution is interior, there exists a Lagrange multiplier λ such that

$$\boldsymbol{\nabla f} + \lambda(\boldsymbol{\nabla f} - \boldsymbol{P_S} - \boldsymbol{q'DP_S}) = \boldsymbol{0}.$$

Substituting that the optimal rewards $\boldsymbol{p} = \boldsymbol{P_S}(\boldsymbol{q})$, we obtain

$$\boldsymbol{p} = \boldsymbol{\nabla f} + \frac{1}{\lambda}\boldsymbol{\nabla f} - \boldsymbol{q'DP_S}.$$

To calculate the Lagrange multiplier, we right multiply by $\boldsymbol{q'}$ and use $\boldsymbol{pq'} = \boldsymbol{\nabla f q'} = f(\boldsymbol{q})$ to obtain

$$\lambda = \frac{f(\boldsymbol{q})}{\boldsymbol{q'DP_S}(\boldsymbol{q})\boldsymbol{q}}.$$

These two formulas imply part 1 of the proposition statement. The observation in part 2 follows directly from the formula in part 1.

□

B Data Appendix

This study used five main anonymized data sets: a database of all kidney exchange transplants done in the US from January 1, 2008 through December 4, 2014 (the OPTN transplant data), databases of all kidney exchange transplants organized by each of the three largest multi-hospital platforms in the US (the NKR, APD, and UNOS transplant data), and a database of all patient and donor registrations to the largest of those platforms (the NKR registration data).

B.1 Transplant data

In order to document the kidney exchange market, we merged the OPTN transplant data with the transplant data from NKR, APD, and UNOS. In what follows, we will describe these data and the merge procedure we used.

Obtaining the datasets The OPTN provided us with a dataset on all transplants conducted in the US, known as the Standard Transplant Analysis and Research (STAR) dataset. The STAR dataset by itself lacks two key pieces of information: the transplant hospitals where the kidney was put into the patient and removed from the donor (which we use to determine whether a transplant is internal or external) and the unacceptable antigens for the patient (which we need to measure sensitization). These supplemental pieces of information are also available from the OPTN on request. Merging is done by using OPTN identifiers. The OPTN database contains records on 4377 kidney exchange transplants.

We obtained each of the platform datasets directly from the platform. The platform datasets contain records on 1400 kidney exchange transplants in total: 1193 from NKR, 100 from APD, and 107 from UNOS.

Dataset merge algorithm In order to identify which transplants in the comprehensive OPTN database were organized by each of the three platforms, we matched records in the platform files to records in the OPTN file. However, because all datasets are anonymized, the merge must be done on the biological characteristics of each transplant's recipient and donor ans logistical information on the transplant itself.

Fortunately, we can also use the fact that the OPTN database is comprehensive while the platform database is not.[1] This implies that the transplants in platform data are a subset of transplants in the OPTN data if record-keeping is perfect. However, without perfect record-keeping, we should be wary of false matches. Our matching procedure is designed to limit false matches by keeping only the highest quality match that meets a minimum threshold. We formalize this idea below.

Let the set of records in the platform data be D_p and the set in the OPTN data be D_o. Let the **universe** of acceptable matches of one platform record $r_p \in D_p$ to one OPTN record $r_o \in D_o$ be represented by $U \subseteq D_p \times D_o$. Define the set of **collisions** a match (r_p, r_o) has relative to some set X to be $\kappa(r_p, r_o, X) \equiv \left\{\left(r'_p, r'_o\right) \in X \mid r'_p = r_p \text{ or } r'_o = r_o\right\} \setminus \{(r_p, r_o)\}$. A set of matches $M \subseteq U$ is called a **merge** if none its matches have any collisions relative to the merge, that is, for each $(r_p, r_o) \in M$, we have $\kappa(r_p, r_o, M) = \emptyset$. Any record that is not part of a match in merge M should be interpreted as unmatched.

Now, we can discuss what makes an individual match good. Define a **ranked criterion** to be a series of N subsets of the universe $C = (C^n)_{n=0}^N$ such that $C^n \subseteq C^{n+1}$, $C^0 = \emptyset$, and $C^N = U$. The ranked criterion codifies a hierarchy of match quality levels. Further, define the **rank** of a match (r_p, r_o) to be $\rho(r_p, r_o) \equiv \min\{n \mid (r_p, r_o) \in C^n\}$.[2]

Finally, we are ready to state our notion of merge quality. Recall that the main idea is that a match of rank n should be in the merge if and only if it is the unique match of rank n or better whose component records aren't part of a better match in the merge. For a given set X and rank level n, define the **universe excluding records that matched better than n in X** to be

$$U^n(X) \equiv U \setminus \left[\bigcup_{(r_p,r_o)\in X\cap C^{n-1}} \kappa(r_p, r_o, U)\right].$$

Then, a set of matches $M \subseteq U$ is **stable relative to ranked criterion C** if it satisfies

$$(r_p, r_o) \in M \Leftrightarrow \left\{\begin{array}{c}(r_p, r_o) \in U^{\rho(r_p,r_o)}(M) \\ \text{and} \\ \kappa\left(r_p, r_o, C^{\rho(r_p,r_o)} \cap U^{\rho(r_p,r_o)}(M)\right) = \emptyset.\end{array}\right\}.$$

The first condition on the right-hand side ensures that (r_p, r_o) doesn't collide with a strictly better match in M, while the second condition ensures that (r_p, r_o) is the unique such match.

[1] The OPTN is required by federal administrative law to "[m]aintain records of all transplant candidates, all organ donors and all transplant recipients" (42 C.F.R. § 121.11(a)(1)(ii)).

[2] Note that we could equivalently start with a rank function and construct C to be the lower contour sets of that function.

Note that any set M that is stable relative to ranked criterion C is necessarily a merge (see Lemma 5 below).

Finding a stable merge relative to some ranked criterion C is a simple matter of following Algorithm 1. It is also true that for a ranked criterion C, the stable merge is unique, so we need not worry that there is some other stable merge that we might prefer on different grounds. Proofs of these claims (which are summarized in Proposition 1) can be found at the bottom of this subsection.

Algorithm 1. *Initialize by setting* $M^0 = \emptyset$. *Then for* $n \in \{1, \ldots, N\}$,

$$X^n = C^n \cap U^n\left(M^{n-1}\right)$$
$$M^n = M^{n-1} \cup \left\{(r_p, r_o) \in X^n \mid \kappa\left(r_p, r_o, X^n\right) = \emptyset\right\}.$$

The output merge is M^N.

Basically, in Step n, the algorithm adds any match to the output set that is the unique rank-n or better match whose records are not involved in a rank $n-1$ or better match that is already in the output set.

Proposition 1. *Algorithm 1 yields the unique set that is stable relative to ranked criterion* C. *Furthermore, this set is a merge.*

Quality of the merge We merged the OPTN and platform transplant databases using Algorithm 1. The set of acceptable matches, U, was the set of all potential matches where the transplant dates are within 31 days of each other and the ages of the donor and the recipient are each within 10 years of each other. The ranked criterion we used is defined as follows.

- C^1 is the set of all acceptable matches where the donor and recipient each match exactly on blood type, sex, the hospital where the transplant was conducted, and all six major HLA alleles (two each on the HLA-A, HLA-B, and HLA-DR loci).
- For $n \in \{2, \ldots, 5\}$, C^n is the set of all acceptable matches where the donor and recipient each match on either blood type or sex and also each match on at least $7 - n$ out of the six major HLA alleles.

Given the ranked criterion described in the previous section, the merge algorithm performed well. The percentage of platform records matched to an OPTN record was 94% overall (94% for NKR, 97% for APD, and 94% for UNOS). Moreover, the matches seem to be high quality: Table B1 reports the percentage of matches that meet various criteria.

Table B1: Agreement Between Matched Records in the Transplant Merge

Platform	Age within 5 years	Transplant date within 1 day	5 or more HLA alleles match	Blood type and gender match	Transplant hospital matches	At most 1 criterion to the left is violated	No criterion to the left is violated
NKR	97.8%	97.4%	94.9%	95.1%	97.2%	97.6%	87.6%
APD	92.8%	95.9%	93.8%	96.9%	95.9%	96.9%	90.7%
UNOS	87.1%	99.0%	96.0%	99.0%	99.0%	99.0%	83.2%
All of the above	96.6%	97.4%	94.9%	95.5%	97.3%	97.6%	89.9%

Confirming the merge is correct It was necessary to use the merge described above to match transplants in the APD and NKR databases to the OPTN transplant database, since all of these datasets are anonymized to different sets of identifiers. However, since UNOS is a contractor for the OPTN, the UNOS and OPTN databases share common identifiers that allow us to see the actual true merge of the UNOS transplant dataset to the OPTN dataset. Comparing true matches to the matches selected by our algorithm, we find that only one UNOS record was incorrectly matched. That is, our algorithm chose the correct match in the OPTN dataset for 99% of the UNOS records it matched. This gives us added confidence that out merge algorithm is working well.

