

Market Failure in Kidney Exchange*

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

This version: September 21, 2017

First version: September 20, 2017

Preliminary: please do not cite without permission

Abstract

Kidney exchange platforms enable transplants for about 800 patients with incompatible living donors. Using novel administrative datasets, we document that the market still shows clear signs of inefficiency because of fragmentation due to hospital participation behavior. We analyze these facts using a model that views an exchange platform as a producer of outputs (transplants) using inputs (donors and recipients) supplied by participants (hospitals). It identifies two market failures that can cause inefficiency: sub-optimal rewards for supplying inputs and agency problems in supply decisions. We then estimate a production function to provide quantitative results in the kidney exchange market. Our results show that the market produces about 400 fewer transplants than feasible because individual hospitals that conduct the majority of transplants are at an inefficiently low scale. While an optimal mechanism increases the number of transplants, it is also necessary to solve agency problems in hospital decisions to eliminate this inefficiency.

JEL: D47, D42, L11

Keywords: Kidney Exchange, Free Riding, Ramsey Pricing

*We are grateful to 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 for facilitating access to the data used in this study. We thank Glenn Ellison, Parag Pathak, Al Roth, Jean Tirole, Utku Ünver for extremely helpful discussions, and seminar participants at various universities for input. We received excellent research assistance from Abigail Ostriker. Azevedo acknowledges support from Wharton's Dean's Research Fund. Agarwal acknowledges support from the National Science Foundation (SES-1729090). Ashlagi acknowledges support from the National Science Foundation (SES-1254768). The data reported here have been supplied by UNOS as the contractor for the Scientific 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: Department of Economics, MIT and NBER, email: agarwaln@mit.edu. Ashlagi: Management Science and Engineering, Stanford University, email: iashlagi@stanford.edu. Azevedo: Business Economics and Public Policy, Wharton, University of Pennsylvania, email: eazevedo@wharton.upenn.edu. Featherstone: Business Economics and Public Policy, Wharton, University of Pennsylvania, email: claytonf@wharton.upenn.edu. Karaduman: Department of Economics, MIT, email: omerk@mit.edu.

1 Introduction

The success of platforms such as online auctions, health exchange marketplaces and ride-sharing hinges on bringing together many different participants in order to facilitate transactions. However, the overall market efficiency depends not only on the platform’s design, but also on incentives to transact elsewhere. In recent years, platforms for kidney exchange have grown to enable life-saving transplants for patients with living but incompatible donors. Yet, little is known about whether the market is operating efficiently and the extent to which improving the design and participation incentives can increase the total number of transplants.

Transplantation is not only the preferred treatment for end stage renal disease (ESRD) but also saves tax payers in the United States approximately \$270,000 per patient over five years due to savings from dialysis. As of 2017, however, more than 97,000 people are waiting for a deceased donor kidney in the United States. The shortage of organs led to an increase of living donor transplants, but not every willing live donor is compatible with her intended recipient. This creates demand for *kidney exchange*, in which two or more incompatible patient-donor pairs exchange kidneys.

A few national kidney exchange platforms emerged in the U.S during the last decade that try to create a thick marketplace and leverage gains from scale (Roth et al. 2004). But participation is not mandatory, and hospitals, upon which patients typically rely on, may decide not to engage in kidney exchange or match some of their patient-donor pairs internally creating even further competition (Ashlagi and Roth 2014). Overall, kidney exchanges in the U.S. accounts for about 800 annual transplants in recent years. The goal of this paper is to evaluate the inefficiency of the kidney exchange market in the U.S and understand how to improve it.

The paper contributions are organized as follows. First, we assemble a new dataset and document the state and evolution of kidney exchange in the U.S. including fragmentation and hospital behavior. Second, we develop a price-theoretic model based on neoclassical producer theory to explain how inefficiency arises in equilibrium and present a solution. Finally, we use data from the largest national platform to estimate the key parameters of the model in order to quantify inefficiency and its sources in order to design better mechanisms.

For the first part we combine data of all transplants in the U.S. with proprietary datasets from the largest kidney exchange platform. These data reveal various signs of inefficiency. Rather than a few large platforms, the market is highly fragmented with 65% of exchange transplants coming from internal matches conducted by hospitals. Further, patients transplanted through internal exchanges are easier to match (lower sensitized) than patients matched through national platforms. Participation in national programs varies across hospitals; consistent with fixed costs, smaller hospitals are less likely to participate in national programs than larger hospitals. The hospitals that do participate don’t perform all their kidney exchanges through the national platform and register particularly hard-to-match patients and donors at the platform. Importantly, this behavior is associated with a significantly lower match efficiency within hospitals exchanges when compared with matches through national platforms. For

example within hospital exchanges match a much larger fraction of O donors with non-O patients (see Roth et al. (2007) for efficient allocations at large scale).

Motivated by these descriptive findings we develop a simple theoretical framework based on neoclassical producer theory and economics of platforms to explain the sources inefficiencies and generate better solutions. We model a kidney exchange clearinghouse as a platform to which hospitals submit patients and donors and are rewarded with transplants. The model unifies and build on observations from the kidney exchange literature (Roth et al. 2007; Ashlagi and Roth 2014). A key component of the platform is a production function f that determines the transplants that can be done given the quantity and composition of the pool. This builds on an observation by Roth et al. (2007) that the potential number of transplants is determined by biological compatibility. They analyze a production function assuming a large market, which suggests that some patient-donor types generate 2 additional transplants upon joining a platform, while other types do not generate any additional transplants but rather compete with other pairs for being matched. We depart from Roth et al. (2007) by allowing for a general production function that need not always obey their large market properties. Although qualitatively similar, our estimates of the production function based on the administrative data from the NKR will show that this departure is quantitatively important for designing an optimal mechanism.

We further assume that hospitals are the decision-makers for whether to submit patients and donors to the platform. This follows observations by Sönmez and Ünver (2013) and Ashlagi and Roth (2014) on the incentives to withhold pairs for within hospital transplants and is consistent with our descriptive analysis. Hospitals in our model have quasilinear preferences over their submissions and the number of transplants they receive from the platform. To capture agency problems, we also allow for hospital welfare to deviate from our preferred utilitarian welfare measure, which is to maximize the total number of transplants.

We analyze the model by considering rewards that maximize aggregate hospital welfare. To do so, a platform must reward hospitals by assigning them transplants according to their contribution. Specifically, let \mathbf{p} be the vector of rewards that the platform provides for different types of submissions. Theorem 1, which is a version of Ramsey’s optimal linear commodity taxation formula, shows that the optimal rewards \mathbf{p} are approximately equal to marginal product of each pair. Intuitively, this rewards structure encourages hospitals to submit types of pairs that allow for a greater number of transplants to occur. At the same time, as opposed to mandatory participation, it does not distort incentives if some hospitals are much more productive when matching internally due to technological differences.

Our results suggest that mechanisms currently used in national platforms are, most likely, sub-optimal. The current rewards system is implicitly determined by the maximal matching algorithms so that \mathbf{p} is equal to the probabilities of matching different types of submissions through the exchange. These algorithms are not designed with the view of encouraging transplants centers to submit types that increase transplants to the exchange. That is, current rewards are based on a completely different quantity than optimal rewards. In principle, this kind of problem could be solved with simple “point mechanisms”, where centers keep track of a point balance and points are added or deducted upon transplants based on marginal products.

Another source of inefficiency in our model is agency problems. Agency problems happen if hospitals do not fully internalize the welfare of the parties that they present (Jensen and Meckling 1976). Agency problems may arise in this market because hospitals currently bear a significant portion of costs associated with participation in kidney exchange platforms (Rees et al. 2012). The social value of transplants, however, are likely small relative to these costs. Nonetheless, hospital financial incentives can be skewed away from participation in national platforms if these costs are not reimbursed.

Our model therefore suggests that solving both these sources of inefficiency can be important for increasing the total number of transplants conducted through kidney exchange. Indeed, if the production function exhibits constant returns to scale and there are no agency problems, then the market will attain the first-best outcome under an optimal rewards mechanisms. However, this model does not quantify the importance of each of these sources of inefficiency and provide quantitative numbers for the rewards that should be implemented.

The third part of the paper turns to these issues. We return to the data to quantify the inefficiency due to the channels described above and to design a better mechanism. To do so we estimate the key primitive of the model, the production function, using administrative data from the largest national kidney exchange program in the U.S., the National Kidney Registry (NKR). Instead of traditional techniques for estimating production functions suitable for low-dimensional data on observed inputs and outputs (Marschak and Andrews 1944; Olley and Pakes 1996), we build a production function using detailed knowledge of the logistics and algorithms involved in operating kidney exchange platforms.¹ Specifically, we set up a detailed simulation procedure for the various stages in organizing an exchange, from initial registration to final transplantation including intermediate frictions due to biological testing. This procedure allows us to estimate the number of transplants per year that would be produced for any combination of inputs. The model requires only a handful of parameters to be calibrated and does very well in replicating the observed outcomes in the NKR data.

The estimated production function yields two striking results. First, we find that the major kidney exchange platforms are at approximately efficient scale and far beyond the point where returns to scale are rapidly increasing. At the same time, almost all single-hospital platforms far below the efficient scale. This difference in scale suggests considerable inefficiencies because single-hospital platforms perform approximately 65% of all live-donor exchanges. It is also possible to calculate the number of transplants lost as a result of single-center hospitals operating at an inefficiently small scale with a given production function and composition of donors and patients for single-hospital platforms. The most conservative estimates across a range of approaches suggest that at least 200 transplants a year are lost due to single-center hospitals operating at inefficient scale. Our central estimates place the inefficiency at twice that number. Thus, consistent with the descriptive evidence and the shape of the production function, fragmentation in the market appears to have large efficiency consequences. Further, because returns to scale are roughly constant at the scale at which the major kidney exchanges operate, further gains due to scale economies alone from merging these exchanges are likely small. This observation on the scale economies can also explain why, instead of tip-

¹Ashlagi has been working closely with multiple platforms including the NKR.

ping to a single platform, multiple large national exchanges co-exist (Ellison and Fudenberg 2003).

Second, we find that the marginal product of different kinds of submissions are extremely varied and considerably different than probabilities of matching. This result explains several descriptive facts documented in the first part of the paper and presents a route to designing a better mechanism. For example, it shows that the difficult to match pairs have marginal products that are much lower than probabilities of matching whereas valuable pairs have lower match probabilities. This is consistent with evidence on the selection of types into the national exchanges. It also suggests the level of rewards implemented through a point system that can improve the participation incentives provided by the platform. Indeed, simple characteristics of submissions, such as patient-donor blood types, patient sensitization and whether a donor is non-directed explain most of the variation in marginal products. Therefore, very simple and practically implementable point mechanisms that depend only on these characteristics would be considerably closer to optimal than the current system.

Finally, we calculate the gain in (hospital) welfare if the exchange could adopt such a mechanism and explore the relative importance of agency problems and the sub-optimal platform incentives. Our findings suggest that while an optimal market mechanism would increase efficiency the market would remain inefficient without also resolving agency problems. We further provide initial supporting evidence for inefficiency due to agency problems.

To summarize, this paper finds that the U.S. kidney exchange market is highly fragmented and identifies two major sources of inefficiencies: platform incentives and agency problems. Moreover we find that while simple point mechanisms would increase efficiency by fixing platform incentives, to attain full efficiency agency problems should be fixed as well.

2 Background and Data

2.1 Institutional background

Importance of kidney transplantation

Medicare spends approximately 7% of its annual budget on patients with kidney failure, with the vast majority of these expenses used to cover dialysis costs. Transplantation is a significantly preferable option because it increases life expectancy by 10 years on average and is cheaper than maintaining a patient on dialysis (Wolfe et al. 1999). A transplant conducted for a Medicare beneficiary saves the government more than \$270,000 in present value (Wolfe et al. 1999; Held et al. 2016). These numbers are larger for privately insured patients (Irwin et al. 2012). In addition, patients derive significant value from the improved life-expectancy due to a transplant. Estimates place the social value of a transplant based on estimates of cost differences between dialysis and transplantation, and the value of life at about \$1.1 million (Held et al. 2016).² Unfortunately, the scarce supply of deceased donor kidneys results in

²Held et al. (2016) conduct a comprehensive cost-benefit analysis with sensitivity to a range of assumptions. They place a value of a year of perfect health at \$200,000 and adjust for differences between the quality of

a growing waiting list, which currently is almost 100,000 patients long. The average waiting time for a deceased donor is now between 3-5 years while a patient can expect to live on dialysis only about five years.

An alternative to a deceased donor transplant is to use one of the two kidneys from a healthy live donor to conduct a transplant. The importance of living donor transplantation has grown because of the scarcity of deceased donor organs and because they result in more life-years for the recipient.. Therefore, the best outcome for a patient with kidney failure is if she can find a suitable living donor..

Basics of kidney exchange

Kidney exchange is a solution for the many patients with willing and healthy donors that cannot directly receive a transplant from this donor because of biological incompatibility. It involves two or more incompatible patient-donor pairs, with each patient in the exchange receiving a compatible kidney from another patient's donor.³ Such exchanges are a natural method for overcoming the double coincidence of wants given the prohibition against financially compensating donors (Roth et al. 2005a).⁴

A kidney exchange typically take one of two forms. The first, called a **cycle**, involves a set of incompatible pairs and transplants the donor from one of the pairs to the patient in the next pair until the loop is closed. All transplants are carried out simultaneously to minimize the risk that a donor donates a kidney without her intended recipient receiving one. This form is usually limited to at most three pairs due to logistical difficulties in coordinating several simultaneous surgeries.⁵ The second common form, called a **chain**, is initiated by an altruistic or non-directed donor donating to a patient in an incompatible pair. The donor from this pair can continue the chain by donating to the next pair and so on until the chain terminates in a patient who does not have a willing live donor. Because transplants in chains can be done non-simultaneously, they can involve dozens of incompatible pairs⁶ but usually involve only four to five transplants.

A donor must be both **blood-type and tissue-type compatible** with the patient for the transplant to take place. For a donor to be blood-type (ABO) compatible with a patient, she should not have a blood protein that the patient lacks. A and B represent the two types of proteins that can be found in the blood, and O represents an absence of both proteins.

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).

³The first kidney exchange was in Korea (Kwak et al. 1999) between two blood type incompatible pairs.

⁴The National Organ Transplant Act (NOTA 1984) makes it illegal to obtain organs for transplantation by compensating donors.

⁵In the early days kidney exchange was limited to pairwise exchanges and was expanded to cycles of length three because of the gains from doing so (Saidman et al. 2006).

⁶See Rees et al. (2009) for the first long non-simultaneous chain

So, an O donor is quite valuable because it is ABO compatible with a patient irrespective of her blood-type. An A donor, on the other hand, can only be compatible with blood-type A or AB patients (Danovitch 2009). For a donor to be tissue-type compatible, she should not have certain tissue proteins against which the patient has an immune response.⁷ The predominant measure for a patient’s immune sensitivity is the **Panel Reactive Antibody (PRA)**: it is the likelihood that the patient is tissue-type incompatible with a representative donor.

