HCEO WORKING PAPER SERIES

Working Paper

The Effect of the National Kidney Registry on the Kidney-Exchange Market*

Roksana Ghanbariamin[†]

Bobby W. Chung[‡]

January 17, 2020

Abstract

We assess the causal effect of the National Kidney Registry (NKR), the largest national kidney-exchange network in the U.S., on kidney-exchange outcomes. Analyzing a unique database hosted by the Scientific Registry of Transplant Recipients (SRTR) that contains information on all kidney donors, wait-listed candidates, and transplant recipients in the U.S., we find that patients in an NKR hospital are 2.5 to 3 times more likely than their counterparts in a non-NKR hospital to receive a transplant from a living donor, conditional on wait-time. At the same time, NKR participation does not have a significant effect on the desirability of donors and the health status of recipients. We employ various approaches to ensure our finding is robust in addressing the non-random sorting of patients and donors. As far as the outcomes we study, the expansion of the NKR brings about an overall positive impact on the kidney exchange market.

JEL Classification: I11, L14

Keywords: Kidney-exchange networks; National Kidney Registry (NKR)

1 Introduction

Previous studies have documented the value of a kidney transplant for End-Stage Renal Disease (ESRD) patients, including longer life expectancy, lower cost compared to dialysis, and lower government health expenditure (Wolfe et al., 1999; Held et al., 2016). However, there is a significant shortage of kidneys available for transplant. In 2018, approximately 110,000 patients were waiting on the wait list for a kidney transplant (Organ Procurement and Transplantation Network, 2018). The creation of kidney-exchange networks, which provide a market that enables patients with living but incompatible donors to swap donors, is the solution proposed by Roth et al. (2004). The previous literature has focused on expanding the number of transplants from creating a large pool of available patient-donor pairs in kidney-exchange networks.[1] However, the effect of the expansion of kidney exchange itself on the welfare of patients is unknown.[2]

This paper is the first to assess the causal effect of the adoption of the largest national network, the National Kidney Registry (NKR), by a hospital on the welfare of patients in the kidney exchange market. Participation in a kidney-exchange network enables patients to swap their incompatible living donors in a bigger pool of incompatible patient-donor pairs. Therefore, it increases the marginal value of donors. However, the effect of these networks on the marginal value of the quality of donors is ambiguous. One might expect to see a decline in the quality of matches after the adoption of a network if patients are willing to trade a relatively lower quality donor for shorter wait-time on the waiting list. Alternatively, kidney-exchange networks can increase the incentive of patients in bringing a high-quality

---

[1] The original idea started with Roth et al. (2004), and Roth et al. (2005) who introduced the idea of emerging national kidney-exchange networks. Ashlagi and Roth (2014) focused on expanding the original model, to an alternative that provides incentives for hospitals to register all of their patient-donor pairs in these national networks. Agarwal et al. (2019) estimated the additional number of transplants from merging all kidney-exchange networks into the largest national network in the United States. Teltser (2018) estimated the additional number of transplants created from kidney exchanges using differences in the patient's local exchange activity over time.

[2] Agarwal et al. (2019) provide a mean-comparison of age, BMI, and blood-type of living-donors involved in exchanges between NKR and non-NKR hospitals, and concludes that except for blood-type, there is no observable difference between donors in these hospitals.

donor that is more desirable for trade to improve their chance of finding a compatible donor. The ambiguity in the quality of matches also imposes an empirical question on whether NKR adoption improves the after-transplant health status of a patient. Therefore, we refer patients' welfare to kidney transplant rate, donors' characteristics, and recipients' health status after the transplant in the current discussion.

We combine the information on NKR participation of all transplant centers with a unique database hosted by the Scientific Registry of Transplant Recipients (SRTR) that contains information on all kidney donors, wait-listed candidates, and transplant recipients in the United States. The first-order effect we show is that NKR adoption increases the likelihood of a wait-listed patient to receive a transplant from a living donor. In obtaining an unbiased estimate in survival analysis, we estimate a duration model proposed by Hinchliffe and Lambert (2013) and Royston and Parmar (2002) to account for the existence of competing events, namely death and receiving a transplant from different sources. We then address the concern of the non-random sorting of patients using two distinct approaches. First, we match NKR and non-NKR hospitals based on observed characteristics using the entropy balancing technique (Hainmueller, 2012). Second, we construct hospital connections based on our unique data about common surgeons among hospitals. We exploit time-by-hospital idiosyncratic variations in the number of NKR connections to instrument for the probability of a hospital joining NKR at a particular year. We find that patients in an NKR hospital are 2.5 to 3 times more likely than their counterparts in a non-NKR hospital to receive a transplant from a living donor. At the same time, NKR adoption does not have any effect on the transplant rate of the liver or intestines. This placebo test further gives credence on the existence of the NKR impact on the kidney transplant rate.

We then employ a difference-in-differences approach to examine the effect of NKR adoption on the characteristics of donors and the after-transplant health status of patients. The magnitude of the estimates is mostly small and insignificantly different from zero. The null effect is consistent when we use the synthetic control method to address policy

2

endogeneity due to time-varying unobserved heterogeneity (Abadie and Gardeazabal, 2003; Abadie et al., 2010). We add to the literature on the kidney-exchange mechanism by showing that expanding the size of a network of this kind increases efficiency by increasing the volume of trades. In addition, this expansion does not show any decrements in the quality of the match.