Proofs concerning Algorithm 1 and stable merges To prove Proposition 1, three lemmas are helpful. First, we show that in Algorithm 1, once a match collides in some step of the algorithm, there is no chance that the match will ever be added to the merge.

Lemma 2. *In Algorithm 1, if* $(r_p, r_o) \in X^n$*, then any* $\left(r'_p, r'_o\right) \in \kappa\left(r_p, r_o, X^n\right)$ *cannot be part of* M^N*.*

Proof. Consider the base case $n = N$. By the definition of collision, $\left(r'_p, r'_o\right) \in X^n$, and by the definition of the algorithm, $\left(r'_p, r'_o\right) \notin M^N$.

Now, we prove the induction step: for $n < N$, if $(r_p, r_o) \in X^n$ and $\left(r'_p, r'_o\right) \in \kappa\left(r_p, r_o, X^n\right)$, then $(r_p, r_o) \in X^{n+1}$ and $\left(r'_p, r'_o\right) \in \kappa\left(r_p, r_o, X^{n+1}\right)$. By the definition of collision, both $\left(r'_p, r'_o\right) \in X^n$ and $(r_p, r_o) \in \kappa\left(r'_p, r'_o, X^n\right)$ must hold, and hence by the definition of the algorithm, $\left\{(r_p, r_o), \left(r'_p, r'_o\right)\right\} \cap M^n = \emptyset$. Now, clearly $\left\{(r_p, r_o), \left(r'_p, r'_o\right)\right\} \subseteq C^{n+1}$ since $C^n \subseteq C^{n+1}$, so to show that $\left\{(r_p, r_o), \left(r'_p, r'_o\right)\right\} \subseteq X^{n+1}$, we just need to establish that $\left\{(r_p, r_o), \left(r'_p, r'_o\right)\right\} \subseteq U^{n+1}\left(M^n\right)$.

By way of contradiction, assume otherwise. Then, there must exist some $\left(r''_p, r''_o\right) \in M^n \cap C^n$ such that either $(r_p, r_o) \in \kappa\left(r''_p, r''_o, U\right)$ or $\left(r'_p, r'_o\right) \in \kappa\left(r''_p, r''_o, U\right)$. Without loss of generality, let $(r_p, r_o) \in \kappa\left(r''_p, r''_o, U\right)$. Since $(r_p, r_o) \in X^n$, we know that $\left(r''_p, r''_o\right) \notin M^{n-1}$. Then, for $\left(r''_p, r''_o\right) \in M^n$ to hold, it must be that $\left(r''_p, r''_o\right) \in X^n$. But then, $(r_p, r_o) \in \kappa\left(r''_p, r''_o, X^n\right)$,

which contradicts $\left(r_p'', r_o''\right) \in M^n$. Hence, we have shown that $\left\{\left(r_p, r_o\right), \left(r_p', r_o'\right)\right\} \subseteq X^{n+1}$. Clearly then, $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, X^{n+1}\right)$. Thus we have proved the result via induction. □

Next, we show that in Algorithm 1, any match is either added to the merge in the step equal to its rank, or it is never added at all.

Lemma 3. *In Algorithm 1, if (r_p, r_o) isn't in $M^{\rho(r_p, r_o)}$, then it is not in M^N.*

Proof. The statement is trivially true if $\rho\left(r_p, r_o\right) = N$, so we consider $\rho\left(r_p, r_o\right) = n < N$. By the definition of the algorithm, Step n is the first step at which (r_p, r_o) could join the merge.

Now, if $(r_p, r_o) \notin M^n$ because $(r_p, r_o) \notin X^n$, then for any $n' > n$, $(r_p, r_o) \notin M^n$. To see this, note that since $(r_p, r_o) \in C^n$, if $(r_p, r_o) \notin X^n$, then it must be that $(r_p, r_o) \notin U^n\left(M^{n-1}\right)$. By definition, we know that $U^{n'}\left(X'\right) \subseteq U^n\left(X\right)$, for any $n' > n$ and $X \subseteq X'$. Hence, $U^{n'}\left(M^{n'-1}\right) \subseteq U^n\left(M^{n-1}\right)$ for any $n' > n$ and hence $(r_p, r_o) \notin X^{n'}$. From this, it follows that for any $n' > n$, $(r_p, r_o) \notin M^{n'}$and hence $(r_p, r_o) \notin M^N$.

Now, if (r_p, r_o) fails to join the merge at Step n and $(r_p, r_o) \in X^n$, then it must be that there exists some $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, X^N\right)$. By Lemma 2, we conclude that $\left\{\left(r_p, r_o\right), \left(r_p, r_o\right)\right\} \cap M^N = \emptyset$. □

Then, we establish the existence of a set that is stable relative to ranked criterion C by showing that Algorithm 1 produces it.

Lemma 4. *Algorithm 1 outputs a set that is stable relative to ranked criterion C.*

Proof. To prove that the algorithm's output, M, is stable relative to C, we must consider both the forward and backward implication in the definition. First, consider the forward implication, which requires that for any $(r_p, r_o) \in M$ both $\kappa\left(r_p, r_o, C^n \cap U^n\left(M\right)\right) = \emptyset$ and $(r_p, r_o) \in U^n\left(M\right)$, where $n = \rho\left(r_p, r_o\right)$. We will prove both implications by contradiction.

By way of contradiction, assume that there exists $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, C^n \cap U^n\left(M\right)\right)$. By Lemma 3, (r_p, r_o) must join the merge in Step n. For this to be true, it must be that $(r_p, r_o) \in X^n$. But since $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, C^n \cap U^n\left(M\right)\right)$ and $U^n\left(M\right) \subseteq U^n\left(M^{n-1}\right)$, it must also be that $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, X^n\right)$, which by Lemma 2 means that $(r_p, r_o) \notin M$, providing the required contradiction.

Now, by way of contradiction, assume that $(r_p, r_o) \notin U^n\left(M\right)$. Since Lemma 3 tells us that $M^{n-1} = M \cap C^{n-1}$, we can conclude that $U^n\left(M\right) = U^n\left(M^{n-1}\right)$. This then tells us that $(r_p, r_o) \notin X^n$, and hence that $(r_p, r_o) \notin M^n$, providing the required contradiction.

Now, consider the backward implication. By way of contradiction, assume that $\kappa\left(r_p, r_o, C^n \cap U^n\left(M\right)\right) = \emptyset$ and $(r_p, r_o) \in U^n\left(M\right)$, but $(r_p, r_o) \notin M$, where $n = \rho\left(r_p, r_o\right)$. It must be that $(r_p, r_o) \in X^n$, since $U^n\left(M\right) \subseteq U^n\left(M^{n-1}\right)$. Hence, $(r_p, r_o) \notin M$ requires that there exists $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, C^n \cap U^n\left(M^{n-1}\right)\right)$. But, Lemma 3 tells us that $M^{n-1} = M \cap C^{n-1}$, so $U^n\left(M^{n-1}\right) = U^n\left(M\right)$. Hence, $\left(r_p', r_o'\right) \in \kappa\left(r_p, r_o, C^n \cap U^n\left(M\right)\right)$, which provides the required contradiction. □

We also need to show that any set that is stable with respect to ranked criterion C is a merge.

Lemma 5. *Any set $M \subseteq U$ that is stable with respect to ranked criterion C must be a merge.*

Proof. By way of contradiction, assume that M is stable with respect to ranked criterion C, but is not a merge. Then, there exists $(r_p, r_o) \in M$ such that $\kappa(r_p, r_o, M) \neq \emptyset$. Let $(r'_p, r'_o) \in \kappa(r_p, r_o, M)$. Now, if $\rho(r_p, r_o) \neq \rho(r'_p, r'_o)$, then without loss of generality, let $\rho(r_p, r_o) > \rho(r'_p, r'_o)$. Then, it must be that $(r_p, r_o) \notin U^{\rho(r_p, r_o)}(M)$, which contradicts the definition of stable with respect to ranked criterion C.

So, assume that $\rho(r_p, r_o) = \rho(r'_p, r'_o)$. Further, assume that $\{(r_p, r_o), (r'_p, r'_o)\} \subseteq U^{\rho(r_p, r_o)}(M)$; otherwise we have already contradicted the definition of stable. Then, $\{(r_p, r_o), (r'_p, r'_o)\} \subseteq C^{\rho(r_p, r_o)} \cap U^{\rho(r_p, r_o)}(M)$, and hence $\kappa\left(r_p, r_o, C^{\rho(r_p, r_o)} \cap U^{\rho(r_p, r_o)}(M)\right) \neq \emptyset$, contradicting the definition of stable with respect to ranked criterion C. □

Finally, we are ready to prove that the stable match relative to C produced by Algorithm 1 is in fact the unique stable merge relative to C.