Organization of kidney exchange platforms and financial considerations

In its early days, kidney exchange was organized by single hospitals or by a small group of hospitals (Roth et al. 2005a). Slowly, several large kidney exchange platforms emerged in an attempt to systematically find possibilities for kidney exchange that are not available at the small scales in which hospitals operate. These platforms vary in geographical scope, size and logistics of how exchanges are organized. National platforms such as the Alliance for Paired Donation (APD), the United Network for Organ Sharing (UNOS), and the National Kidney Registry (NKR), which is the largest platform, involve the participation of many hospitals.⁸ Others include much smaller regional networks or single hospitals like Methodist Hospital in San Antonio. Most large platforms identify exchanges using optimization software to maximize weighted number of transplants, but differ on the precise weights and other operational details.⁹ In addition to large organized platforms, many smaller hospitals internally match their patient-donor pairs with each other when feasible. These exchanges are usually organized on an ad-hoc basis.

Kidney exchange, especially when conducted through multi-hospital platforms, presents several financial obstacles for hospitals. Rees et al. (2012) and Wall et al. (2017) discuss the costs of conducting kidney exchange that are not reimbursed to hospitals within the current

⁷Broadly speaking, the immune system tags foreign objects (antigens) with antigen-specific antibodies so that white blood cells (leukocytes) can defend against them. Hence, if we put 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 immune system will attack the kidney, leading to rejection. A recipient is tissue-type compatible with a donor kidney if she has no antibodies corresponding the antigens of the donor kidney (Danovitch 2009). Note that a transplant between certain incompatible patient-donor has become possible due to development of desensitization technologies (see, e.g., Orandi et al. (2014)).

⁸Historically, live donation mostly involved donors with an immediate relationship to the patient, obviating the need for any larger scale coordination. Over time, however, the practice of kidney paired donation grew in prevalence, culminating in the passage of the Charlie W. Norwood Living Organ Donation Act, which explicitly states that paired kidney donation is not forbidden by NOTA’s provision outlawing the transfer of an organ in exchange for “valuable consideration”. This legal clarification allowed for the formation of private clearinghouses for matching incompatible donor-patient pairs and altruistic donors. Large clearinghouses include the Alliance for Paired Donation, founded in 2006, and the National Kidney Registry, founded in 2007. More recently, in 2010, UNOS formed its own clearinghouse, called the OPTN KPD Pilot Program.

⁹Abraham et al. (2007) design an algorithm to handle large populations that is being used by United Network for Organ Sharing (UNOS) exchange program. See also Ashlagi et al. (2016); Anderson et al. (2014); Dickerson et al. (2012) for related work on matching policies.

payment system. Many of these costs are associated with extensive biological testing required for evaluating potential donors for a patient. These issues are further complicated if the donor belongs to a different hospital than the patient. In addition, fees charged by large platforms such as the NKR add to the costs borne by a hospital. In addition to approximately \$4,000 charged for each transplant and costs of shipping a kidney, the NKR also charges annual fees of about \$10,000 to each participating hospital (National Kidney Registry 2016). In addition, hospitals participating in platforms may need to hire dedicated transplant co-ordinators to deal with the additional logistical issues involved in multi-hospital kidney exchanges. These costs can be a significant portion of variable profits from a transplant and result in financial barriers to kidney exchange. In particular, they discourage participation in multi-hospital platforms. Hospitals receive about \$100,000 to \$160,000 for conducting a typical kidney transplant (Held et al. 2016; USRDS, United States Renal Data System 2013). This amount includes payments for hiring surgery teams, drugs, equipment and capital. Indeed, surveys point to financial considerations coming in the way of participation in kidney exchange platforms (ASTS, American Society of Transplant Surgeons 2016). These costs are small relative to the costs saved on dialysis treatments, let alone the value of improved quality and length of life post-transplant. Similar reasons have motivated practitioners to propose improvements in the financing model for kidney exchange (Rees et al. 2012).

2.2 Data

To track the kidney exchange market in the US, we start with de-identified administrative datasets from the largest American clearinghouse: the National Kidney Registration (NKR). These data contain essentially all the information used as inputs in the matching algorithm in addition to information used to manage the logistics of kidney exchange at the platform. It contains rich information on the patients and donors registered at the platform, including their age, gender, blood-type and antigens. Each patient and donor is associated with a home hospital, a registration data. The related incompatible donor for each patient that registers as a pair is also recorded. In addition, we also know the unacceptable antigens for each patient and the criteria of acceptable donor quality (age, weight, height, medical conditions etc.). Given this information, we can very accurately determine the set of feasible transplants. The data also records the transplant date, the donors and patients involved in the transplants, the chain or cycle configuration and the hospitals where the transplants were conducted. We obtained records on all transplanted patients from the NKR until December 2014. Data on patients and donors that were registered but not transplanted through the NKR is only available after April, 2014.

The platform data alone has several limitations. First, it tells us nothing about exchanges that are not organized through the NKR. Second, it tells us nothing about what happens to pairs that aren't matched by the clearinghouse. Fortunately, these problems can be remedied by combining the data from the platforms with administrative records collected by United Network for Organ Sharing (UNOS), a non-profit federal contractor with administrative responsibilities over all organ transplants in the US.¹⁰ Specifically, UNOS's Standard Transplant

¹⁰The National Organ Transplant Act of 1984 (NOTA) established the Organ Procurement and Trans-

Analysis and Research (STAR) dataset contains a record for each transplant done in the US, as well as records for a patient’s registration status on the deceased donor waitlist.¹¹ These records contain most of the fields about individual patients and donors that are included in the clearinghouse datasets.¹² The UNOS data we use in this paper contains all kidney transplants from February 2008 until September 2015, as well as any waitlist registration that was active during that window.

The one limitation of the STAR dataset for our purposes is that it does not contain information about which patients and donors form incompatible pairs as waitlist registrations do not mention whether the candidate is considering kidney exchange. However, for each transplant, the data contains not only the information about the the patient and the donor, but also indicates f whether the organ came from a live donor and if sowhether the transplant was an exchange or a direct transplant.

To leverage the pair information in the clearinghouse data, it must be linked to the STAR data. Since both the clearinghouse and STAR datasets are anonymized separately before they were given to us, this is non-trivial. We summarize our linking procedure here. Because the STAR dataset contains the universe of transplants, it is straightforward to find the STAR transplant that is closest on dimensions of transplant date, transplant center, and patient and donor age, gender, blood type, and HLA type. Ultimately, we were able to match around 95% of clearinghouse transplants in this way to a very high degree of certainty.¹³ For donors and patients who were never transplanted by the clearinghouse, the matching process is more difficult because it is possible that they won’t match to any STAR record at all.¹⁴ We deal with this problem by implementing a naïve Bayes classifier, which outputs a probability that a match is genuine, and setting a cutoff (Russell and Norvig 2003). These procedures create “synthetic matches” based on patient and donor age, blood-types, tissue-types and transplant center and date information. Therefore, they do not involve identifying any individuals.

3 Descriptive Evidence

The datasets assembled for this study allow us to take a fresh look at the avenues through which patients with a willing but incompatible live donor obtain a transplant. Section 2 described the growing number of kidney exchange platforms and the financial considerations for hospitals involved in kidney exchange. This section starts by documenting the fragmentation

plantation Network (OPTN). Since 1986, UNOS has served as the OPTN systems provider under contract with the Health Resources and Services Administration (HRSA) of the Department of Health and Human Services (DHHS).

¹¹The STAR dataset should be close to the universe of such observations because UNOS runs the deceased donor waitlist, and US transplant centers are required by law to submit information to UNOS about each completed transplant. (OPTN 2017)

¹²The standard analysis files from the STAR data do not contain information on acceptable antigens, home centers, and transplant centers. These fields are available from UNOS on request.

¹³In fact, 85% of the matches were a perfect match along the dimensions just described.

¹⁴Donors only appear in STAR if they ultimately donate an organ, while patients only show up if they receive an organ or if they register for the deceased donor waiting list.

in the market. We then describe hospital participation decisions that are consistent with financial incentives limiting full participation in the National Kidney Registry (NKR). Next, we show that the set of patient-donor pairs registered by hospitals are hard-to-match. This selective registration in the NKR occurs concurrently with many inefficient transactions conducted through ad-hoc within-hospital kidney exchanges. These facts will motivate a model with two sources of market failure, the first due to improper financial incentives and the second due to a design of the NKR that does not adequately reward easy-to-match patients and donors.

3.1 Growth and fragmentation

The goal of kidney exchange platforms is to maximize total transplants while respecting constraints due to biological compatibility and logistical issues. In the U.S., the total number of transplants through kidney exchange grew rapidly from about 400 in 2008 to about 800 transplants in 2014, but this growth has recently slowed down (Figure 1).¹⁵ The growth is concurrent with the growing importance of co-operative, cross hospital platforms. In 2008, almost all kidney exchange transplants involved a patient and a donor that belonged to the same hospital. We will refer to such kidney exchange transplants as **within hospital** exchanges. By the end of our sample period, about half of the transplants involved donors and recipients from different hospitals.¹⁶ Indeed, most exchanges across hospitals were facilitated by the National Kidney Registry (NKR), which has rapidly grown to become the largest kidney exchange platform in the U.S. Towards the end of our sample, it accounts for about a third of all kidney exchange transplants in the U.S..

Despite NKR’s growth, the figure shows that the market remains fragmented. Many hospitals in the U.S. continue to conduct ad-hoc within hospital kidney exchanges amongst their patients and donors. Indeed, about 40% of transplant centers that conduct kidney exchanges do not participate in the NKR and the total number of transplants from due to within hospital exchanges hasn’t substantially diminished.

Such fragmentation can undermine the number of transplants conducted nationwide. Larger platforms and those with easy-to-match patients and donors have more ways in which they can arrange cycles and chains that involve a given patient or donor. Coordinating on a single large kidney exchange platform in which all incompatible patient-donor pairs and altruistic donors are enrolled could help increase the total number of transplants. Potentials costs due to additional logistical challenges are likely small in comparison to the value of additional life-years afforded and costs savings discussed in Section 2.1. However, one potential drawback of coordinating on a single platform is that competition across platforms may allow for continued experimentation and innovation in the design of kidney exchange platforms.¹⁷

¹⁵Our data for the NKR extend until December 1, 2014. This censoring may account for the slight drop in transplants in the last year of this figure.

¹⁶Because a donor’s kidney is typically transported to the patient’s hospital after extraction, we use the hospital in which the donor’s kidney was extracted as the donor’s hospital.

¹⁷The Alliance for Paired Donation (APD), for example, has recently started organizing kidney exchange

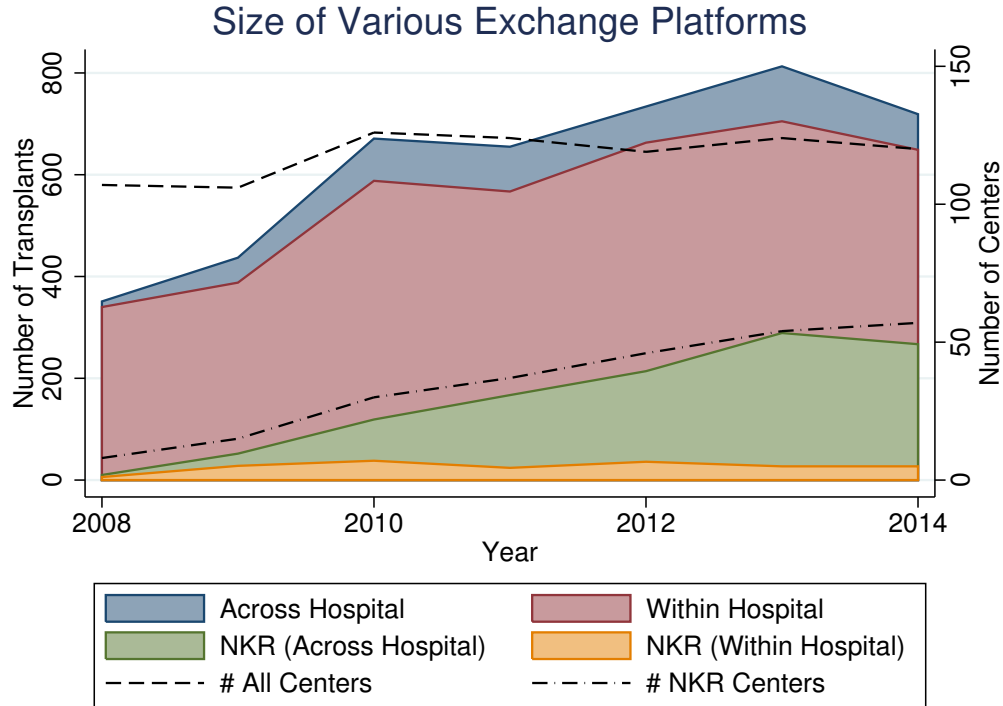


Figure 1: Market fragmentation and trends in live-donor kidney exchange

3.2 Hospital participation decisions

Hospitals and surgeons play an important role in steering patients and donors to a national platforms or in attempting within hospital exchanges with their patients and donors.¹⁸ Anecdotes and institutional features discussed in Section 2.1 suggest that many hospitals find the logistics involved in coordinating with large platforms demanding on their limited manpower. Hospitals may also prefer within hospital exchanges for their patients because there is a chance that not all the patients that they can match themselves will be matched by the national platform (Sönmez and Ünver 2013; Ashlagi and Roth 2014). Other hospitals have decided, as a matter of policy, to enroll all their patients and donors in a large platform. While hospital behavior is heterogeneous, these anecdotes suggest that many hospitals narrowly consider private costs and benefits, either in the interest of their patients or to maximize profits.

Figure 2 shows a bin-scatter plot of the fraction of hospitals that conduct some exchanges

on a global scale to facilitate more transplants in ways that benefits both U.S. patients in need of a kidney and patients in the developing world that cannot afford the costs of a kidney transplants Nikzad et al. (see 2017, for details).

¹⁸In other countries such as the U.K. (Johnson et al. 2008), the Netherlands (De Klerk et al. 2005), and Canada (Malik and Cole 2014), there is a single national exchange program and participation is mandatory. However, similar challenges may raise when such countries attempt to merge their pools as attempted currently by Israel and Cyprus (Siegel-Itzkovich 2017).

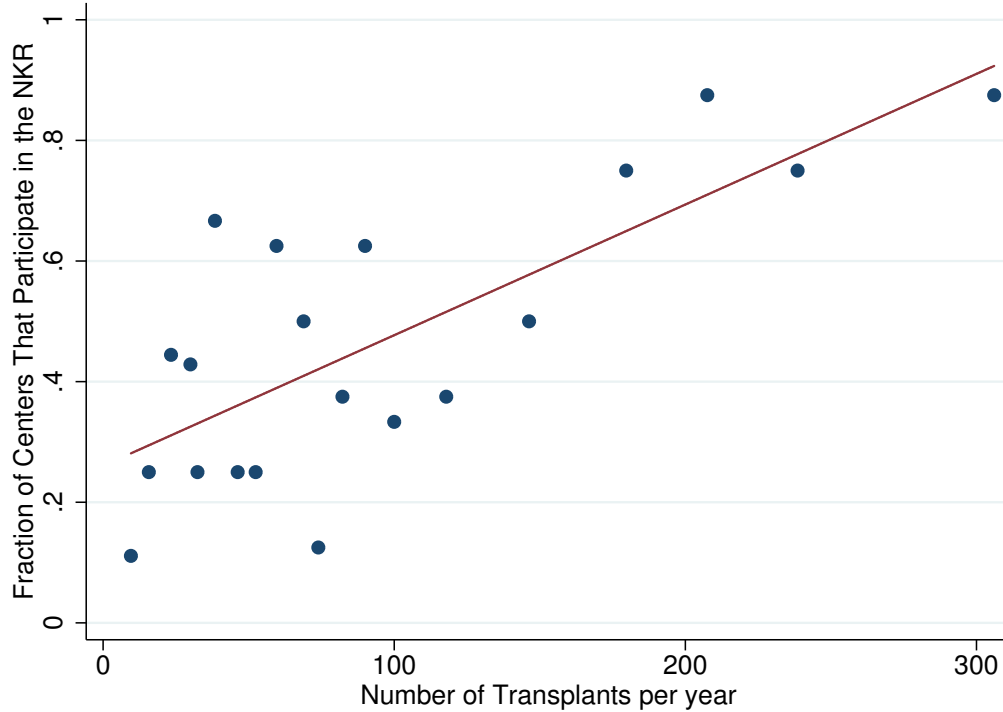


Figure 2: Heterogeneity in participation in the NKR

through the NKR against size. We measure hospital size by the number of transplants conducted per year, including deceased donor transplants to limit an endogenous increase in this measure due to participation in the NKR. The figure shows that although the NKR consists of a mix of small and large hospitals, smaller hospitals are much less likely to participate in the NKR. This is consistent with participation barriers arising from fixed costs of co-ordinating kidney exchanges with a national platform, such as costs of hiring a full-time kidney exchange co-ordinator. Larger hospitals are more likely to be able to recover fixed costs because they are spread over more transplants. Section 2.1 describes some of the costs of participating in kidney exchange platforms that can be a barrier to participation. These include membership costs, costs of increased blood and tissue-type tests that may not be fully reimbursed and possibly hiring additional staff to coordinate with a platform. However, these costs are likely small in comparison to the value of additional transplants. Even the additional cost-savings from one additional transplant can be expected to cover hiring costs of a transplant coordinator, the membership costs and other increases due to blood-tests. These savings are in addition to the large value in terms of the additional life-years afforded to a transplanted patient. Nonetheless, these costs can limit participation in national platforms because hospitals are only paid a fraction of the value for each transplant.

In addition to incomplete participation on the extensive margin, Figure 3 shows that even hospitals that participate in the NKR often conduct within hospital kidney exchanges. It presents a scatter plot of the fraction of kidney exchanges at a hospital that are facilitated by

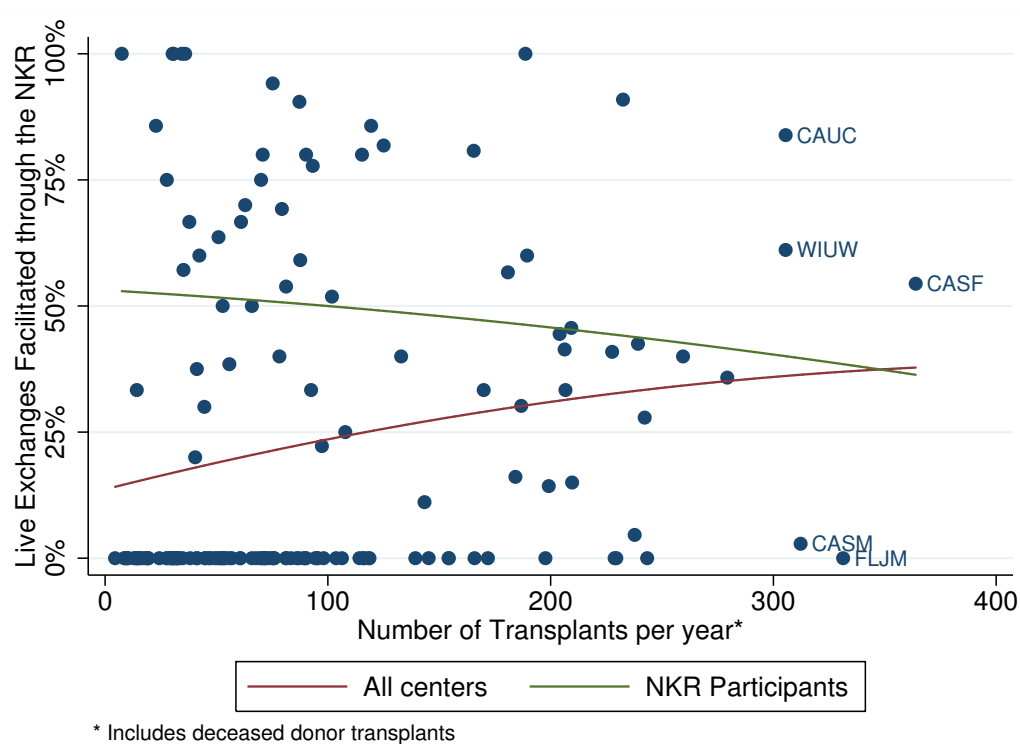


Figure 3: Reliance on the NKR for live-donor exchanges

the NKR. The curves represent the best quadratic fit for the points, estimated for all hospitals and separately for hospitals that participate in the NKR. The plot also shows remarkable heterogeneity both across and within hospital size in the reliance of hospital on the NKR for kidney exchange. Although most small hospitals do not participate in the NKR, most of of small hospitals that participate in the NKR rely on the platform almost exclusively to organize the kidney exchanges for them. Because these hospitals are small, it is not surprising that they will not be particularly successful at organizing exchanges with the few patients and donors they are responsible for without cooperating with other hospitals via an exchange. While some larger hospitals almost exclusively rely on the NKR to organize their kidney exchanges, many conduct kidney exchanges outside the NKR, including within hospital exchanges. These hospitals are likely making decisions about which patients and donors to enroll in the NKR.

3.3 Selective registration and inefficient matching

The nature of blood-type and tissue-type compatibility implies that certain patient-donor pairs are easier to match than others. For example, patients with blood-type O or high PRA are more likely to be incompatible with a related live donor and therefore more likely to need an exchange. On the other hand, blood-type O donors are likely to have directly donated to their intended recipient. This makes blood-type O donors relatively scarce and blood-type

O patients relatively abundant. More generally, as Roth et al. (2007) argue, one should expect that exchange pools have more overdemanded pairs than underdemanded pairs; an **overdemanded pair** has a donor who is blood-type incompatible with her related patient and **underdemanded pair** has a patient who is blood-type compatible with her related donor. Similarly, it is difficult to match patients with **high Panel Reactive Antibody (PRA)** that have immune systems more likely to react with a randomly chosen donor’s tissue-type (Ashlagi et al. 2012).

Table 1 describes the altruistic donors, patient-donor pairs and unpaired patients registered with the NKR between April, 2012 and December 2014, the period during which we have complete registration data. The blood-type of both altruistic and paired donors is skewed away from O donors and in favor of A donors relative to the US population. The deceased donor population in a typical year has about 45% of 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 (31.5%). The excess supply of O patients relative to O donors registered with the NKR results in a group that is difficult to transplant. Indeed, 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 an A blood-type. This distribution is consistent with centers responding to incentives created by NKR practices that avoid transplants to unpaired O patients. This policy has been instituted to make maximal use of the large number of donors in pairs with O patients.

Additionally, 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 of the reference donor population. On the kidney waiting list, in comparison, the PRA is much lower, just under 20% on average. Another convenient summary of the difficulty of matching pairs in the NKR is given by the match power, which calculates the fraction of other registered patients/donors that a given patient/donor is compatible with. The table shows that there is less than a one-third chance that any given donor is compatible with a typical patient in the NKR pool.

This adverse composition of patients and donors limits the NKR’s ability to organize a large number of transplants. If it had more easy-to-match types, the NKR could assemble cycles and chains with harder-to-match types to facilitate a greater number of transplants. However, there are remarkably few incentives for enrolling pairs.¹⁹ If hospitals are not rewarded by the platform for registering easy-to-match types at the NKR, then they will have an incentive to retain and use these patients and donors internally to organize exchanges with their other patient-donor pairs (Ashlagi and Roth 2014). Under this theory, one would expect that the patients and donors transplanted through within-hospital exchanges are significantly easier to match. Further, we should expect to see some signs of inefficient matching. We now turn to these issues.

¹⁹An exception is that some platforms actively incentivize hospitals to enroll only their altruistic donors by guaranteeing to terminate chains with their own patients. These incentives may not violate NOTA because donors are not compensated. Arguably, a key to the success of the NKR is the attraction of the large number of altruistic donors since inception (the vast majority of NKR transplants have been facilitated through chains).

Table 1: Patients and donors registered with the NKR

N	Altruistic Donors		Pairs		Unpaired Patients	
	164		1265		501	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Patient Blood Type						
A	---	---	23.8%	(0.43)	51.1%	(0.50)
B	---	---	15.0%	(0.36)	16.0%	(0.37)
AB	---	---	2.6%	(0.16)	19.0%	(0.39)
O	---	---	58.6%	(0.49)	14.0%	(0.35)
Donor Blood Type						
A	44.5%	(0.50)	44.8%	(0.50)	---	---
B	14.0%	(0.35)	18.5%	(0.39)	---	---
AB	3.7%	(0.19)	5.1%	(0.22)	---	---
O	37.8%	(0.49)	31.5%	(0.46)	---	---
Match Power						
Recipient/Pair	---	---	21.6%	(0.21)	43.0%	(0.39)
Donor	27.6%	(0.16)	25.4%	(0.16)	---	---
Panel Reactive Antibody (PRA)	---	---	48.8%	(0.41)	44.4%	(0.45)
Pair Type						
Overdemanded	---	---	13.8%	(0.35)	---	---
Underdemanded	---	---	42.2%	(0.49)	---	---

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.

Table 2: Efficiency of live-donor exchanges by platform

	NKR		Non-NKR Across Hospital		Within Hospital	
N	1118		480		2781	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Patient Blood Type						
A	34.7%	(0.48)	37.1%	(0.48)	37.0%	(0.48)
B	19.0%	(0.39)	18.5%	(0.39)	17.1%	(0.38)
AB	5.7%	(0.23)	6.7%	(0.25)	5.4%	(0.23)
O	40.6%	(0.49)	37.7%	(0.49)	40.5%	(0.49)
Donor Blood Type						
A	36.8%	(0.48)	37.5%	(0.48)	33.2%	(0.47)
B	18.2%	(0.39)	16.5%	(0.37)	13.8%	(0.35)
AB	3.9%	(0.19)	5.6%	(0.23)	2.8%	(0.16)
O	41.1%	(0.49)	40.4%	(0.49)	50.2%	(0.50)
Panel Reactive Antibody (PRA) (Sensitization)						
Mean	34.95%	(0.40)	34.77%	(0.39)	17.80%	(0.31)
Fraction >90%	16.4%	(0.37)	15.3%	(0.36)	5.3%	(0.23)
Transplant Outcomes						
O Donor → Non-O Patient	2.7%	(0.16)	4.4%	(0.20)	11.3%	(0.32)
O Donor → PRA < 90%, Non-O Patient	1.4%	(0.12)	3.1%	(0.17)	10.5%	(0.31)
Mean Days on Dialysis	1026.6	(1088.1)	1061.1	(1134.8)	960.8	(984.7)
Transplanted Donor Quality Measures						
Age	44.1	(11.8)	44.0	(11.4)	43.2	(11.8)
Body Mass Index (BMI)	26.5	(4.0)	26.7	(4.0)	26.5	(4.3)
Height (cm)	169.4	(9.8)	169.1	(10.1)	169.3	(9.9)
Weight (kg)	76.3	(15.1)	76.8	(14.9)	76.3	(15.2)

Note: Sample of all PKE transplants between January 1, 2008 and December 1, 2014.

Table 2 shows the blood-types, patient PRA and other characteristics of of transplants that were conducted via the NKR, non-NKR across hospital exchanges, and within hospital exchanges. Between 2008 and 2014, 40% of the patients transplanted through the NKR transplanted were blood-type O and a 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 cross-hospital kidney exchanges not facilitated by the NKR.

In contrast, among within hospital kidney exchanges outside the NKR, half 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 the NKR. These statistics indicate that the transplants occurring within hospitals involve much easier to match patients and donors. They are consistent with hospitals retaining easy-to-match patients and donors to conduct within hospital exchanges and steering hard-to-match types to the NKR. Such behavior could be inefficient as NKR can possibly make better use of easy-to-match types.

One easily detectable sign of inefficiency is a transplant between an O donor and a non-O patient. To gain intuition for why such transplants are inefficient, let's assume that all patient-donor pairs are either blood-type A-O or O-A. The vast majority of the U.S. population has either an A or an O blood-type. This simplification primarily abstracts away from blood-types A-A and O-O as these patient-donor pairs can be transplanted amongst themselves. Except for patients with extremely high PRA, tissue-type compatibility is unlikely to be barrier to getting her transplanted because the probability that she is tissue-type incompatible with all donors in a large platform is small. Roth et al. (2007) argue that a similar stylized model captures the main considerations for a large platform such as the NKR, but also consider B and AB blood-types. Their characterization of efficient matching implies that a transplant between an O donor and an A patient would result in a loss of one transplant because one could instead transplant a O patient (O-A pairs are in excess supply) and use the patient's related A donor to transplant the original A patient as well.

Table 2 shows that the NKR uses an O donor to transplant a non-O patient in only 2.7% of cases.²⁰ In about half of these cases, the patient happened to have a very high PRA (above 90%), and therefore was not very likely to match with other donors. Non-NKR cross-hospital exchanges are less efficient than the NKR, but still use O donors quite efficiently. Except in 3% of cases, O donors are used to either to transplant O patients or very high PRA patients. In contrast, 10.5% of within hospital exchange transplants involve an O donor and a relatively low sensitized non-O patient. If each of these transplants sacrifices another transplant as described above, then simply achieving the level of efficiency obtained by the NKR using these patients and donors could have resulted in about 250 additional transplants between 2008 and 2014. This "smoking gun" evidence of inefficient matching within hospitals is a lower bound for the total inefficiency in the market – it ignores other forms of inefficient transplants that are more difficult to detect.

Although exchanges conducted outside platforms seem inefficient, one may wonder whether there are other dimensions on which within hospital exchanges are superior. Possible differences include the time-lag involved in coordinating exchanges across hospitals or the quality of the donor. Table 2 shows that patients receiving a transplant from the NKR or from other platforms typically wait for only about 2 more months on dialysis relative to within hospital transplants. Given that the average patient waits for about 32 months, this difference represents an 8% extra waiting-time. A longer waiting time at the NKR should be expected because, as we showed earlier, patients transplanted through the platform are harder to match on average. Further, it does not appear that there are differences in how desirable the donors might be to patients. The indicators of donor quality such as age, weight, height, BMI are extremely similar across platforms. One reason why patients considering the NKR need not worry about donor quality is that the NKR allows patients and doctors to specify

²⁰A careful reader will notice that 2.7% is larger than the difference between the number of O donors and O patients transplanted through the NKR. This difference is due to the small number of cases in which a kidney from an A donor with subtype 1 can be transplanted into an O patient if she does not have an immune response to this subtype. This also explains why the percentage of Non-NKR across hospitals transplants between O donors and non-O patients is larger than the difference between O donors and O patients within this group.

acceptability criteria for the donor. It also allows patients to refuse transplants even after they have been proposed if the donor is considered unsuitable.