Proof of Proposition 1. Lemma 4 shows that Algorithm 1 outputs a stable match. Uniqueness can be established by induction on rank. For the base case, consider some $(r_p, r_o) \in M \cap C^1$. Since M is stable, it must be that $\kappa(r_p, r_o, C^1) = \emptyset$. But this would require that (r_p, r_o) be included in M' as well. So, M and M' must agree on their intersection with C^1.

Now, for the inductive step, assume that $M \cap C^n = M' \cap C^n$. Take a match $(r_p, r_o) \in (M \setminus C^n) \cap C^{n+1}$. Since M is stable, it must be that $\kappa(r_p, r_o, C^{n+1} \cap U^{n+1}(M \cap C^n)) = \emptyset$. But since $M \cap C^n = M' \cap C^n$, it is also true that $\kappa(r_p, r_o, C^{n+1} \cap U^{n+1}(M' \cap C^n)) = \emptyset$, which means that $(r_p, r_o) \in M'$. So, by induction, any two merges must contain the same matches, that is, the stable match is unique. □

B.2 Registration data

In this subsection, we describe how the list of registrations to the NKR was assembled. The NKR provided us with snapshot files of the patient and donor pool between April 1, 2012 and December 4, 2014. These files are typically daily snapshots except for some missing periods, each of which is up to a month in length. Each snapshot corresponds to a different date and includes basic medical records for each patient and donor in the pool, their listing dates, the related patient for each donor (if any), and whether a patient is unpaired. From these snapshots we recover patient and donors departures, which may be due to being transplanted or other non-documented reasons. A small number of patients and donors depart without a transplant during the period of a missing snapshot; for these, we use bounds on the departure time using the two closest available snapshots, before and after the real departure date. These snapshots also include each patient's set of donors that may not be matched (i.e. are blocked) despite being virtually compatible. Some of these blocked donors are due to patient preference (not to match with these donors) and others are due to match failures.

C Simulation Details

We now provide details on the procedure used in Section 5.1.

C.1 Matching Offers

C.1.1 The algorithm and examples

We solve the linear programming problem described in Section 5.1 as follows. Because it is computationally burdensome to compute all cycles in a pool with many patients and donors, we first solve the relaxed problem by ignoring the constraint that cycles cannot involve more than three transplants. Specifically, we solve:

$$\max_{x_{jk}\in\{0,1\}} \sum_{jk} c_{jk} w_{jk} x_{jk}$$
$$\text{s.t.} x_{jk} - \sum_{l} x_{kl} = 0 \text{ for all } k \in \mathcal{P}$$
$$\sum_{j} x_{jk} \leq 1, \sum_{k} x_{jk} \leq 1$$

where $x_{jk} = 1$ denotes a proposed transplant from the donor in $k \in \mathcal{A} \cup \mathcal{P}$ to the patient in $j \in \mathcal{P} \cup \mathcal{U}$, w_{jk} is the weights described in Section C.1.2 below, and $c_{jk} = 1$ if a transplant from k to j is allowed and 0 otherwise. The first constraint ensures that a donor who is part of a pair is only asked to donate an organ if the intended recipient has been proposed a transplant. The second constraint ensures that no donor or recipient is involved in more than one transplant.

If the solution to the problem does not involve any long cycles, i.e. there do not exist $j_1, \ldots, j_4 \in \mathcal{P}$ such that $x^*_{j_{k+1}j_k} = 1$ for $k \in \{1, \ldots, 3\}$, then it must be that x^* is optimal, given the no long-cycle constraints, and is our desired solution. In our simulations 87.2% of "simulation days" the solution to this relaxed problem yields a feasible match without any further cycle restrictions.

If the solution to this problem contains at least one long cycle, then we proceed as follows. We begin by following the algorithm in Anderson et al. (2015). The algorithm includes a constraint that explicitly prohibits all long cycles in x^*, i.e. for each sequence $j_1, \ldots, j_K \in \mathcal{P}$ such that $x^*_{j_{k+1}j_k} = 1$ for $k \in \{1, \ldots, K-1\}$, we include a constraint in the problem above to ensure that $\prod_{k=1}^{K-1} x_{j_{k+1}j_k} = 0$. If the solution to the modified problem also contains long cycles, we modify the problem again to prohibit those cycles. We iterate this procedure up to ten times. This procedure yields a feasible solution in about 50% of the remaining cases (about 7.8% of all cases) with an average of approximately 6.3 iterations.

If the algorithm above does not yield a feasible solution even after 10 repetitions, we proceed to the next phase in which we use Johnson (1977)'s algorithm to compute cycles and explicitly add constraints that prohibit long cycles. This algorithm searches the compatibility graph

induced by c to calculate cycles. We enumerate and add a constraint to our program to prohibit any long cycles we have found. Since the number of feasible long cycles is extremely large, we search for cycles with a time-out of one second. We find a solution to the problem with these additional constraints and terminate our algorithm if the solution is feasible. This procedure is repeated once more, if necessary. At the end of this phase, we are able to find an optimal solution to the full problem in about 99.0% of the "simulation days."

For the remaining 1.0% of days in the simulation, our matching algorithm still ends up with long cycles. Whenever this is the case, we abandon the match on that "simulation day" that no match is offered, and all patients and donors are returned to the kidney exchange pool to wait for the next day for the transplant offer. The typical average duration for the heuristic of "no match" is around 2.9 days.

Figure C1 illustrates a few kidney exchange pools. The left panel shows compatibility as captured by c, feasible transplants, and the optimal match. Blue dots denote patient-donor pairs, and magenta dots denote altruistic donors. We ignore unpaired patients in this illustration for simplicity. A blue arrows depicts a feasible transplant with the origin of the arrow denoting the donor. Red and green arrows depict cycles and chains, respectively, in the optimal match. Given feasible transplants on left, our match algorithm offers the one on the right. The figure shows there may be several feasible transplants, and in these cases, the optimal match may be relatively easy to determine. Figure C2, on the other hand, illustrates a relatively hard-to-match problem where the optimal match is relatively more difficult to determine.

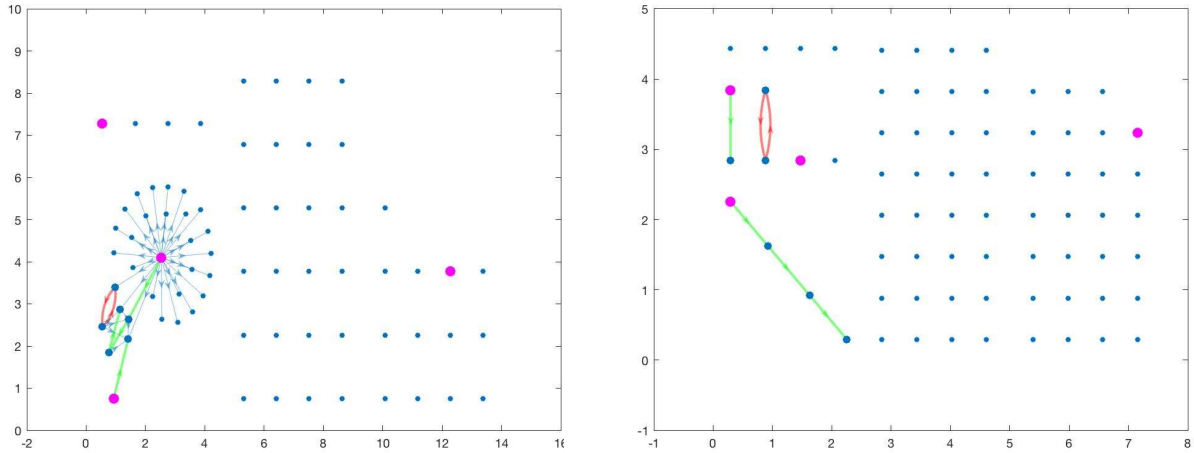


Figure C1: An optimal match in a simple pool

C.1.2 Weights

We attempt to closely match the weights, w_{ij}, on a NKR transplant between patient i and donor j. These weights are designed to favor patients who are highly sensitized, in other words, who are harder to transplant. To define weights, w, NKR first defines a matching power for each submission. Each patient has a Patient Match Power (PMP), a number

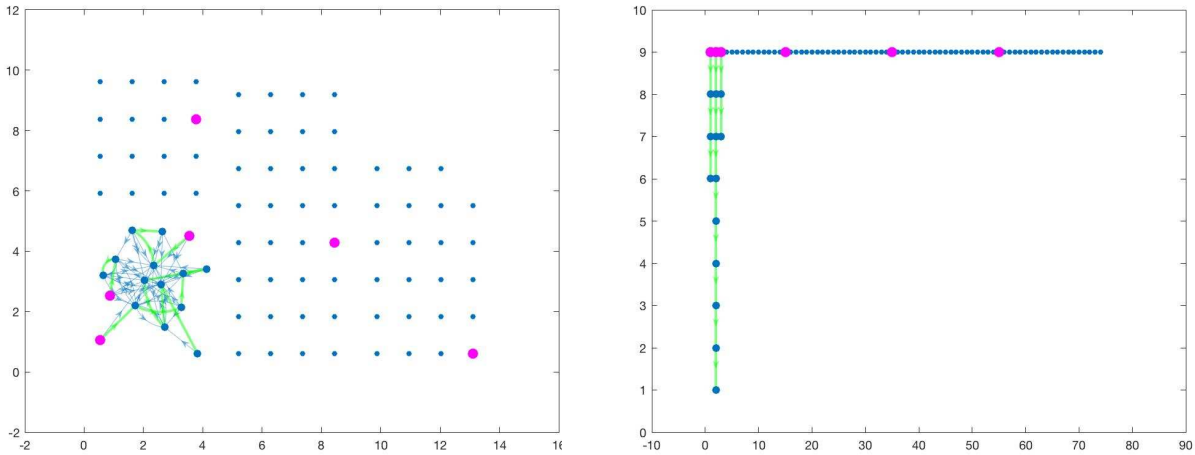


Figure C2: A complicated kidney exchange pool

between 0 and 1, that is a fraction of compatible donors in the NKR pool for that patient. A low PMP for patient i implies that few donors are compatible with patient i. Similarly, the Donor Match Power (DMP) is defined as the fraction of patients in the NKR pool with whom that donor is compatible. Because these quantities and the pool used by the NKR to compute these match powers are not directly observed in our dataset, we calculate them using our sample.

Given these characteristics, NKR calculates a scaled measure of how likely a feasible transplant can occur between i and j, $WNKR_{ij}$. Specifically,

$$WNKR_{ij} = PMP * DMP * 10200.$$

A low $WNKR_{ij}$ correlates with a transplant between i and j being unlikely. It is important to note that the magnitude of $WNKR_{ij}$ is not related to the success of a transplant if it turns out to be feasible. These weights therefore accord higher priority to hard-to-match patients and donors. Using $WNKR_{ij}$, NKR assigns the weights w_{ij} as follows:

$WNKR_{ij}$ *interval*	w_{ij}
$WNKR_{ij} > 70$	1
$25 < WNKR_{ij} \leq 70$	1.01
$5 < WNKR_{ij} \leq 25$	1.2
$WNKR_{ij} \leq 5$	1.5

Because these weights are less than 2, they typically maximize the total number of transplants. However, these weights may sometimes result in two transplants, each with weight 1.5, instead of three transplants with weight 1 each.

C.2 Arrival and Departure

C.2.1 Arrival process

We assume the daily number of submissions in the NKR is given by the a poisson distribution with parameter λ. We estimate that parameter $\lambda = 1.975$, the mean arrival rate for NKR. In each period, our simulations draw a number, say n_t,from this distribution. Then we draw n_t submissions with replacement from the entire pool that ever registered in the NKR during the April 2012 to June 2014 sample period.

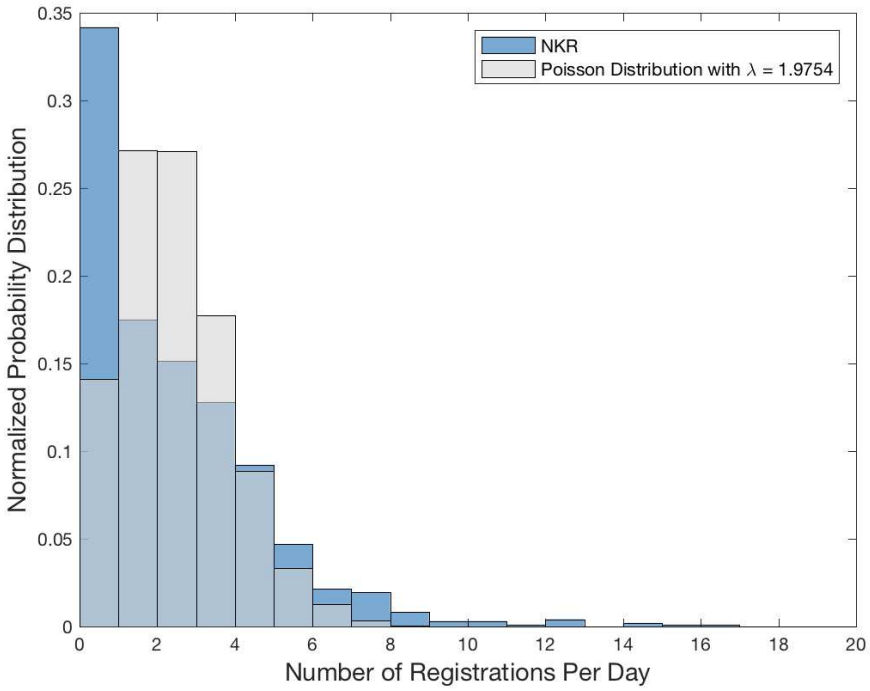


Figure C3: Distribution of NKR and Poisson Number of Submissions Per Day

Figure *C3* shows the fit of the arrival per day distributions of NKR and Poisson. Notice that NKR's distribution has more 0 arrivals per days than the poisson distribution. This mass point is explained by weekends, which appear to have a much lower arrival rate.[3] Figure *C4* shows the arrival per day distribution of NKR for weekdays and our estimated poisson distribution, which shows a better fit.

C.2.2 Departure process

To model departures, we estimate an interval censored hazard model to calculate the rate at which patients and/or donors depart the NKR without a transplant. Specifically, let t_i^a,

[3] Only 40 arrivals in weekends over the course of 140 weeks.

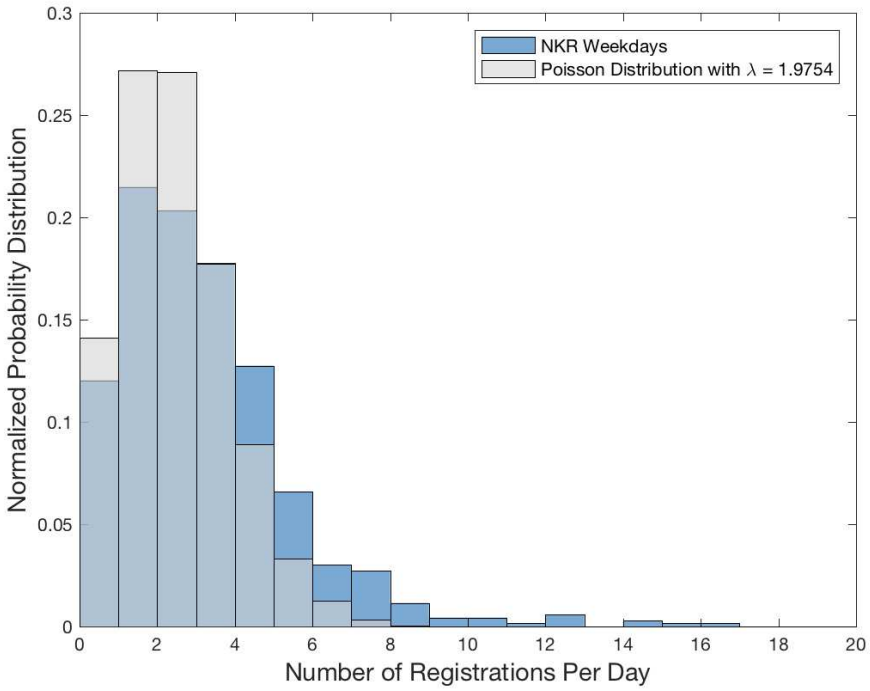


Figure C4: Distribution of NKR and Poisson Number of Submissions Per Day for Weekdays

t_i^*, and t_i^d be the (latent) arrival, transplant, and departure dates for an unpaired patient, donor, or patient-donor pair i. Our dataset records t_{0i} and t_i^* if i was transplanted. Further, if i is transplanted, then we know that $t_i^d > t_i^*$. If i is not transplanted, then in most cases we observe t_i^a, but in some cases, we only know that t_i^a belongs to an interval $\left[t_i^{a-}, t_i^{a+}\right]$ (typically within a week). If i departed without a transplant, we observe t_i^deither exactly or up to a small interval. If i remains in the NKR at the end of our sample, then we know that $t_i^d > T$. Using these observations, we can construct bounds on the duration τ_i that each unit i remains in the NKR without a transplant.

With these observations, we estimate the exponential hazards model. The model is characterized by a survival function

$$S_i(\tau) = exp(-\lambda_i t),$$

where we use the parametric form

$$\lambda_i = \alpha + z_i\beta.$$

The likelihood of the model for the interval censored survival data is straightforward to derive, and estimation via intcens in STATA is straightforward.