While suggestive of a general tendency of hospitals to maximize profits by prioritizing in-house exchanges to the National Kidney Registry, it is important to note that these patterns do not imply that all centers are engaging in this behavior. Especially given the evidence presented in figures 2 and 3, registration policies are likely to be heterogeneous by centers. Although we have documented significant heterogeneity in hospital behavior, the patterns consistent with health professionals distorting care in response to financial incentives have been documented in the health economics literature more broadly (see Clemens and Gottlieb 2014, for example).

4 Theory

The evidence above shows that kidney exchange markets are fragmented, selected, and inefficient. Further, hospitals seem to be avoiding the national platform because of participation costs. We now develop a simple framework to explain how this inefficiency arises. The framework will also suggest the primitives that need to be estimated in order to measure the overall inefficiency, and to propose better mechanisms.

4.1 Definitions

We model a kidney exchange clearinghouse as a platform that procures submissions (donors and patients) from hospitals, and rewards these hospitals with transplants. The platform has a production function, which describes how many transplants it can produce with a given set of submissions. The model has three key ingredients: exchanges, hospitals, and a utilitarian welfare notion.

Consider a monopolistic platform, that 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 $i = 1, \dots, I$. Given a vector of **quantities** $\mathbf{q} = (q_i)_{i=1, \dots, I}$, with q_i being the quantity of type i submissions, the platform can produce $f(\mathbf{q})$ transplants. The model can be interpreted either as static or as a steady-state from a dynamic model. We will use the steady-state interpretation in the empirical analysis (section 5). All variables are measured in flows, such as transplants per year, and are real numbers.

It is easy to understand the production function with a simple, concrete example. Roth et al. (2007) calculated such a production function, assuming that the market is large enough so that only ABO blood type compatibility matters. They assumed that all submission types are pairs, and that a submission type specifies the blood types of the donor and of the recipient, so that there is a total of $I = 16$ types. In their model, the production function is linear. The marginal product equals 2 for over-demanded pairs, 0 for under-demanded pairs and 1 for all other pairs. Our empirical implementation in section 5 will be more flexible,

allowing submissions to differ by whether they are patient-donor pairs, altruistic donors or unpaired patients. It will also allow types to differ by a host of variables including blood types, antigens, and antibodies. Thus, readers should think of I as being potentially large.

Submissions are provided by **hospitals** indexed by $h = 1, \dots, H$. We restrict attention to the case where hospitals are rewarded with a number of transplants that is linear and anonymous in their submissions. That is, there exists a **vector of rewards** $\mathbf{p} = (p_i)_{i=1, \dots, I}$ such that a hospital that submits a flow of submissions \mathbf{q}^h receives a flow of transplants $\mathbf{p} \cdot \mathbf{q}^h$. This is a good approximation to current platforms' rules, because current matching algorithms maximize a weighted sum of the number of matches without consideration of the pool of patients and donors submitted by the hospital. When a hospital submits a pair, the change in the probability that the platform matches another pair from the same hospital is negligible. That is, current platforms reward hospitals linearly, with p_i equal to the probability that a pair of type i is matched.

We assume that hospitals have quasilinear preferences over the number of transplants they receive from the platform and their submissions. Formally, hospital h maximizes

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

where $C^h(\mathbf{q}^h)$ is the **private cost** of submissions. One particular case is when hospitals maximize the number of their patients that are transplanted, in which case $C^h(\mathbf{q}^h)$ is the number of within hospital transplants that the hospital would have to forfeit in order to submit \mathbf{q}^h .

The final piece of the model is a utilitarian welfare notion. Welfare is defined over an **allocation** $(\mathbf{q}^h)_{h=1, \dots, H}$ that specifies the quantity of pairs supplied by each hospital. We will use two such notions. The first notion is **hospital welfare** $W^H(\mathbf{q}^1, \dots, \mathbf{q}^H)$, which is the total welfare measured from the point of view of hospitals. We evaluate hospital welfare in terms of transplants, because this is the resource that is most easily transferred across hospitals, due to the abundance of underdemanded pairs and unpaired patients (see table 1). Formally,

$$W(\mathbf{q}^1, \dots, \mathbf{q}^H) = f(\mathbf{q}) - \sum_{h=1}^H C^h(\mathbf{q}^h), \quad (1)$$

where \mathbf{q} is the aggregate quantity. Hospital welfare measures efficiency according to hospitals' objectives. This is compelling if, as is common in the market design literature, the goal is to help the agents (hospitals in this case) achieve their objectives.

Hospital welfare is not compelling if there are agency problems between hospitals and the patients and insurers that they represent. As argued in section 2.1, hospitals face a large share of the costs of kidney exchange, but gain few of the benefits. Thus, hospitals may overweight the opportunity cost of making submissions. Define $SC^h(\mathbf{q}^h)$ as the **social cost** for hospital h to supply a vector \mathbf{q}^h submissions. If there are agency problems, social and private costs are different, and there is an **externality** from hospital h 's submissions given by

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

Given an allocation, let $E(\mathbf{q}^1, \dots, \mathbf{q}^H)$ be the **aggregate externality**, which equals the sum of externalities from all hospitals. The externality represents the benefits to other stakeholders that are not internalized by hospitals. While we will use the terminology of agency problems, because the key stakeholders are patients and insurers that the hospitals represent, this wedge includes any possible deviation between hospital welfare and social welfare. These deviations may also be due to the social welfare measure attaching different weights to different individuals, or due to hospitals not behaving optimally for behavioral reasons. In the particular case where there are no agency problems, we have $\partial E = \mathbf{0}$.

Define **total welfare** as

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

Thus, the difference between social and hospital welfare is equal to the aggregate externality E .

Define the **aggregate cost function** $C(\mathbf{q})$ as the minimum of the sum of costs of all hospitals subject to the total quantity submitted being equal to \mathbf{q} . For simplicity, we assume that the production, cost, social cost, and aggregate cost functions are defined over all non-negative real vectors and smooth. Aggregate cost is strictly convex. Quantities are column vectors, while vectors of rewards and gradients are row vectors.

Our model finesses a number of important issues, such as efficiency costs of transferring transplants, the choice of a particular welfare function, and the case of multiple competing exchanges. We will return to these issues later, after deriving the main conclusions from the model.

4.2 Illustrative example

Our model clarifies the similarity between kidney exchange, neoclassic producer theory and the economics of platforms. This connection requires some abstraction, because kidney exchange markets do not involve monetary transfers, and the relevant numeraire are transplants. In this section we consider a concrete example to clarify how the model connects to actual kidney exchange markets, and to be clear about our substantial assumptions.

Consider the following example. There is a monetary cost $K^h(\mathbf{q}^h)$ for hospital h to supply \mathbf{q}^h submissions to a kidney exchange platform, which is borne by the hospital. The cost can include fees from the platform, costs or rearranging the hospital schedule around the platform, and the cost of additional transplant coordinators (see section 2.1). Let $T^h(\mathbf{q}^h)$ be the flow of internal kidney exchange transplants that hospital h foregoes when supplying \mathbf{q}^h .

Hospitals value transplants at v dollars. This value includes profits, and the value that hospitals place on transplanting their patients. Gross revenues from a transplant are of the order of \$150,000. For illustrative purposes, take v to be \$50,000, which represents a generous 50% mark-up on costs. Therefore, hospital utility, in transplant units, equals the number of

transplants minus the monetary costs divided by the value per transplant,

$$W^h = \mathbf{p} \cdot \mathbf{q}^h - T^h(\mathbf{q}^h) - \frac{K^h(\mathbf{q}^h)}{v}.$$

This fits our model by setting

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

Society values transplants at V dollars. Society takes into account the value of transplants for insurers, taxpayers, patients, and for other patients that benefit from reductions in the deceased donor waitlist. Insurers and the taxpayers benefit because transplants save over \$200,000 in medical costs over dialysis, a large share of which is borne by Medicare. Patients have large gains in life expectancy and quality. The net present value of this gain has been estimated to be over \$1,000,000 using an estimates for the value of quality-adjusted life years and medical cost savings (Held et al. 2016). Naturally, some patients who receive a kidney exchange transplant would otherwise receive a kidney from a deceased donor. But, whenever that happens, another patient in the waitlist receives this kidney. So the social benefit of the kidney exchange transplant should still be about the same as the gain from a single transplant. Thus, it is reasonable to select V in the ballpark of \$1,300,000 (Held et al. 2016). There is a substantial wedge between the social and private value of a transplant, because V is more than an order of magnitude larger than v . Moreover, from society's perspective, the additional costs of participating in a platform $K^h(\mathbf{q}^h)$ are negligible compared to the benefits of performing additional transplants. Under these assumptions, social welfare equals

$$SW^h = \mathbf{p} \cdot \mathbf{q}^h - T^h(\mathbf{q}^h) - \frac{K^h(\mathbf{q}^h)}{V}.$$

This fits our model if we set

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

The externality term equals

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

That is, the externality equals the difference, measured in transplant terms, of how much more hospitals care about the monetary costs of kidney exchange compared to society. Our use of the term ‘‘agency’’ for the wedge E^h is somewhat inaccurate. We use agency because it is plausible that a large share of the benefits that are not internalized by hospitals accrue to patients and insurers who contract with the hospital. But, even in our example, E includes benefits to other third parties, so that one may wish to use term it externalities instead of agency. In fact, the general model allows for many other wedges, such as hospitals being unaware of kidney exchange or misperceiving its benefits and costs.

To develop intuition, assume that the monetary cost is linear in the number of submissions, so that $K^h(\mathbf{q}^h) = k \cdot \|\mathbf{q}^h\|_1$. Then the externality is

$$E^h(\mathbf{q}^h) \approx \frac{k}{v} \cdot \|\mathbf{q}^h\|_1.$$

The externality, representing the agency wedge, depends on the cost of kidney exchange as a percentage of the private value of a transplant, and not the social value. Therefore, the agency wedge is substantial, even though the costs of kidney exchange are small relative to V . For example, if k is \$10,000 and v is \$40,000, then the wedge per submission is $k/v = 0.25$ transplants per submission. This is a substantial wedge because hospitals compare it to probabilities of matching in the rewards vector \mathbf{p} . The intuition for this wedge is that hospitals receive a very small share of the social benefits of transplants, but pay many of the costs, so that even small costs are magnified when divided by v . This wedge would not exist if hospitals were reimbursed for the costs of participating in kidney exchange platforms $K^h(\mathbf{q}^h)$.

Figure 4 presents a graphical illustration for this example to clarify the two sources of inefficiency in our model: agency problems, and inefficient platform incentives. The horizontal axis plots aggregate supply \mathbf{q} . The vertical axis plots marginal products, social costs, social benefits, assuming that hospitals choose individual supply optimally given a rewards vector. \mathbf{p}_0 is the current vector of rewards set by the platform, which is equal to the probability of matching each pair. The current quantity is \mathbf{q}_0 , where hospitals choose supply given current rewards \mathbf{p}_0 . Current supply is inefficient, even from hospitals' perspective. The current mechanism rewards hospitals according to probabilities of matching \mathbf{p}_0 , instead of according to marginal products $\nabla \mathbf{f}$, as in the hospital-optimal quantity \mathbf{q}^* . Thus, the first inefficiency is that the platform gives inefficient incentives, $\mathbf{p} \neq \nabla \mathbf{f}$. The second inefficiency is that there are agency problems, so that private cost and social cost differ, $C^h \neq SC^h$, which is the same as saying that the wedge $\partial \mathbf{E} \neq 0$. The socially efficient quantity is achieved if both the platform uses efficient incentives, $\mathbf{p} = \nabla \mathbf{f}$, and agency problems are solved so that $\partial \mathbf{E} = 0$. In the example this will happen, for example, if hospitals are reimbursed for the costs of paired kidney exchange K^h .

This intuitive explanation finessed two subtleties, as will be made clear by the formal results. First, efficient platform incentives are only approximately equal to marginal products. The reason is that, because there are increasing returns to scale in kidney exchange, marginal products are higher than average products, so that it is not feasible to reward hospitals according to marginal products (Theorem 1). However, estimates in section 5 will show that this adjustment term is negligible for the National Kidney Registry. Second, it is not possible to reach the first-best using a better mechanism if there are agency problems. The reason is that 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 reward hospitals for marginal products, so that it is not possible to reward hospitals for the externalities. This suggests considering a two-pronged approach: design optimal mechanisms with a narrow market design perspective of helping hospitals to achieve their collective goals, and implement policies to solve the agency problems.

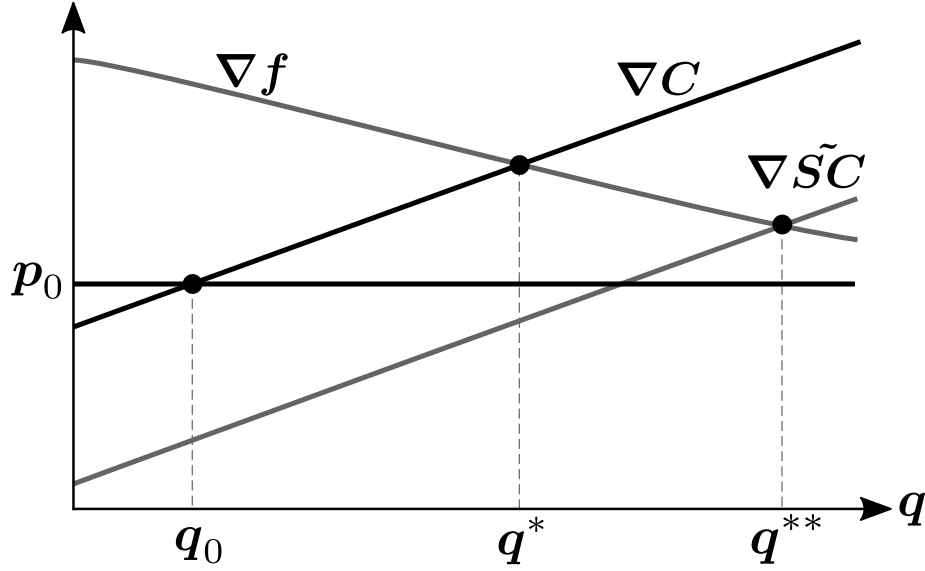


Figure 4: The two Sources of Market Failure

Notes: The horizontal axis represents aggregate quantity of submissions into the kidney exchange platform. The curves represent the marginal product of submissions ∇f , the marginal private cost of submissions ∇C from hospital's perspective, and the marginal social cost of submissions $\nabla \tilde{S}C$. The figure depicts the current quantity q_0 , with agency problems and a suboptimal mechanism, the quantity q^* from a hospital-optimal mechanism but with agency problems, and the first-bets quantity q^{**} with an efficient mechanism, and no agency problems, so that hospital and social incentives are aligned, $C^h = SC^h$. The social cost as a function of aggregate supply is defined assuming that the relevant q vectors are strictly positive, and that hospitals optimize given linear rewards, so that $\tilde{S}C(q) = \sum_{h \in H} SC^h(S^1(\nabla C(q)), \dots, S^H(\nabla C(q)))$. For simplicity, the figure assumes that individual supply is uniquely defined function S^h of rewards and that the exchange has approximately constant returns to scale.

4.3 Optimal mechanisms

We can now give a detailed description of optimal mechanisms. The following theorem collects the main insights.

Theorem 1. *Consider a vector of rewards \mathbf{p} and allocation $(q^h)_{h=1, \dots, H}$ with strictly positive aggregate quantity q that maximizes hospital welfare subject to all hospitals choosing supply optimally given \mathbf{p} , and subject to not promising more transplants than produced. Then:*

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

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

where

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

2. If f has constant returns to scale, then this allocation attains the first-best hospital welfare, and the reward for each type of submission is exactly equal to the marginal product.
3. If, in addition, there are no agency problems, in the sense that $\partial E = \mathbf{0}$, then this allocation also maximizes total welfare.

The theorem focuses on mechanisms that maximize hospital welfare. That is, on how to help hospitals organize a platform that will achieve their collective goals. The theorem has three parts.

The first part shows that, in an optimal mechanism, submissions should be rewarded approximately by their marginal product. The intuition is simple if we ignore the constraint that the platform cannot promise more transplants than it produces. The exchange is similar to a firm that produces a consumption good (transplants), using intermediate goods (submissions). The supply of intermediate goods is efficient when prices \mathbf{p} are equal to marginal products ∇f . In kidney exchange, although there are no financial transactions and no prices, the proof is identical. The first order condition for the first-best aggregate supply is $\nabla C = \nabla f$. When we take hospital incentives into account, the marginal cost curve equals the supply curve, so that optimal rewards are $\mathbf{p} = \nabla f$.

The only complication is the constraint on not promising more transplants than those that are produced. If f has increasing returns to scale, then this constraint is binding, because the number of transplants produced, $f(\mathbf{q})$, is smaller than the number of transplants that would have to be paid in the first-best, $\nabla f(\mathbf{q}) \cdot \mathbf{q}$. This case is easy to visualize in the one-dimensional case, where tangents to the production function cross the vertical axis below zero. Thus, the exchange has to shade how much it rewards hospitals relative to the marginal products. The exact amount of shading for each type of submission is given by the adjustment term \mathbf{A} . Optimally, the exchange shades more aggressively on submissions with less elastic supply. In fact, our formula is identical to the Ramsey (1927)-Boiteux (1956) formula from optimal linear commodity taxation. 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 exchange rules are very unlikely to be efficient. Current exchange rules reward submissions with \mathbf{p} equal to the probability that they are transplanted, instead of the optimal rewards. So, there should be a wedge between the social and private benefits of submissions. That is, hospitals face the dilemma of choosing whether to perform actions that help their patients, or actions that help the system as a whole. The clearest example of this dilemma is a hospital with two overdemanded pairs, who chooses to match them internally instead of submitting them to an exchange, resulting in the kind of inefficiency that we 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, so that the adjustment term equals zero, and achieves first-best hospital welfare. This is important because, as we will show, the NKR is well within the region of approximately constant returns to scale.

Therefore, optimal mechanisms can be calculated only using information about marginal products, and do not depend on detailed information about 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 practical approach for is to introduce a simple dynamic points mechanism. For example, hospitals can be credited a number of points equal to the marginal product of each submission to an exchange. Then, a point is subtracted whenever a hospital transplants a patient. The exchange performs optimal matches, but gives higher priority to patients in hospitals with a higher point balance, and with a constraint that no balance can go below a certain bound. Naturally, there are important theoretical issues about how to give incentives in this kind of mechanism, while maintaining efficiency (Hauser and Hopenhayn (2008)), and important implementation details. We will return to these issues in section 6.

The third part of the theorem states that with constant returns to scale, and with no agency problems, the optimal mechanism achieves first-best welfare. This clarifies that there are two possible sources of inefficiency in our model: inefficient platform incentives, and agency problems. Inefficient platform incentives happen if rewards deviate from the optimal, $\mathbf{p} \neq \nabla \mathbf{f}$. In the platforms literature, this problem is usually attributed to wedges between the goals of the platform and of society (Rochet and Tirole (2003); Armstrong (2006); Weyl (2010)), or to some prices being rigid at zero. Agency problems happen if hospitals do not fully internalize the welfare of the parties that they represent, $\partial \mathbf{E} \neq \mathbf{0}$ (Jensen and Meckling (1976)). If platform incentives are efficient ($\mathbf{p} = \nabla \mathbf{f}$) and if there are no agency problems ($\partial \mathbf{E} = \mathbf{0}$), the market will work efficiently. Therefore, the significant inefficiency documented in section 3 is due to a combination of inefficient platform incentives and agency problems.

The upshot of this analysis is that, much like in standard neoclassic producer theory, many key questions about kidney exchange depend on the production function. In the next section we estimate the production function, and use it to measure the magnitude of the inefficiencies, and to develop simple policy responses. The key results come from three analyses. First, we can measure total inefficiency by estimating the returns to scale, and estimating how many more transplants would be performed if production was moved to the efficient scale. This is similar to the misallocation literature in macroeconomics (Hsieh and Klenow 2009). Second, we can design optimal mechanisms using marginal products. This is similar to the derivation of optimal policy formulas based on a small set of statistics commonly used in public finance (Dixit and Sandmo 1977; Saez 2001; Chetty 2009). Finally, we can use our price-theoretic framework to measure the gains from optimal mechanisms. This is similar to the standard Harberger triangle analysis. Using these results, we can gauge the importance of policies that address agency problems versus platform incentives.

5 Production Function Estimates and Results

5.1 Estimation

This section describes how we estimate the production function $f(\mathbf{q})$ for a large kidney exchange platform such as the National Kidney Registry. The most common approach for estimating a production function is based on observed input-output data from multiple firms (c.f. Olley and Pakes 1996). This approach uses variation in chosen inputs to learn about how firms transform them into outputs. Empirical implementations usually have to deal with the endogeneously chosen inputs and typically specify parametric forms for the production function.²¹

There are several challenges to implementing this approach in our setting. First, \mathbf{q} is high dimensional with potentially each patient-donor pair, altruistic donor or unpaired patient being a separate type. The literature on production function estimation, in contrast, is typically concerned with only a handful of inputs (capital, labor and materials, for example). Second, the theory provides little guidance on the functional form for $f(\mathbf{q})$. Finally, our ideal units for $f(\mathbf{q})$ and \mathbf{q} are in terms of transplant rates and arrivals, respectively. This creates challenges in associating specific arrivals with transplants.

Instead, we use detailed institutional knowledge on the operations of the NKR to build a simulation model for computing $f(\mathbf{q})$. The procedure mimics the various phases in the functioning of a kidney exchange platform, from registration to transplantation. Using knowledge of the technological and business processes has several advantages over an approach based on observed input-output data. First, it allows us to replicate the features of $f(\mathbf{q})$ without imposing restrictions on the substitutability or complementarities between various components of \mathbf{q} . Second, we circumvent standard endogeneity concerns about the chosen inputs because we build a simulation that calculates $f(\mathbf{q})$ for any vector of inputs. Third, we can estimate the production function by observing data from a single “firm,” the NKR. The remainder of this subsection described this procedure.

Estimating the production function

1. **Registration:** The first phase is registration with the NKR in each period, denoted by t . We simulate the number of registrations in each day using a Poisson distribution with parameter λ , which is the mean number of registrations in the NKR per day. The specific patient/donor/pair characteristics are then drawn from the dataset with replacement. Each specific registration j , belongs to the set of altruistic donors \mathcal{A} , the set of patient-donor pairs \mathcal{P} or the set of unpaired patients \mathcal{U} . We use a time-average from these simulated registrations to construct estimated registration rates in the NKR \hat{q}_i for any classification of individual patients, donors or pairs into types, $i = 1, \dots, I$.

²¹Another empirical concern in the productivity literature is the selection problem: firms that survive in an industry may be selected to be unobservably more productive. This issue is less of a concern because the primary goal here is to estimate the production function for a single large platform, the NKR. We discuss potential sources of biases when we extrapolate to other hospitals.

Given a stock of patients and donors registered with NKR from the previous period, denoted \mathcal{J}_t^0 , this procedure yields a new stock \mathcal{J}_t .

2. **Transplant proposals:** Each day the NKR identifies an optimal weighted set of potential exchanges within the stock of patients and donors registered with the platform. Specifically, the NKR solves the problem:

$$\begin{aligned} & \max_{x_{jk} \in \{0,1\}} \sum_{jk} c_{jk} w_{jk} x_{jk} \\ & \text{s.t. (feasibility constraints),} \end{aligned}$$

where $x_{jk} = 1$ denotes a proposed transplant from donor $k \in \mathcal{AUP}$ to patient $j \in \mathcal{PUU}$, 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 allowed and 0 otherwise. In the NKR, a transplant may be infeasible if the donor is biologically incompatible with the patient or is unacceptable by the patient. The weights w_{ij} prioritize transplants that are unlikely in an attempt to utilize hard-to-match donors and transplant hard-to-match patients whenever possible. They typically only break ties between two matches with the same number of transplants in favor of retaining patients and donors that are likely to match in the future.²²

There are three feasibility constraints. First, no donor or recipient is involved in more than one transplant:

$$\sum_j x_{jk} \leq 1, \sum_k x_{jk} \leq 1.$$

Second, a donor that is part of a pair is only asked to donate an organ if her intended recipient has been proposed a transplant. Formally, if $\mathcal{P}_t \subseteq \mathcal{J}_t$ denotes the set of pairs, then

$$x_{jk} = 1 \implies \sum_l x_{kl} = 1 \text{ for } k \in \mathcal{P}_t.$$

Finally, cycles that involve more than three transplants are infeasible due to the logistical difficulty in organizing many simultaneous surgeries. Specifically, a cycle of length n is an ordered tuple, (j_1, j_2, \dots, j_n) where $x_{j_k j_{k+1}} = 1$ for $k < n$ and $x_{j_n j_1} = 1$. We prohibit all cycles of length four or greater. Because there are a very large number of cycle length constraints, we first solve a relaxed problem without these constraints and iteratively add constraints prohibiting such cycles. Appendix C provides details on the algorithm.

3. **Frictions in consumating proposed transplants:** Let $\mathbf{x}^{*,t}$ be the transplants proposed in period t . After solving the problem above, patients are first required to approve the proposed transplants in the solution in consultation with their doctors. The acceptance phase is followed by biological compatibility tests.²³ Proposed transplants that

²²Appendix C contains further details on the specific weights that are used.

²³A crossmatch test is performed using blood samples to verify that the donor's kidney will not be rejected by the patient.

are accepted and approved after testing are consumated. We denote these consumated transplants with $\mathbf{y}^{*,t}$, where $y_{jk}^{*,t} = 1$ only if $x_{jk}^{*,t} = 1$.

Both these phases can take several days. Cycles in which any patient refuses or is incompatible with her proposed donor are abandoned. Consistent with NKR policy, we also abandon any chains in which either the first or the second patient cannot be transplanted. For other chains, all proposals until the first failure are consumated. Finally, all patients and donors in failed cycles and links in chains following the first failure are returned to the pool, and are recorded by setting $c_{jk} = 0$ for future iterations if the donor in k was refused by the patient in j . Platforms often retain the donor belonging to the final patient-donor pair in a proposed chain as a **bridge donor** that may initiate new chains in the future much like an altruistic donor.²⁴

Our simulations suggest that a two-week period for both the acceptance and the biological testing phase and a failure rate of one-fifth for each phase are best calibrated to the observed 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.

4. **Departure and entry phases:** Patients and donors often depart the NKR without a transplant. This may occur if the patient dies, becomes “too sick to transplant” or receives a kidney transplant elsewhere. To account for these departures, we simulate a phase with departure rates calculated using the exponential hazards model (or the constant risks hazard model). 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 $j \in \mathcal{A}$, a patient-donor pair $j \in \mathcal{U}$ or an unpaired patient $j \in \mathcal{P}$; 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 include the fraction of donors ever registered with the NKR that are compatible with the patient in j , the fraction of patients in ever registered in the NKR that are compatible with the donor in j , blood-type dummies for the donor and the patient, and the patient and donor’s ages at registration. We estimate this hazard model using 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

²⁴Based on our simulations and statistics from the data, we assume that all O donors are retained as bridge donors. Allowing other blood-types are allowed to initiate chains for another thirty days provides the best fit with the data. If this period expires, we transplant the “bridge donor” to an unpaired patient. 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 center.

sample period.²⁵ The point estimates for λ_{g_j} and β are presented in the appendix. The new pool, \mathcal{J}_{t+1}^0 , at the start of the next period is given by removing all transplanted patients and donors, and any that have departed from \mathcal{J}_t .

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. The dependence on the initial pool eventually fades away. For the baseline simulations, we initialize the NKR pool with the set of patients and donors registered on April 1, 2012, and burn-in 2000 days of each simulation. We then compute the time average of the total number of transplants to estimate f for each simulation s :

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

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 Efficiency of the Current Mechanism and Optimal Mechanisms

This section uses the estimated production function to study the quantitative implications of the model. Our preferred social welfare measure, SW is the total number of kidney exchange transplants performed, both within hospitals and in platforms. We use this welfare measure because the social value of transplants is high relative to the costs of organizing kidney exchange. It is also convenient because our dataset primarily contains information on transplants.

We start by documenting the main features of the production function, and use them to measure efficiency losses and design better mechanisms. We will focus on our baseline specification before discussing the robustness of the results to our assumptions. The qualitative findings and the policy implications are not sensitive to a broad set of assumptions.

5.2.1 Returns to scale

We first document the returns to scale of the production function, that is, how productivity changes with the size of a platform.

We calculated the production function for pools of submissions \mathbf{q} with the same composition as NKR, but with different scales, as measured by the total flow of altruists and pairs per

²⁵Additionally, our dataset on departure dates is constructed from snapshots of the pool of patients and donors present in the NKR. These snapshots are typically recorded daily, but were sometimes missing for up to a month. Although this missing data does not affect the set of registered and transplanted patients and donors, it results in interval censoring. Because these snapshots are most likely missing at random, we don't expect biased estimates of our hazard model once this interval censoring has been accounted for in our likelihood function.

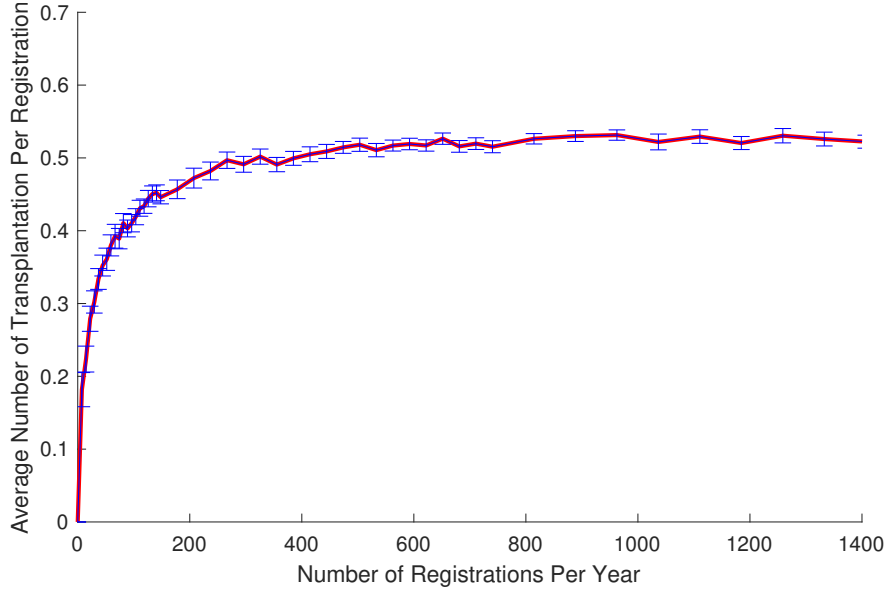


Figure 5: Production Efficiency versus Scale

Notes: The figure plots the average product of a kidney exchange platform versus its scale. The vertical axis represents average products, defined as the share of pairs and altruists who are transplanted. 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 NKR.

year. Figure 5 plots average products, equal to $f(\mathbf{q})$ divided by the flow, as a function of the flow (see Appendix D for a plot of f as a function of scale).