Table C2 presents the estimates. Note that the hypothesis tests in a hazard model are reported relative to 1, which implies no effect. As can be seen, patients with blood types that are easier to match or who are paired with easier-to-match donors have a higher departure rate. This finding is consistent with patients and patient-donor pairs departing in response

Table C2: Departure Hazard Rate Estimates

	(1) Patient-Donor Pairs	(2) Unpaired Patients	(3) Altruistic Donors
Patient Matching Power	1.824***	16.35***	
	(0.244)	(2.446)	
Donor Matching Power	0.0699		0.000167
	(0.137)		(0.00126)
Patient Age	0.994*	1.002	
	(0.00349)	(0.00383)	
Donor Age	1.008*		1.011
	(0.00442)		(0.0140)
AB Blood-type Patient	2.465***	1.557*	
	(0.698)	(0.390)	
A Blood-type Patient	1.184	1.294	
	(0.160)	(0.265)	
B Blood-type Patient	1.077	0.635*	
	(0.172)	(0.158)	
AB Blood-type Donor	0.584**		1.249
	(0.150)		(1.357)
A Blood-type Donor	0.667***		1.562
	(0.0832)		(0.654)
B Blood-type Donor	0.608***		0.764
	(0.0957)		(0.520)
Constant	0.00578***	0.000892***	0.0656
	(0.00565)	(0.000203)	(0.234)
Observations	1,264	498	164

Note: Interval censored exponential hazard model. Patient (Donor) Match Power is the fraction of donors (patient) in the NKR pool over the course of a sample a given patient (donor) is compatible with. Sample restricted to patients and donors that registered after April 2012.

to transplantation opportunities elsewhere, either through direct donation, deceased donor transplants, or live-donor exchanges outside the NKR.

C.3 Compatibility and acceptance

To calculate whether donor j is compatible with patient i, we use the blood types of the patient and donor, the tissue type of the donor, and the list of unacceptable antigens listed by the patient. There are three additional ways in which a transplant between a patient and a donor can be prohibited.

First, upon registration, each patient can declare criteria for excluding donors based on a variety of characteristics. These include thresholds for the maximum donor age and minimum donor weight that are acceptable. These criteria are recorded in our dataset.

Second, upon arrival a patient can list as unacceptable any number of specific donors who were in the NKR pool at the time. This rejection can be done for any reason, including known pathologies. Patients can also exclude donors later, but according to our understanding, the practice is most common during registration. Our dataset includes the anonymized identifiers for these excluded donors.

Third, when a transplant is proposed, a patient may refuse the specific donor. A first phase of refusals is at the patient's discretion (with advise from his/her surgeon). If a patient chooses to proceed after the first phase, a final tissue-type compatibility test is conducted. We refer to this as the second phase.

In our simulations, we initialize $c_{ij} = 1$ if j is compatible with i and if j is not excluded by i. Otherwise, we set $c_{ij} = 0$. If j is offered to i during the simulation and a transplant is ruled out during either the first or second phases in the third type of exclusion, then we set $c_{ij} = 0$ for future "simulation days."

C.4 Burn-in and calculating standard errors

We may start our simulations from any initial state for the NKR because the effect of the chosen initial state fades over time. A convenient choice is to pick an initial pool with no unpaired patients, altruistic donors, or patient-donor pairs. Although the initial pool does not affect long-run averages with enough simulations, it is advisable to discard or burn-in a portion of the initial chain in order to improve the estimates' precision. A burn-in of about 2000 days appears to yield potential scale reduction factors for the number of transplants per day that is close to 1, suggesting that the chain is likely to have converged at that point.

Our simulations produce a series $y_1, \ldots, y_T$ of the transplants that occur on each day after an initial burn period. We estimate $f(\boldsymbol{q})$ as the sample mean of the y_t and calculate the standard errors of this estimate using the non-overlapping batch means estimator by following Chapter 12 in Robert and Casella (2004). The method divides the time series of y_t into batches, calculates the sample mean in each of those batches, and uses the variability in sample means to estimate the standard error of $f(\boldsymbol{q})$. We use the commonly recommended batch

size of approximately $\sqrt{T}$. The procedure is a simple and popular method that accounts for autocorrelation of the y_t.

C.5 Calibration

C.5.1 Calibration Procedure

Our simulation procedure is tailored to match the procedures and practices used by the NKR. In most cases, the data or institutional knowledge directly tell us the parameters; e.g., the weights w_{ij} are chosen to match NKR's practices. However, there are a few aspects of the real-world procedures and outcomes in the NKR on which we don't directly have information. We model and parametrize these aspects in our simulation model and calibrate them to match the realized number of matches in the NKR.

There are two main sets of parameters we need to calibrate. First, we do not have direct data on the frictions of translating proposed transplants into surgeries. As mentioned in Section C.3, the various acceptance phases may result in some transplants not being consummated. Each phase introduces a time-lag between transplants being proposed and finalized as well as the chance of a match being aborted. Roughly speaking, these frictions reduce the number of transplants facilitated by the NKR.

We parametrize these phases and calibrate the parameters to best fit the observed number of transplants by patient type. To do so, we simulate outcomes predicted by our model for various lengths (number of days) and various failure probabilities for both phases. The first phase parameters can be interpreted as controlling the frictions in the system because proposed matches are refused, whereas the second phase parameters govern the frictions due to biological compatibility tests.

Second, as mentioned earlier, when chains are aborted because of a refusal, NKR usually tries to use the donor, called the bridge donor, of the last transplanted patient for a new chain. However, the exchange prefers not to wait too long to start a chain with this donor. If a new chain cannot be found, the donor is offered to a patient without a related donor. Unfortunately, we do not know of a consistent policy rule followed by the NKR. We therefore also experimented with the number of days the NKR tries to match a bridge donor.

In summary, we calibrated five parameters: (i) the number of days a bridge donor can initiate a new chain, (ii) the number of days for consent and the probability of consent for the two phases of match acceptance, and (iii) the number of days taken for testing and the probability of failing the tissue type biological test.

This approach parsimoniously parametrizes the nature of frictions in the NKR. However, these models' single parameter versions are clearly a simplification. Most likely, NKR policies and their ability to translate proposed matches into transplants evolves over time and includes some ad hoc modifications to their basic procedure.

For our calibrations, we conduct our simulations by setting an initial market with the patients and donors who were present in the NKR on 1 April, 2012. This is the date from which we

have clear registration data. Then, for each parameter set, we run 100 simulations until December 2014, the last date of the available data.

We calibrated our parameters to match the transplantation probabilities, days in the exchange broken by whether or not the patient/donor was transplanted, the total stock of patients and donors in the exchange, and the total number of transplantations. The best fitted parameters were 14 days of waiting and 80% success rate for each of the two phases.

C.5.2 Calibration Results

Figures $C5$ - $C8$ compare the trends for these statistics across the data and the simulations for our chosen parameters. The dashed lines depict data, darker lines are the mean of 100 simulations, and dashed lines are 95% confidence interval from the simulations.

Green lines show the number of submissions of a given type (altruistic, pair, or unpaired), red lines show the cumulative number of submissions that departed the market without a transplant, and blue lines show cumulative number of submissions that were transplanted.

Figures $C5$ and $C6$ show that, at the calibrated parameters, the model fits the data extremely well. Although the averages are not as well matched in Figures $C8$ and $C7$, the observed quantities are within the 95% confidence intervals for the model.

Table $C3$ presents the summary statistics comparison between simulations and NKR data. Although we calibrated only five parameters, the table shows that the model matches several data moments extremely well.