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 0.52, which varies only marginally once scale is sufficiently large. A platform program that is half the size of NKR has an average product of 0.49, while a platform with double the size has an average product of 0.54. Therefore, the market would operate at a high level of efficiency even if there exist a handful of competing platforms. So mergers of sufficiently large platforms would have small effects on efficiency. The estimated relationship between platform size and average product has important implications for the overall market efficiency. Section 3 showed that approximately two thirds of kidney exchange transplants are performed within hospitals that are likely to be much smaller than the NKR. Unfortunately, we do not have direct data on the patients and donors interested in kidney exchange that are available to a hospital. 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 this assumption, one can use the observed rate of kidney exchange transplants to infer the scale for each hospital. Specifically, let y^h be the flow of within hospital kidney exchange transplants conducted at a hospital. The scale of a hospital performing within hospital transplants should be $q^h = \hat{f}^{-1}(y^h)$ where \hat{f} is our estimated production function.

This calculation would suggest that within-hospital transplants, which comprise two-thirds of all kidney exchanges, are taking place at an inefficiently low scale. The median hospital has a scale of 10. The 90th percentile is 29. The largest, Methodist at San Antonio, has a scale of 114. The average product at these efficient scales is, respectively, 0.15, 0.27 and 0.41. Thus, even the largest single-hospital platform is not large enough to operate efficiently. The efficiency losses are considerable, even for the largest single-hospital program, and efficiency losses are larger for the bulk of the market. This is consistent with the finding that hospitals often perform matches that are inefficient from a social perspective.

Of course, these calculations may be biased if the assumptions laid out above are violated. We discuss these biases and robustness analysis below where we calculate the overall inefficiency of the market relative to the first-best.

5.2.2 Misallocation: inefficiency due to small production scale

We start by using the baseline approach in the previous section to give a rough estimate of the inefficiency due to market fragmentation. That is, we estimate how many additional transplants would be performed if the entire kidney exchange market worked at the efficiency of NKR’s scale. To do this, we use a hospital’s estimated scale, $q^h = \hat{f}^{-1}(y^h)$, to calculate the gap in efficiency between the hospital and NKR. The efficiency gap multiplied by the hospital scale is the total number of transplants that are lost due to the inefficiently small scale. These lost transplants equal the total deadweight loss, because we are interested in a social welfare function that is equal to the number of transplants. Due to data limitations, we can calculate lower bounds for the inefficiency, but cannot calculate an unbiased point estimate.

This approach allows for a simple and transparent estimation strategy but suffers from four potential biases. First, the composition of submissions in hospitals may differ from that in NKR. To assess robustness to this assumptions, we estimated the inefficiency using patient and 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.²⁶ Because hospitals in the top quartile of participation should be submitting a less selected pool of patients and donors, robustness of the estimates to this estimates should provide a sense of whether compositional differences may be introducing biases in our approach. Second, while some hospitals are “islands” that perform virtually all kidney exchange transplants within their hospital, other hospitals perform both internal and external matches. Our inefficiency estimate assumes that a hospital is an island, and is biased for hospitals that are not islands, with no clear direction in the bias. To address this issue, we calculated the efficiency loss disaggregated by whether a hospital participates in the NKR, and by the fraction of the hospital’s paired kidney exchanges that are conducted through the NKR. Third, hospitals may use a different matching technology than NKR. For example, Bingaman et al. (2012) report that Methodist at San Antonio (perhaps the most sophisticated single-hospital

²⁶We 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.

program) initially used a Microsoft Access Database, and that their algorithm was “stratified by ABO compatibility and then by HLA compatibility.” This kind of greedy priority-match algorithm is less efficient than the linear-programming algorithms used by the NKR.²⁷ On the other hand, single-hospital programs face simpler logistical constraints, which may increase their productivity vis-à-vis our estimates. The direction of this bias is not signed in general, but it is more likely that single-hospital platforms are less efficient than our estimated production function. Fourth, our approach is likely conservative for the overall inefficiency of the market because it keeps the patients and donors that are interested in kidney exchange fixed. However, this flow is endogenous, and affects the magnitude of the deadweight loss. In general, the direction of this bias is ambiguous. The chief concern is that hospitals value transplants at less than the social value, and due to administrative costs, expend inefficiently low effort in recruiting patients and donors. If incentives were optimal, hospitals may try to recruit a greater number of and more valuable donors into kidney exchange. Our approach does not account for this margin because we do not observe recruitment effort and is therefore likely to underestimate the overall inefficiency of the market.

The inefficiency estimates are displayed in Table 3. Estimates under the baseline assumptions suggest that 527.9 transplants are lost per year due to market fragmentation. This estimate is likely to be biased for the reasons discussed above. Nevertheless, the extensive robustness checks indicate that the true inefficiency is large. The first issue is related to the composition of patients and donors submitted to the NKR. Columns (1) to (3) present estimates under alternative assumptions on the composition of patients and donors that are available to the hospital. A comparison of the estimates suggest that the overall inefficiency is not particularly sensitive to these compositional differences. One reason for this robustness is that tissue-type compatibility is more likely to be the main barrier to kidney exchange the scale of a hospital a compared to the scale at which a large platform such as the NKR operates (Roth et al. 2007). The second potential issue is the inability to accurately determine the scale of hospitals that partially participate in the NKR. If we restrict attention only to the 177 hospitals that do not participate in NKR, across all pool compositions in our exercise, the efficiency loss in column (1) is 230.0 transplants per year. Within the set of NKR participants, the 17 hospitals that participate in NKR and are in the lowest quartile of fraction of transplants performed in NKR alone contribute to an efficiency loss of 111.2 transplants per year. These subsamples and alternative specifications for the pool composition assuage concerns about the first and second biases. These estimates are likely coarse lower bounds to the true inefficiency because they consider subsamples and because of the third and fourth sources of bias. Thus, we are left with the robust finding of considerable inefficiency due to market fragmentation. This is in line with our descriptive finding that hospitals often perform inefficient matches.

²⁷Methodist at San Antonio adopted software written by one of us (Ashlagi) in 2013.

Table 3: Efficiency Loss Table

Sample of Centers	Number of Centers	Total Number of live transplants per year	Total Number of Kidney Exchange Transplants per year	Total Number of Internal Kidney Exchange Transplants per year	Efficiency Loss (foregone transplants through kidney exchange/total transplants through kidney exchange)		
					(1)	(2)	(3)
<i>Panel A: All Centers</i>							
All Centers	256	5657.3	791.0	462.0	550.0	405.8	559.7
<i>Panel B: By center size (number of live transplants per year)</i>							
Top Quartile	60	3814.7	627.4	386.7	344.0	245.4	366.8
2nd Quartile	57	1139.5	118.3	52.8	125.2	100.1	126.0
3rd Quartile	59	542.8	38.2	18.7	60.6	46.6	54.7
Bottom Quartile	58	160.2	7.1	3.7	20.2	13.7	12.2
<i>Panel C: By center size (number of PKEs per year)</i>							
Top Quartile	42	2996.4	596.4	360.5	286.2	200.8	311.4
2nd Quartile	47	1246.6	136.3	65.9	135.8	108.3	140.1
3rd Quartile	39	635.3	45.3	29.2	90.7	71.9	87.8
Bottom Quartile	35	329.4	13.1	6.4	37.4	24.9	20.4
<i>Panel D: By NKR Participation</i>							
Yes	78	3275.3	575.8	294.9	305.5	218.8	321.2
No	177	2506.6	213.4	162.4	244.5	187.0	238.5
<i>Panel E: By NKR Participation Rate (Fraction of PKEs facilitated through the NKR)</i>							
Top Quartile	17	418.9	66.26	8.6	19.7	15.1	19.7
2nd Quartile	17	562.3	100.3	27.7	53.8	41.1	51.8
3rd Quartile	17	1028.0	196.5	97.7	97.1	69.7	104.6
Bottom Quartile	17	1061.7	217.5	165.5	112.2	77.2	121.9

Note: Column (1) assumes that the typical transplant center 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 centers using only the centers with the top quartile of participation rates in the NKR. Column (3) assumes a composition based on centers with the lowest quartile of participation rates.

Table 3 gives tentative insights into the sources of inefficiency. Consider column (1), ignoring the biases discussed above for the purposes of this decomposition. Panel B shows that about 64.1% of the efficiency losses come from hospitals in the top quartile of number of live transplants. Panel E shows that about 53.9% of the losses come from the hospitals in the top quartile of number of kidney exchange transplants. Even though large hospitals perform internal exchanges more efficiently, their market share is higher, so that they account for most of the efficiency losses. Panel C shows that about half of the efficiency losses are due to hospitals that do not participate in the NKR at all, and the other half is due to hospitals that participate partially. Panel D shows that among hospitals that participating at NKR, a disproportionately large share of the efficiency loss is due to the hospitals with low participation.

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 current rules are optimal, then current rewards for different submission types, \mathbf{p} , should equal optimal rewards, $\nabla \mathbf{f} - \mathbf{A}$. We will test this equality at the composition and rate of submissions \mathbf{q} for the NKR during our sample period.

Current rewards \mathbf{p} 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.7). Marginal products $\nabla \mathbf{f}$ can be estimated by numerically differentiating the production function with respect to its 1429 dimensions.

In principle, the adjustment term is harder to estimate, because it depends on the elasticity matrix of the supply of submissions. An additional complication is that the supply elasticity matrix is high-dimensional, making it difficult to estimate using observed submissions to the NKR. Fortunately, because returns to scale are approximately constant for NKR's size, this adjustment term is small, and optimal rewards are approximately equal to marginal products. The intuition is similar to optimal commodity taxation: if the government is primarily concerned with efficiency as opposed to revenues, then optimal taxes are close to zero. We can understand this quantitatively with Theorem 1. The quantity-weighted average of the adjustment term equals

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

That is, the average amount of shading is the difference between the average of marginal product and the average product. Our point estimate is a negligible amount of shading with absolute value 2.16e-04. This suggests that shading is not a major concern. For that reason, we will simply use the approximation of optimal rewards with marginal products.

Figure 6a plots current rewards, the probabilities of matching \mathbf{p} , versus optimal rewards, marginal products $\nabla \mathbf{f}$, for each type of submission in the data. The marginal product of each of the 1429 types is measured with noise due to the small number of simulations for

each type. Figure 6b plots the same figure with pairs aggregated by category, and is much more precise.

The figures are qualitatively similar to the theoretical predictions on which types of submissions are valuable that were derived by Roth et al. (2007) for idealized large markets. For example, in Roth et al. (2007), the marginal product of underdemanded pairs is 0. Our estimates also show that the underdemanded pairs have a marginal product of close to 0. In Roth et al. (2007), the marginal product of an overdemanded pair is 2 whereas our estimates suggest that the marginal product of an overdemanded pair with low sensitization is only 1.40. One reason for this difference is that we estimate that these pairs only get matched with probability 0.79 because many of these pairs depart the platform without getting matched. Our empirical model also reveals interesting information about how productivity varies with sensitization. For example, the marginal product of overdemanded and self-demanded pairs goes down considerably as these pairs become more sensitized. These quantitative differences and heterogeneity across sensitization can be important in practical implementations of points systems.

Both figures show that there is a large wedge between the current rewards and optimal rewards. Altruists and overdemanded pairs with low PRA are far below the 45-degree line. Some types of submissions are drastically over-priced. Overdemanded pairs with low sensitization have marginal products of 1.40, but their probability of matching is 0.79. Even more extreme, altruistic O donors have a marginal product of 1.88, but a probability of matching of only 0.93. Other types of submissions are overpriced. Underdemanded pairs with low sensitization have marginal products of approximately 0.04, but have a substantial probability of being matched of around 0.37. These differences suggest that the platform can do considerably better, by increasing rewards to the productive and undervalued submissions while reducing rewards to the unproductive but overvalued submissions.

5.2.4 Optimality of simple mechanisms

The marginal products suggest that current platform rules are far from optimal, and that point mechanisms like the ones described in Sections 4 and 6 are likely to do better. Moreover, because the adjustment term \mathbf{A} is negligible, the marginal products alone will allow us to describe optimal mechanisms. However, in principle, the high dimensional marginal products estimated above may be noisy and too complex to implement in practice. We now develop a simple point mechanism based on a few patient and donor characteristics that is nearly optimal.

Figures 6a and 6b show that simple mechanisms are nearly optimal because it is possible to categorize submissions in a handful of types that are highly predictive of marginal products and of the probability of matching. A mechanism that assigns points based on these categories can be explained to participants with a simple table. Another, more systematic, way to design a simple, and approximately optimal mechanism is to estimate a regression tree. Figure 7 depicts the best cross-validated tree predictor using default tuning and cross-validation parameters for the marginal products. We allowed the tree to depend on the patient's PRA,

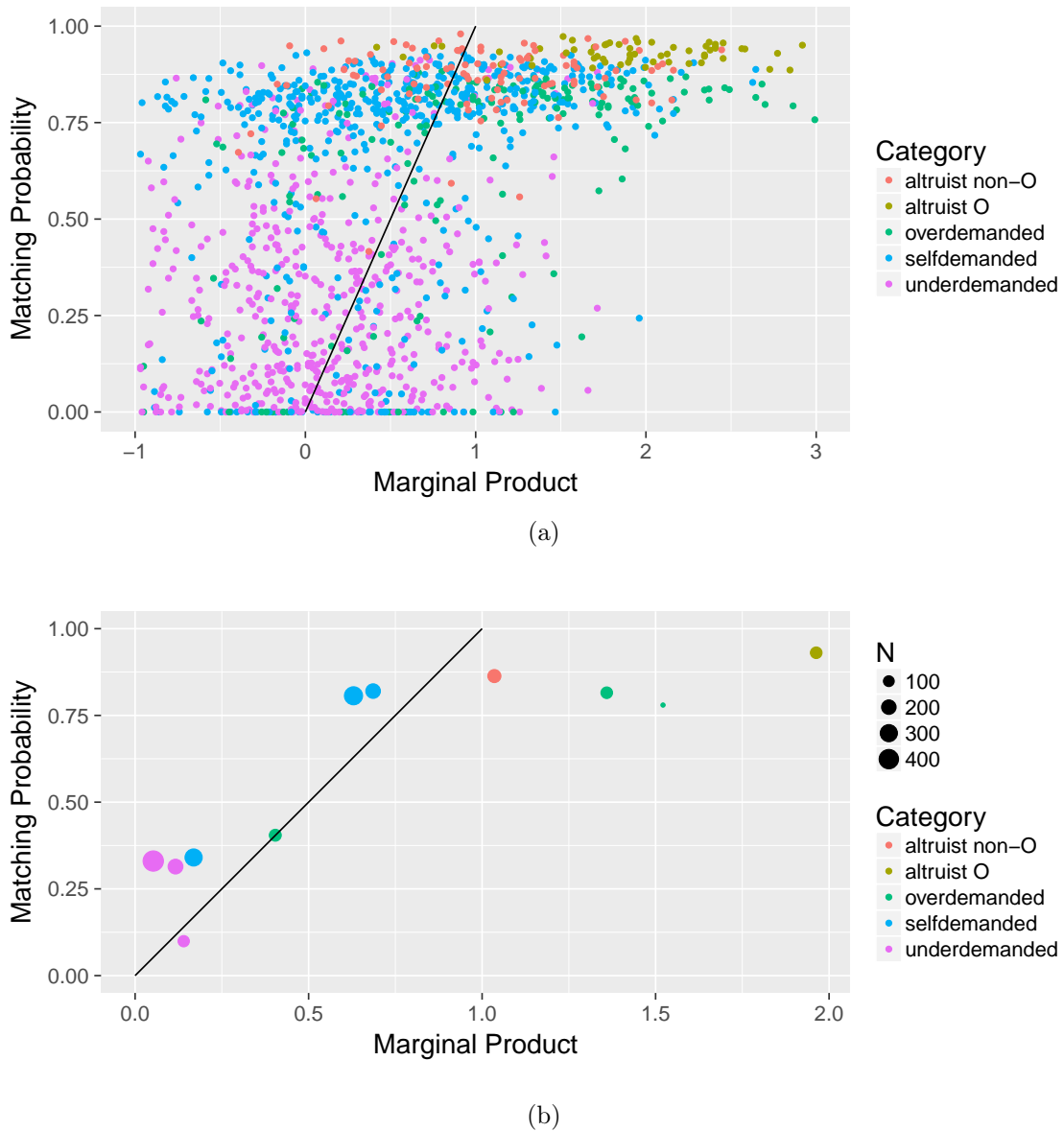


Figure 6: Private versus Socially Optimal Rewards for Submission Types

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.

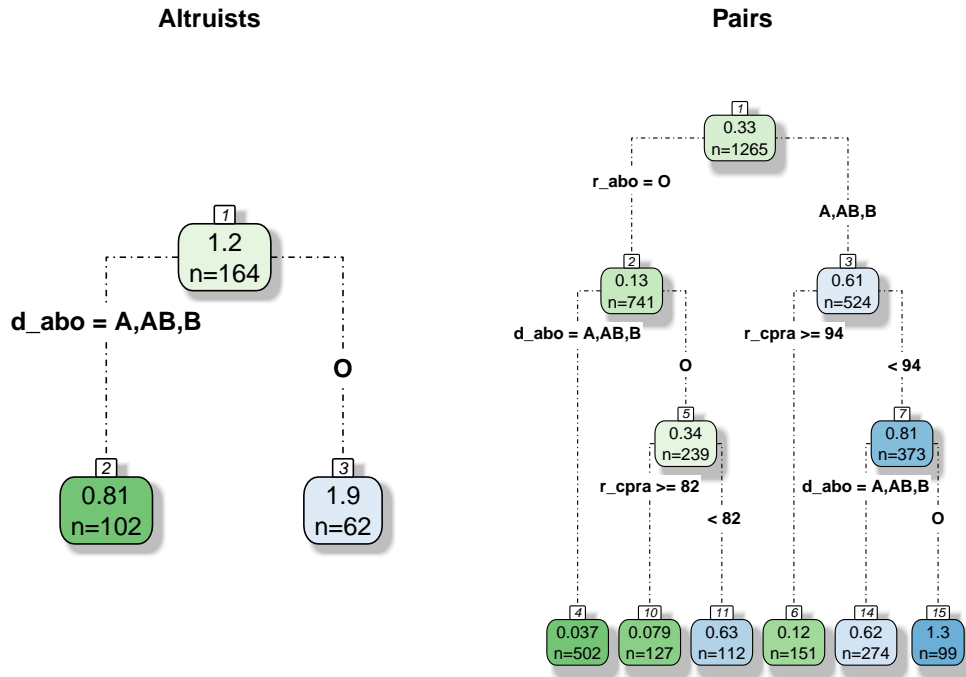


Figure 7: Regression Tree for Marginal Products

Notes: This is the best cross-validated regression tree for predicting marginal products as a function of whether a submission is a pair or altruist, blood type, and PRA. The branching algorithm is the default in the R package `rpart`, with the option to create at least 30 splits. The pruning algorithm selects the shortest tree that is within one standard error of the best cross-validated mean-squared error, following Friedman et al. (2001).

submission type (altruistic, patient-donor pair, unpaired patient), and ABO blood type. It is remarkable that cross-validation in the pruning algorithm chooses a simple tree, suggesting that marginal products can be well approximated with a simple and intuitive function of characteristics (see the figure for details on the algorithm). Moreover, this kind of tree can be easily explained to market participants (Gigerenzer and Kurzenhaeuser 2005 and Gigerenzer and Goldstein 1996). These results show that approximately optimal mechanisms are sufficiently simple to be used in practice.

5.2.5 Welfare gains from simple optimal mechanisms

We now estimate the gain in welfare from moving to the simple optimal mechanism described. The gain equals 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 total welfare.

The calculation of deadweight loss is similar to a multi-dimensional version of the deadweight

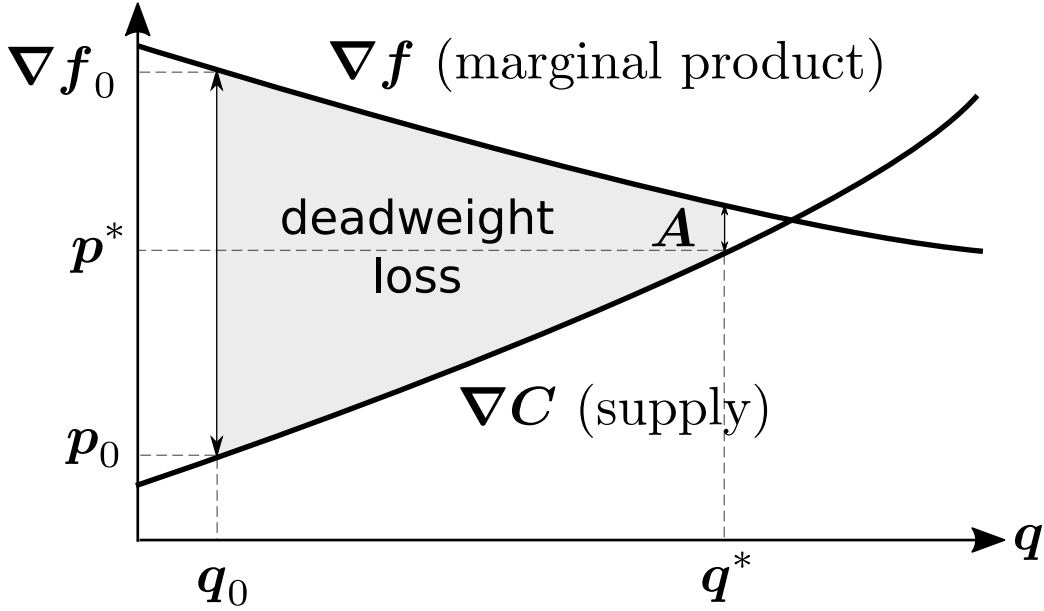


Figure 8: Hospital-Welfare Deadweight Loss from the Current Mechanism

Notes: The deadweight loss from the current mechanism is the shaded area between marginal products and the supply curve. Current rewards are p_0 , equal to the probability of matching each type of submission, while optimal rewards p^* equal marginal products. Current quantities q_0 and rewards p_0 are observed. Marginal products ∇f , including the current value ∇f_0 can be calculated from the production function. In contrast, the supply curve ∇C and optimal rewards p^* and quantities q^* are not observed, and depend on the elasticity of supply. The figure represents multidimensional objects, so that this area is a path integral going from current rewards p_0 to optimal rewards p^* .

loss from linear commodity taxation. Figure 8 illustrates a one-dimensional projection. It depicts the current aggregate supply q_0 , the current rewards p_0 , the current marginal products ∇f_0 and the optimal aggregate supply q^* . Intuitively, the hospital deadweight loss $W(q^*) - W(q_0)$ equals the area between marginal product curve ∇f and marginal cost curve ∇C between q_0 and q^* . Formally, the deadweight loss is the integral of $\nabla f(q) - \nabla C(q)$ as q goes from q_0 to q^* . This is the multidimensional version of the Harberger triangle formula, that is, the area between the marginal benefit and marginal cost curves.

Again, the intuition is similar to that in linear commodity taxation. In linear commodity taxation, the deadweight loss is proportional to the square of the tax wedge, so that large taxes lead to large welfare losses. In fact, a similar approximation to the deadweight loss holds in our setting, which we formalize as the following proposition.

Proposition 1. *Consider a strictly positive aggregate supply of pairs of q_0 , which is produced when hospitals choose supply optimally given rewards p_0 . Consider aggregate supply q^* and rewards p^* that maximize hospital welfare as in Theorem 1. Assume that the matrix $D^2C(q^*) - D^2f(q^*)$ is non-singular. Then the deadweight loss in hospital welfare at q_0 is*

approximated by

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

Alternatively, the deadweight loss is approximated by

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

with the errors bounded by expressions (A10) and (A11) in the appendix.

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 $\nabla \mathbf{f}_0 - \mathbf{p}_0$. The term is the inverse of the derivative of the supply function, so that more elastic supply leads to larger deadweight losses. The term $\mathbf{D}^2 \mathbf{f}$ takes into account how increasing the supply of some types of submissions can change their marginal product. For example, if increasing the supply of overdemanded pairs leads these pairs to become less useful, then the deadweight loss is smaller.

The proposition shows that estimating the deadweight loss requires estimates of $\nabla \mathbf{f}_0 - \mathbf{p}_0$, and either $\mathbf{q}^* - \mathbf{q}_0$ or $\mathbf{D}^2 \mathbf{C}(\mathbf{q}_0) - \mathbf{D}^2 \mathbf{f}(\mathbf{q}_0)$. It is easy to estimate $\nabla \mathbf{f}_0$, \mathbf{p}_0 and \mathbf{q}_0 using the observed data and our production function estimates. Unfortunately, we do not have a good estimate of how hospitals supply pairs and therefore we cannot directly estimate \mathbf{q}^* or $\mathbf{D}^2 \mathbf{C}(\mathbf{q}_0)$. Nevertheless, the large wedge between the current private and social incentives suggests that, unless the elasticity of supply is extremely small, the deadweight loss is significant. Further, the second formula provides a path for estimating the deadweight loss under a broad range of assumptions about supply elasticities. To do so, we restricted attention to mechanisms that set vectors of rewards for the seven categories of pairs that are predictive of marginal products in the regression tree analysis. For these categories, we estimated the wedge $\nabla \mathbf{f}_0 - \mathbf{p}_0$ and the curvature matrix $\mathbf{D}^2 \mathbf{f}$ directly from the data. We then calculated the deadweight loss assuming that the elasticity of supply is constant, and varying assumptions on cross-elasticities.

Figure 9 plots the deadweight loss for a constant elasticity supply function with own-price elasticities ranging from 0 to 6. The middle curve describes the results for zero cross elasticities and the lower and upper curve works with non-zero cross-price elasticities. The deadweight loss is zero when elasticities are zero, because in that case changing the mechanism has no effect on the supply of pairs. The deadweight loss is estimated to be zero if supply is perfectly inelastic and is monotonically increasing in the elasticity. For small elasticities the deadweight loss is approximately linear, and equal to about 35 transplants per year times the elasticity. The deadweight loss is significant for most of this range and above 50 transplants per year if the elasticity is larger than 2. For very high elasticities, the deadweight loss is no longer linear because the curvature of the production function matters. The deadweight loss

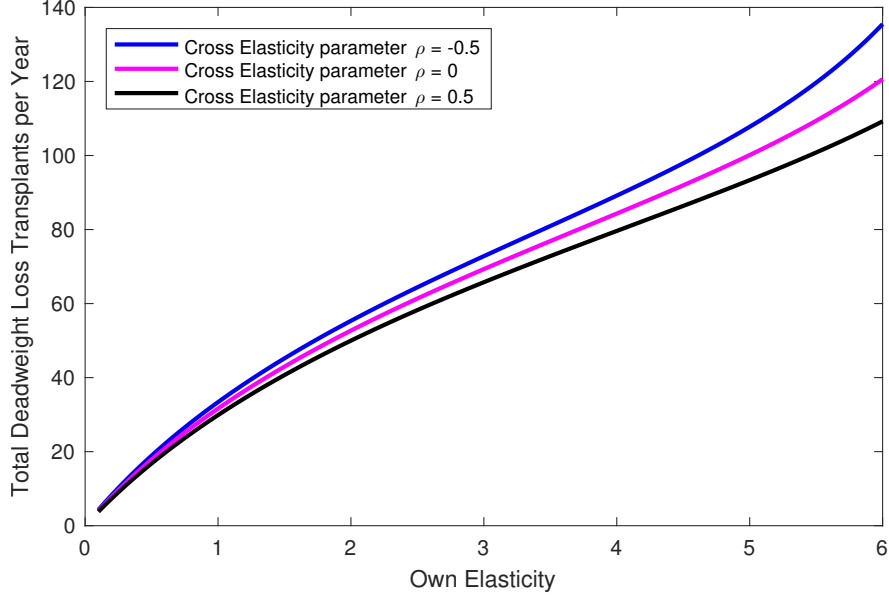


Figure 9: Hospital-Welfare Deadweight Loss from the Current Mechanism

Notes: The estimated hospital-welfare deadweight loss from the current mechanism, using the approximation from Proposition 1, as a function of the elasticity matrix of supply. Own-elasticities are in the horizontal axis. The elasticity of q_j with respect to p_i is assumed to be the revenue share of submission type i , times a cross-elasticity parameter ρ , times the own elasticity. Different curves correspond to different values of ρ , and the middle curve is the case of no cross elasticities.

at an elasticity of 6 is only 120, because of decreasing marginal products of the productive pairs that the optimal mechanism attracts.

These results imply that addressing the inefficient platform incentives has a large positive impact, unless the elasticity of supply is extremely low. However, evidence in section 3 suggests that most hospitals only register a subset of their patients and donors interested in kidney exchange with the NKR and many other hospitals do not participate. Further, observed behavior was consistent with hospitals responding to their financial incentives. These observations suggest that the supply elasticity is unlikely to be very small. Taken together, our results suggest that optimal point mechanisms are not only simple, but also likely to have a substantial effect on the total number of transplants.

5.2.6 Importance of agency problems and inefficient platform incentives

Inefficiency in kidney exchange is driven by two market failures, which have different policy responses. To derive policy implications, it is important to evaluate whether each source of market failure is quantitatively important.

While the results do not allow us to precisely evaluate the importance of each market failure, the results give us useful information on whether each market failure is important. First of all, a conservative misallocation analysis shows that the total deadweight loss is about 180

transplants per year and possibly much higher, with most specifications yielding numbers in the range of 350 transplants. 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 are important as long as supply is not inelastic. The gain in hospital welfare from moving to an optimal mechanism is zero for zero elasticity of supply, 35 transplants per year for an elasticity of 1, and 50 transplants per year for an elasticity of 2. Thus, unless supply is extremely inelastic, optimal mechanisms generate appreciable gains in hospital welfare. Moreover, if there are agency problems, the gains in social welfare are even higher because hospitals undervalue transplants. Specifically, hospital welfare deducts the transplant-denominated private cost of hospitals providing more submissions. When there are agency problems, these private costs are significantly inflated relative to social costs, resulting in a social welfare gain from optimal mechanisms that is even higher.