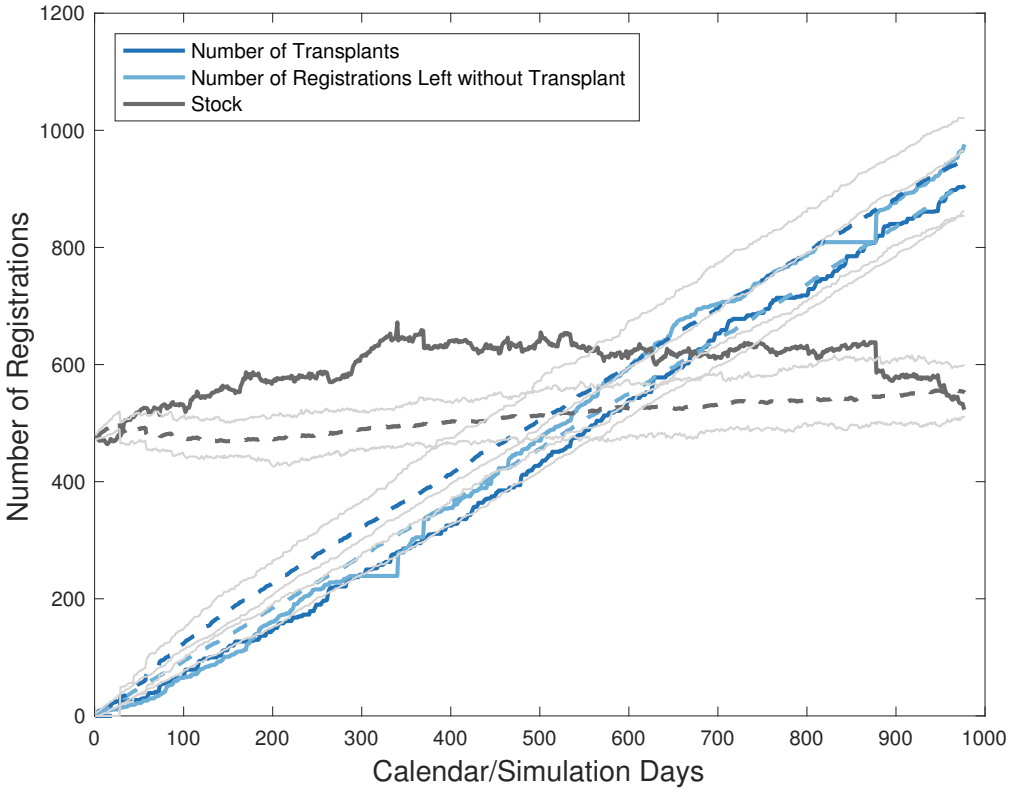


Figure C5: Calibration for All Submissions

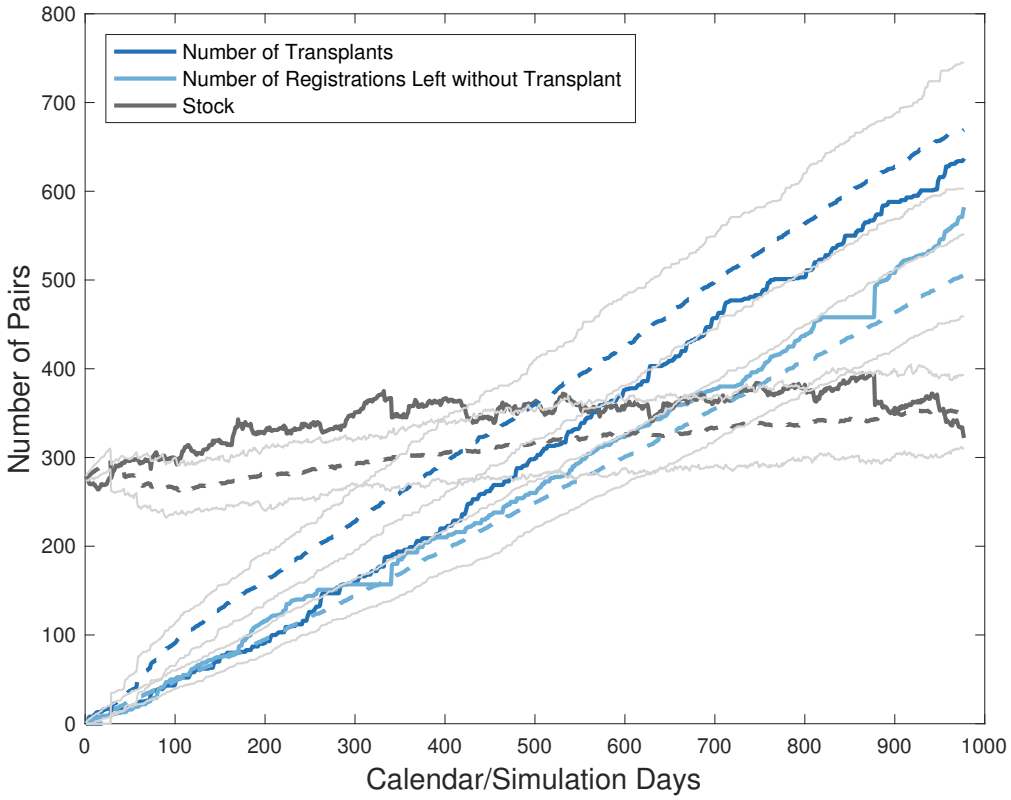


Figure C6: Calibration for Pairs

Note: Observed Quantities and Simulations are shown with solid and dashed lines respectively. Solid grey lines represent 95% confidence intervals.

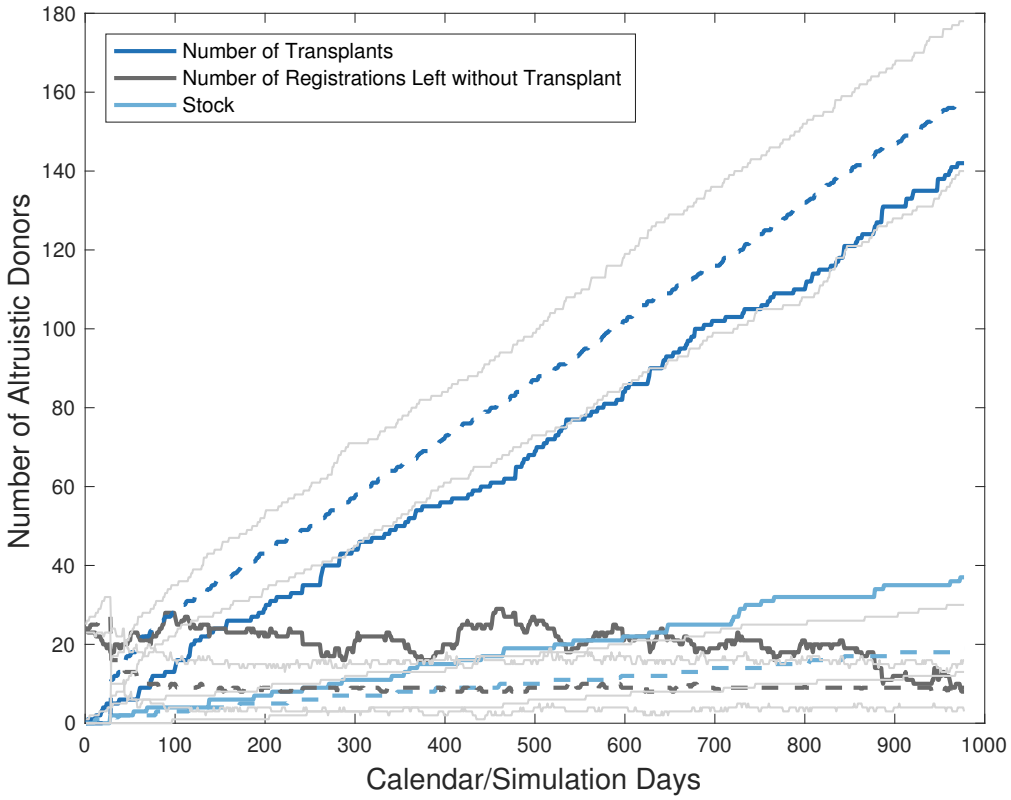


Figure C7: Calibration for Altruistic Donors

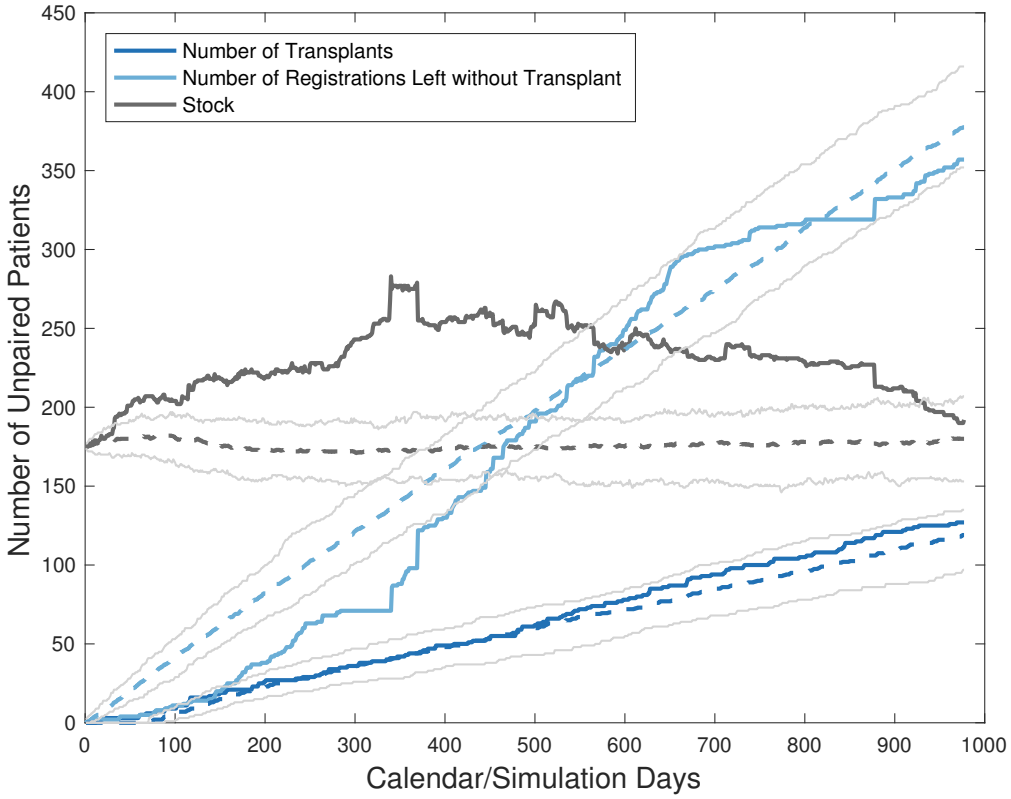


Figure C8: Calibration for Unpaired Patients

Note: Observed Quantities and Simulations are shown with solid and dashed lines respectively. Solid grey lines represent 95% confidence intervals.