Taken together, these analyses imply that agency problems are important unless elasticities are extremely high and there is a precise combination of factors. The misallocation analysis implies that the total deadweight loss is at least 180 transplants per year. 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. But, even for a high elasticity of 6, the Harberger triangle yields a deadweight loss of 120, still considerably below our most conservative estimate of 180. The only way that these estimates can overlap is if we have both high elasticities, and the approximation in Proposition 1 is significantly downward biased. The bias in the approximation depends on how much the production function deviates from the quadratic Taylor series, so that the bias is high if ∇f is very convex. Thus, the only way that we can attribute all of the deadweight loss to inefficient platform incentives is if we have a combination of high elasticities, ∇f being sufficiently convex, and the downward biases in the lower bound of the misallocation inefficiency being small.

The upshot is that policies that address either market failures are likely to be valuable. Solving either market failure is likely to generate gains in the order of hundreds of transplants per year. Except for very unlikely form of the supply function, there are significant gains both from platforms switching to efficient mechanisms, and from aligning hospital and social incentives.

6 Discussion and Robustness

6.1 Implementing a point mechanism

Our results suggest that a point mechanism would lead to higher efficiency. However, our steady-state model abstracts from several details. Designing these details raises interesting practical and theoretical questions.

A point system is similar to what Möbius (2001), Hauser and Hopenhayn (2008), Friedman et al. (2006) and Guo and Hörner (2015) call a chips, scrip, or token mechanism. In such a system, each hospital has a balance of points. Hospitals are credited points for their submissions, with the number of points equal to the marginal product, and debited points for each transplant that they receive. There are various ways a platform can utilize points in practice and several “plumbing” details must also be decided.

A simple approach would be to find a maximum (weighted) cardinality match periodically (as currently done) and the point system would be overlaid on the algorithm.²⁸

If there are multiple optimal solutions, select the one with the smallest absolute value of balances. If there are enough underdemanded pairs and unpaired patients, there will often be multiple maximal cardinality matches that will allow the platform to avoid large positive balances. Negative balances require more consideration. Should the algorithm impose a strict bound on negative balances by constraining matches so that no hospital goes under, say, -5 points? 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 (in which case the credit would equal marginal product divided by probability of matching)? Rewards at submission time are less noisy, but raise the risk that hospitals will make shill submissions. 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.

Existing theory on dynamic mechanism design and monetary economics offers insights on the effectiveness of point systems. A number of papers have considered settings where similar issues arise. Möbius (2001), Hauser and Hopenhayn (2008), and Abdulkadiroğlu and Bagwell (2013) consider dynamic favor exchange, and Guo and Hörner (2015) considers 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. Crucially, token mechanisms are close to the first-best if players are patient and there are many time periods (Guo and Hörner 2015). Jackson and Sonnenschein’s (2007) general results on linked decisions imply that the inefficiency of token mechanisms declines as square root of the number of periods. This is consistent with the literature on monetary economics, where money can achieve high levels of efficiency and trade even with simple institutions (Kiyotaki and Wright 1989), even though optimal dynamic mechanisms can often improve on money (Kocherlakota 1998).

An interesting avenue for future research is to develop models in this line for kidney exchange. A plausible conjecture is that the fact that kidney exchange markets are imbalanced with many underdemanded pairs and unpaired patients, makes token mechanisms very efficient. The idea is that market imbalance creates number of maximum cardinality matchings and

²⁸In some cases, weights are used in a static optimization problem as a stand-in for real-world considerations not explicitly considered by the algorithm. For example, the NKR and Methodist at San Antonio assign high priority to compatible pairs (typically pairs with an O donor) in order to match them quickly so that they do not leave the platform. Weights might also be motivated by distributional considerations (Roth et al. 2005b). Our points system is targeted at hospital incentives, and can work along-side such mechanisms with coarse priorities.

thus a large liquidity for breaking ties, making it very easy to transfer transplants across hospitals. Thus, hospitals would get paid back quickly, and incentives are likely to be very similar to those in our steady-state model. Similar “scrip” systems have been shown to improve efficiency of allocation in real-world settings where monetary transactions are prohibited (see Prendergast 2017).

6.2 Maximizing social welfare

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

Proposition 2. *Consider a vector of rewards \mathbf{p} and strictly positive aggregate quantity \mathbf{q} that maximize social welfare subject to all hospitals choosing supply optimally given \mathbf{p} , and subject to not promising more transplants than produced. Assume that the production function has constant returns to scale, that private costs are strictly convex. Define the aggregate externality as a function of aggregate quantity $\tilde{\mathbf{E}}$ as in equation (A12), and assume that it is smooth at \mathbf{q} . Then:*

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

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

where

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

and

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

2. *The adjustment term 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, $\nabla \tilde{\mathbf{E}}(\mathbf{q}) \cdot \mathbf{q}$, is zero.*

Part 1 shows that 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, the planner would like to reward hospitals for what they generate in the platform, plus the externalities. However, if there are not enough transplants to pay for the externalities, the planner has to shade rewards, and it is better to shade rewards for submissions with more inelastic supply, as in optimal linear commodity taxation.

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 has to know what externalities are generated by each kind of submission. This means knowing for what kinds of submissions hospital objectives deviate the most from social objectives. Moreover, it is necessary to know the elasticity matrix, to measure how much shading has to be done for each type of submission. Elasticities matter so long as the average externality is non-zero, in which case the multiplier λ^{SW} is non-zero, and the adjustment term depends on elasticities. Finally, part 3 shows that the optimal reward vector and allocation do not attain first-best social welfare. This implies that allocations that achieve first-best social welfare depend on even more complex incentives, where hospitals may be rewarded non-linearly.

The upshot is that maximizing social welfare, as opposed to hospital welfare, presents some practical challenges. Optimal rewards are more complex, depend on more data, and are sensitive to changes in the incentives facing hospitals that can affect the externality functions.

6.3 Competing platforms

Another interesting question is what is the optimal strategy for competing platforms, and what are the efficiency costs of imperfect competition. To address these issues, consider a platform that faces an inverse supply of submissions $\mathbf{P}_S(\mathbf{q})$. For simplicity, assume that the platform has an empire-building objective, where it maximizes the number of transplants $f(\mathbf{q})$ facilitated by the platform. The following proposition describes the optimal rewards.

Proposition 3. *Consider a platform facing a smooth inverse supply curve of submissions $\mathbf{P}_S(\cdot)$. Consider a vector of rewards \mathbf{p} and strictly positive aggregate quantity \mathbf{q} that maximize the number of transplants in the platform subject to not promising more transplants than produced. Assume that 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,*

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

where

$$\mathbf{A}^C = \frac{\mathbf{q}' \mathbf{D}\mathbf{P}_S(\mathbf{q}) \mathbf{q}}{f(\mathbf{q})} \nabla f(\mathbf{q}) - \mathbf{q}' \mathbf{D}\mathbf{P}_S(\mathbf{q}).$$

2. *If supply is close to perfectly elastic, so that the matrix $\mathbf{D}\mathbf{P}_S$ is close to zero, then rewards are close to marginal products.*

The proposition shows that empire-building platforms deviate from setting rewards equal to marginal products. Instead, the platform subsidizes submissions that are very productive, and whose rewards have a larger effect on supply. To clarify this intuition, consider the case where supply has zero cross elasticities, and own-elasticities ϵ_i . Then the optimal rewards formula simplifies to

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

This formula describes how much the platform marks down 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 are equal the marginal product. When there are multiple types of submissions, 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 that 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 try to exploit their market power, and set rewards inefficiently. The proposition also implies that, if the market is very competitive, platforms set efficient rewards. When supply is very elastic, so that the matrix DP_S is close to zero (and in the particular case of no cross elasticities the ϵ_i are close to infinity), rewards are close to marginal products.

6.4 Discussion of assumptions

An important simplification in Theorem 1 is that we assume that it is always possible to transfer transplants to hospitals without changing the total number of transplants that are produced. This assumption is appropriate for kidney exchange markets because we observe imbalance that makes transferring transplants easy: hospitals have a large number of over-demanded pairs and patients on the deceased donor waiting list that do not have related live donors. Therefore, there are usually multiple matches that maximize the number of transplants and we can reward hospitals with a high balance in a point system at a minimal efficiency cost. Note that, even if we consider additional constraints in Theorem 1, it is still the case that current rules are not optimal. The additional constraints would have associated Lagrange multipliers, and would change the optimal rewards formula. But, even with the added terms, optimal rewards will depend on marginal products, as opposed to only the probabilities of matching.

7 Conclusion

This paper documents significant fragmentation and inefficiencies in the market for kidney exchanges in the U.S. We develop a simple theoretical framework to show that, despite superficial differences with typical economic models, these inefficiencies can arise due to a couple of classical market failures: platforms incentives and common agency problems. Using novel datasets, we estimate that these failures result in a waste of hundreds of transplants per year.

Our model, based on producer theory and economics of platforms, adopts major observations from the kidney exchange literature: first, marginal contribution to the platform varies across types of patients and donors, and second, hospitals are key decision-makers regarding registrations to the platform. The model predicts that simple point mechanisms would increase

efficiency by correcting platform incentives. Our estimates suggest that full efficiency will not be attained, however, without further fixing agency problems.

The estimated production function shows increasing returns to scale for small platforms that quickly plateau. This suggests that there are advantages of co-ordinating at a single national platform although a few large exchanges may co-exist in equilibrium without significant efficiency losses (Ellison and Fudenberg 2003).. While the U.S market is fragmented, several countries such as the U.K. (Johnson et al. 2008) the Netherlands (De Klerk et al. 2005) and Canada (Malik and Cole 2014) have mandated participation at a single national program. One disadvantage of mandating full participation at a single platform is that competition among platforms may be an important driver of innovation.²⁹ Similar considerations are important when countries attempt to merge their kidney exchange markets(see Siegel-Itzkovich 2017, for the experience of Israel and Cyprus).

A central argument in this paper is in favor of rewards in order to encourage hospitals to register valuable pairs at national platforms. Some platforms in the U.S. have been providing hospitals partial incentives. These platforms promise a hospital that submits an altruistic donor (a valuable type) to end a chain with one of its unpaired patients, essentially maintaining a point system. Arguably, the use of this policy by the NKR from a very early stage is a key to its success because it was able to attract a large number of altruistic donors. Our results suggest that all types of registrations (not only altruistic donors) should be rewarded based on their value to the platform. Further, we provide a recipe for calculating these rewards.

It is worth noting that our exercises abstract away from changes in overall supply in the kidney exchange market. This approach is conservative because alleviating the sources of inefficiency that we identify would likely further increase in supply by making kidney exchange more attractive. Hospitals may have increased incentives to recruit valuable patients and donors that do not currently participate in kidney exchange. The effects of the design of the market on this extensive margin is important for further expanding the number of patients that get a life-saving transplant.

A related line of theoretical work is concerned with the incentives of individual patients and donors for participating in kidney exchange (Sönmez and Ünver 2014; Sönmez et al. 2017; Veale et al. 2017). They suggest designs that are likely to make kidney exchange more attractive, particularly to patients and donors that are valuable to a platform. We view these approaches to expanding the kidney exchange market as complementary to ours. We expect hospital incentives, through the design of the platform or through agency problems, to be important irrespective of the individual incentives assigned to patients and donors interested in kidney exchange.

²⁹The various kidney exchange platforms in the United States have individually experimented with different rules and business models in the past that have arguably resulted in many additional transplants. 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.

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.
- 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.
- 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.
- , **David Gamarnik, Michael A Rees, and Alvin E Roth**, “The need for (long) chains in kidney exchange,” Technical Report, National Bureau of Economic Research 2012.
- , **Maximilien Burq, Patrick Jaillet, and Vahideh Manshadi**, “On matching and thickness in heterogeneous dynamic markets,” *arXiv preprint arXiv:1606.03626*, 2016.
- ASTS, American Society of Transplant Surgeons**, “Kidney Paired Donation: Community Perspectives and Best Practices,” 2016.
- 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’équilibre budgétaire,” *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).
- 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, 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.
- Hsieh, Chang-Tai and Peter J Klenow**, “Misallocation and manufacturing TFP in China and India,” *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.
- 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.
- Jensen, Michael C and William H Meckling**, “Theory of the firm: Managerial behavior, agency costs and ownership structure,” *Journal of financial economics*, 1976, *3* (4), 305–360.
- 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.
- Kwak, JY, OJ Kwon, Kwang Soo Lee, Chong Myung Kang, Hae Young Park, and JH Kim, "Exchange-donor program in renal transplantation: a single-center experience," *Transplantation Proceedings*, 1999, 31 (1), 344–345.
- 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.
- Möbius, Markus, "Trading favors," *Mimeo, Microsoft Research*, 2001.
- National Kidney Registry, "Member Center Terms& Conditions," http://www.kidneyregistry.org/docs/NKR_MC_Terms_Conditions.pdf 2016.
- Nikzad, Afshin, Mohammad Akbarpour, Michael A Rees, and Alvin E Roth, "Financing transplants' costs of the poor: A dynamic model of global kidney exchange," 2017.
- Olley, G Steven and Ariel Pakes, "The dynamics of productivity in the telecommunications equipment industry," *Econometrica*, 1996, 64 (6), 1263–1297.
- OPTN, "Organization Procurement and Transplantation Network Policies," <https://optn.transplant.hrsa.gov/governance/policies/> April 2017. Accessed: 2017-05-12.
- 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 multi-center 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.
- 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, Tayfun Sönmez, and M Utku Ünver**, “Kidney exchange,” *The Quarterly Journal of Economics*, 2004, *119* (2), 457–488.
- , – , and – , “A kidney exchange clearinghouse in New England,” *American Economic Review*, 2005, pp. 376–380.
- , – , and – , “Pairwise kidney exchange,” *Journal of Economic Theory*, 2005, *125* (2), 151–188.
- , – , and – , “Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences,” *The American economic review*, 2007, pp. 828–851.
- Russell, Stuart and Peter Norvig**, *Artificial Intelligence: A Modern Approach*, second ed., Prentice Hall, 2003.
- Saez, Emmanuel**, “Using elasticities to derive optimal income tax rates,” *The review of economic studies*, 2001, *68* (1), 205–229.
- Saidman, Susan L, Alvin E Roth, Tayfun Sönmez, M Utku Ünver, and Francis L Delmonico**, “Increasing the opportunity of live kidney donation by matching for two-and three-way exchanges,” *Transplantation*, 2006, *81* (5), 773–782.
- Siegel-Itzkovich, Judy**, “Israel and Cyprus sign accord for sharing kidneys from live donors,” <http://www.jpost.com/Israel-News/Israel-and-Cyprus-sign-accord-for-sharing-kidneys-from-live-donors-497422> 2017. [Online].
- Sönmez, Tayfun and M Utku Ünver**, “Market design for kidney exchange,” *The Handbook of Market Design*, 2013, pp. 93–137.
- Sönmez, Tayfun and M Utku Ünver**, “Altruistically unbalanced kidney exchange,” *Journal of Economic Theory*, 2014, *152*, 105–129.
- Sönmez, Tayfun, M Utku Ünver, and M Bumin Yenmez**, “Incentivized kidney exchange,” Technical Report, Boston College Department of Economics 2017.

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.