Table C3: Model Fit

	Patient Match Probability	Donor Match Probability	Mean Days in the NKR for Transplanted Patients	Mean Days in the NKR for Untransplanted Patients	Mean Days in the NKR for Transplanted Donors	Mean Days in the NKR for Untransplanted Donors
Panel A: Observed Quantities from the National Kidney Registry (NKR)						
Pair	0.56	0.55	146.67	197.58	140.53	189.92
Chip Patients	0.30	-	63.85	135.40	-	-
Altruistic Donors	-	0.79	-	-	105.98	74.91
Under-Demanded	0.41	0.40	195.19	223.27	187.97	215.60
Over-Demanded	0.65	0.66	117.55	129.45	115.37	128.52
Self-Demanded	0.64	0.62	133.09	188.89	126.22	178.06
Panel B: Simulated Quantities						
Pair	0.60	0.57	99.85	183.35	120.00	179.42
Chip Patients	0.27	-	87.87	105.40	-	-
Altruistic Donors	-	0.89	-	-	50.00	47.72
Under-Demanded	0.38	0.36	164.01	211.06	190.64	209.50
Over-Demanded	0.71	0.69	69.62	135.61	78.82	132.20
Self-Demanded	0.73	0.69	84.87	153.87	106.51	148.88

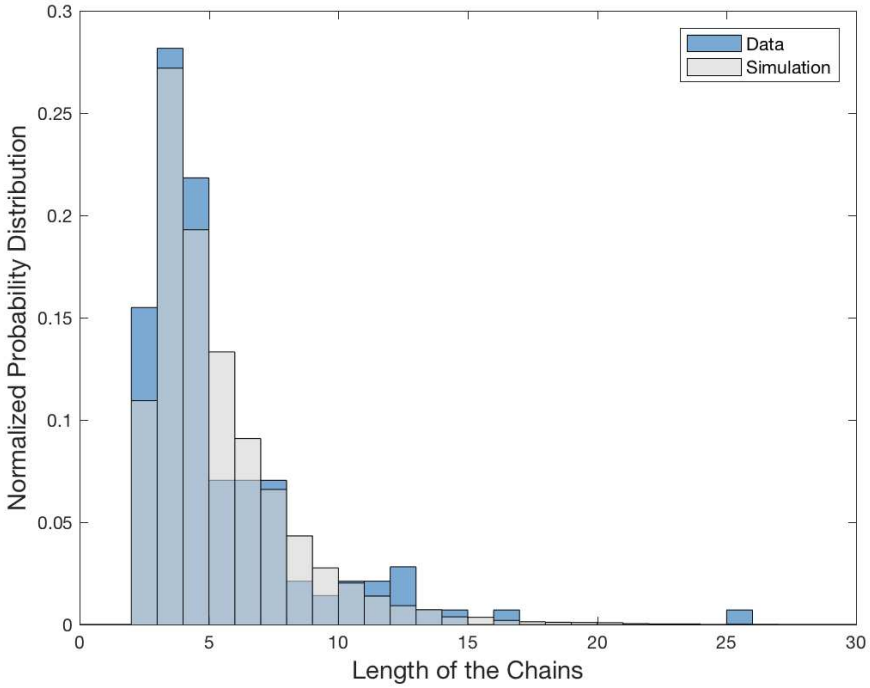


Figure C9: Chain Length Distribution

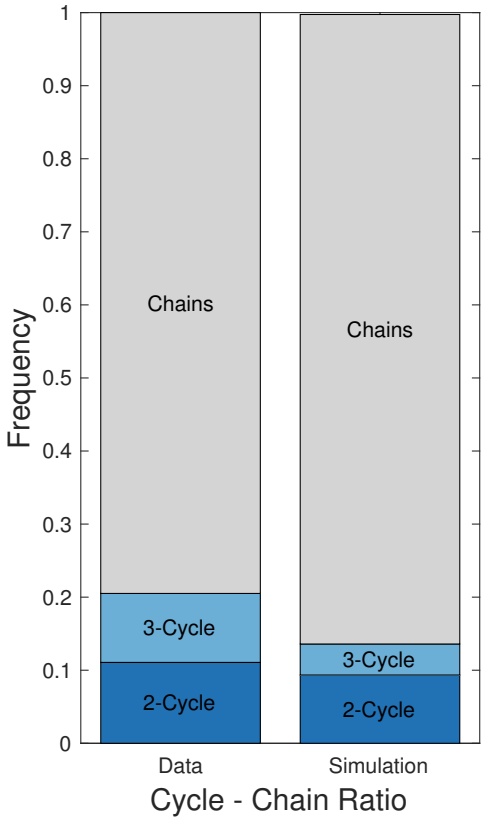


Figure C10: Cycle - Chain Ratio

Table C4: Regression Tree Summary Statistics

	N	Match Probability			Marginal Product		
		Mean	S.E.	Within Category Standard Deviation	Mean	S.E.	Within Category Standard Deviation
				Panel A: Altruistic Donors			
Non-O Donor	102	0.86	(0.01)	0.04	0.84	(0.05)	0.21
O Donor	62	0.94	(0.01)	0.00	1.86	(0.06)	0.17
				Panel B: Patient-Donor Pairs			
O Patient, Non-O Donor	493	0.28	(0.01)	0.16	0.08	(0.02)	0.19
O Patient, O Donor, PRA >= 91%	100	0.30	(0.01)	0.23	0.13	(0.05)	0.23
O Patient, O Donor, PRA < 91%	148	0.81	(0.01)	0.03	0.72	(0.04)	0.22
Non-O Patient, O Donor, PRA >= 96%	124	0.28	(0.01)	0.21	0.09	(0.05)	0.21
Non-O Patient, Non-O Donor, PRA < 96%	292	0.85	(0.01)	0.03	0.69	(0.03)	0.26
Non-O Patient, O Donor, PRA < 96%	108	0.84	(0.01)	0.01	1.44	(0.05)	0.30
				Panel C: Unpaired Patients			
Unpaired Patients	501	0.24	(0.01)	0.13	0.07	(0.02)	0.20

Notes: Categories are determined by regression tree analysis to predict marginal products as a function of whether a submission is a pair or altruist, 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 does not increase the overall fit by at least 2%. The resulting tree is depicted in Figure 8. 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).

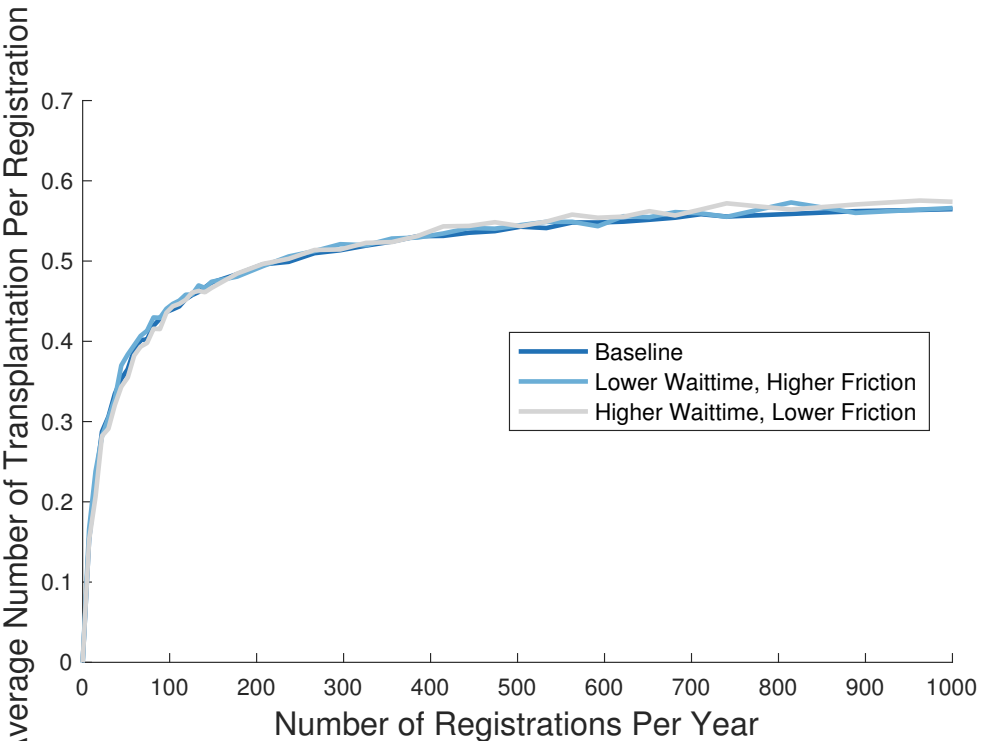


Figure D11: Robustness: Production Function versus Scale

Notes: Constructed as in figure 6.

D Robustness analyses

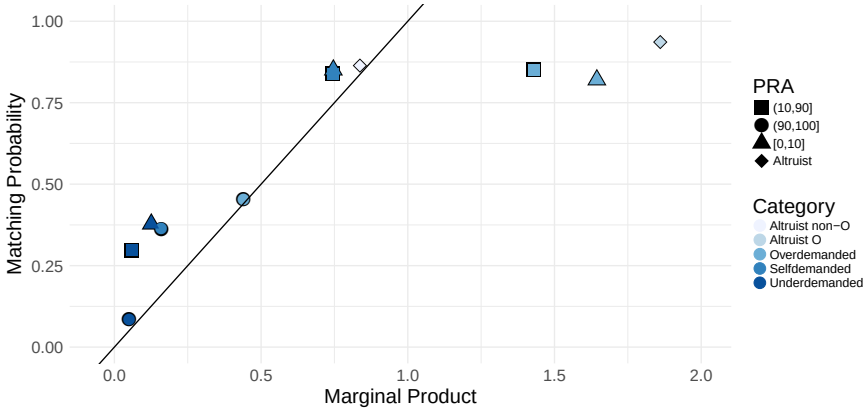
This section assesses the robustness of our results to calibrated parameters. Specifically, frictions in consummating proposed transplants due to longer waiting times but higher approval rates produce similar moments as lower waiting times and lower acceptance rates. Chain lengths, however, are increasing in acceptance rates and are best matched by our baseline parameters. We compare our baseline results with two substantially different parameters. The first, labelled "Higher Wait-time and Lower Frictions," has two weeks and three weeks for each of the two approval periods (approval and biological testing) but increases the acceptance rates in each phase from 80% to 85%. The other, labelled "Lower Wait-time and Higher Friction," uses three days and three weeks for each phase, respectively, but decreases the acceptance rates in each phase from 80% to 75%.

The qualitative and quantitative findings are robust to these alternative parameters. Figure D11 plots average products, as in Figure 6. These alternative parameters yield average product functions that closely follow the baseline. Table D5 shows the inefficiency estimates as in Table 3. The estimated inefficiency is within 5-10% of the baseline. Figure D12 shows marginal product versus matching probability of registrations aggregated by category, as in 7b. These results are also qualitatively similar. Table D6 shows marginal product, matching probability, and point system summary statistics, as in Table C4. Again, the points system under the alternative parameters are similar in magnitude.

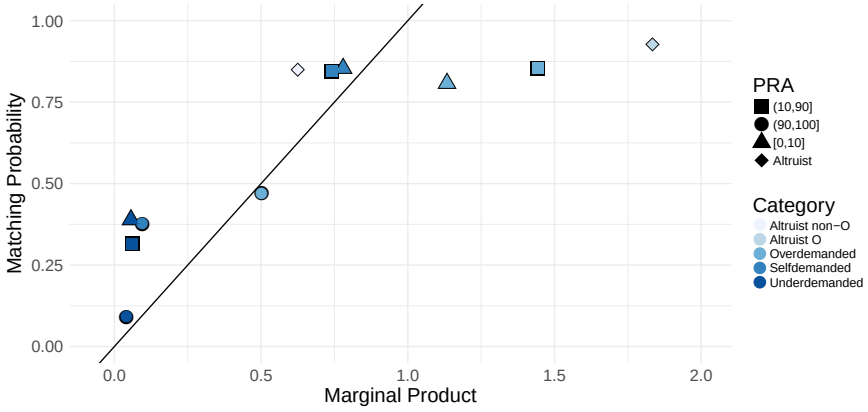
Table D5: Robustness: Total Efficiency Loss

	Number of Hospitals	Efficiency Loss: Additional Kidney Exchange Transplants		
		Base	Higher Waittime Lower Friction	Lower Waittime Higher Friction
		Panel A: All Hospitals		
All Hospitals	164	447.7	454.2	357.1
		Panel B: By hospital size (number of PKEs per year)		
Top Quartile	42	237.5	229.7	238.0
2nd Quartile	48	132.7	126.8	77.4
3rd Quartile	40	57.9	72.1	22.0
Bottom Quartile	34	19.7	25.7	19.7
		Panel C: By Platform Membership		
NKR	68	234.8	231.2	215.6
Only UNOS and APD	45	106.0	107.2	73.4
None	51	106.9	115.8	68.0
		Panel D: By NKR Participation Rate (Fraction of PKEs facilitated through the NKR)		
Top Quartile	17	14.5	15.8	9.0
2nd Quartile	17	44.2	44.2	28.7
3rd Quartile	17	81.5	80.9	78.3
Bottom Quartile	17	94.7	90.3	99.6

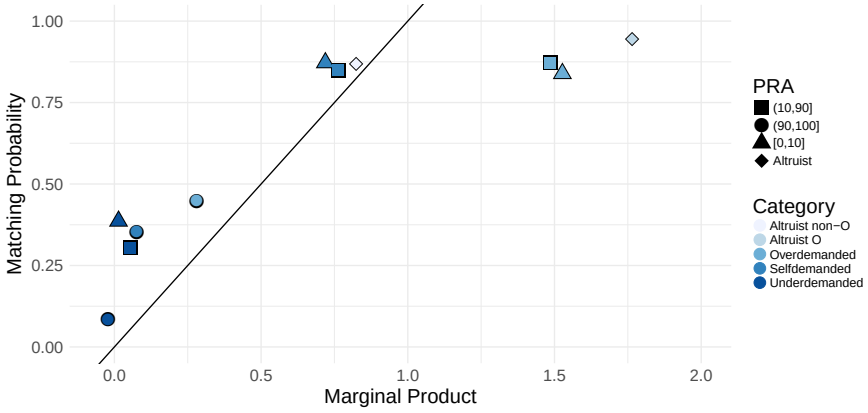
Notes: Constructed as in Table 3.



(a) Baseline



(b) Higher Wait-time Lower Friction



(c) Lower Wait-time Higher Friction

Figure D12: Robustness: Private versus Socially Optimal Rewards for Submission Types

Notes: Constructed as in Figure 7.

Table D6: Robustness: Points System

	Match Probability			Marginal Product			Points per Transplantation		
	Baseline	Higher Waittime Lower Friction	Lower Waittime Higher Friction	Baseline	Higher Waittime Lower Friction	Lower Waittime Higher Friction	Baseline	Higher Waittime Lower Friction	Lower Waittime Higher Friction
Panel A: Altruistic Donors									
Non-O Donor	0.86	0.85	0.87	0.84	0.62	0.82	0.97	0.74	0.95
O Donor	0.94	0.93	0.94	1.86	1.83	1.76	1.99	1.98	1.87
Panel B: Patient-Donor Pairs									
O Patient, Non-O Donor	0.27	0.29	0.28	0.08	0.03	0.00	-0.72	-0.89	-1.00
O Patient, O Donor, PRA >= 91%	0.29	0.30	0.28	0.13	0.10	0.06	-0.55	-0.67	-0.78
O Patient, O Donor, PRA < 91%	0.81	0.81	0.82	0.72	0.72	0.79	-0.11	-0.12	-0.04
Non-O Patient, O Donor, PRA >= 96%	0.28	0.29	0.27	0.09	0.08	0.01	-0.66	-0.73	-0.96
Non-O Patient, Non-O Donor, PRA < 96%	0.85	0.86	0.87	0.69	0.68	0.62	-0.19	-0.21	-0.28
Non-O Patient, O Donor, PRA < 96%	0.84	0.84	0.86	1.44	1.40	1.44	0.72	0.66	0.68
Panel C: Unpaired Patients									
Unpaired Patients	0.24	0.23	0.24	0.07	-0.03	0.02	-0.71	-1.12	-0.92

Notes: Constructed as in Table C4.

References

Anderson, Ross, Itai Ashlagi, David Gamarnik, and Alvin E Roth, "Finding long chains in kidney exchange using the traveling salesman problem.," *Proceedings of the National Academy of Sciences of the United States of America*, jan 2015, *112* (3), 663–8.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani, *The elements of statistical learning*, Vol. 1, Springer series in statistics New York, 2001.

Johnson, Donald B, "Efficient algorithms for shortest paths in sparse networks," *Journal of the ACM (JACM)*, 1977, *24* (1), 1–13.

Mas-Colell, Andreu, Michael Dennis Whinston, and Jerry R Green, *Microeconomic Theory*, Oxford University Press, 1995.

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

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