In the next section, we discuss the background of the kidney exchange market and introduce the variables that we use to evaluate NKR. In Section 3, we employ a competing risk survival model to study the impact of NKR on kidney transplant rate. Entropy balancing technique and an instrument are used to validate the unbiasedness of the estimates. We then turn to a diff-in-diff model and the synthetic control method in Section 4 to examine the impact of NKR adoption on donors' characteristics and recipients' health status. Section 5 concludes the overall welfare implication of NKR.

2   Institutional Background

2.1   Kidney Exchange

For an end-stage renal disease (ESRD) patient, a kidney transplant is the preferred method of treatment compared to dialysis, because of longer life expectancy and lower cost (Wolfe et al., 1999; Irwin et al., 2012). Before 2004, the only option for an ESRD patient with a willing but incompatible living donor was waiting on the deceased-donor list for a cadaver kidney. The shortage of transplantable kidneys, under the prohibition against monetary compensation of human organs by United States law, generates the need for creative solutions to make better use of the incompatible living donors.[3] The introduction of kidney exchanges provided an opportunity of donor swapping for incompatible patient-donor pairs (Roth et al.,

---

[3]According to 42 U.S. Code §274e, "It shall be unlawful for any person to knowingly acquire, receive, or otherwise transfer any human organ for valuable consideration for use in human transplantation if the transfer affects interstate commerce."

3

2005).[4] Initially, kidney exchanges were managed by single hospitals within the population

Figure 1: NKR Adoption and Exchange-Transplant Trends
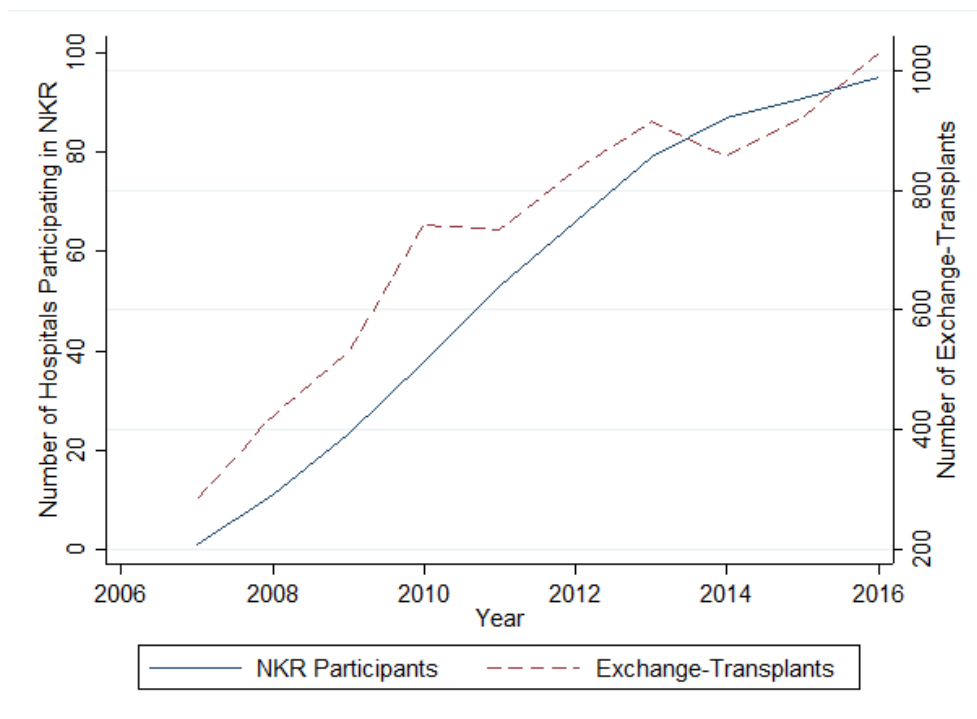
of incompatible patient-donor pairs in that hospital. The need for a thick marketplace constituting a larger pool of patients and donors emerged across hospital kidney-exchange programs, starting from small, local networks, and expanding to three national ones. Today, there are numerous single-hospital programs, several multi-hospital ones, and three major national pairing organizations each with some hospital participants in the United States.

The leading national pairing organization in the U.S. is the National Kidney Registry (NKR). The NKR was the first major network to organize living-donor kidney exchanges nationwide since 2007 effectively. Medical-compatibility and some recipients' preferences regarding their donors is the base of pairing in this network. The NKR runs the matching on a daily basis. Match-runs maximize the number of potential transplants, after giving priority to some hard-to-match patients. Once a match was accepted, the NKR imposes a strict timetable and sets forth rigid guidelines for the transportation of kidneys between

[4]The alternative proposed solution to resolve this shortage is providing monetary incentives for donors. See Bilgel and Galle (2015) and Lacetera et al. (2014) for related discussions.

4

hospitals. The NKR charges hospitals roughly 4,000 dollars per transplant to cover operational expenses, in addition to the annual membership fee (Ellison, 2014). The number of exchange transplants grew gradually as hospitals' adoption of NKR expanded following its introduction in 2007 (Figure 1), and there are 98 hospitals that are currently participating in this network. The NKR is still the leading network in accommodating across-hospital kidney-exchange transplants. As of July of 2018, they have facilitated about 2,800 kidney transplants (National Kidney Registry, 2018).

The other national networks are the Alliance for Paired Donation and Kidney Paired Donation Pilot Program of United Network for Organ Sharing which started operating in 2007 and 2010, respectively.[5] Despite a large number of hospital participants in these two networks, the total number of transplants they have facilitated together is about one-fifth of the number of transplants in the NKR.[6] This is mainly due to their stricter matching criteria, and their weaker reputation compared to the NKR. Given the scale of the system, we focus our analysis on NKR.

2.2 Evaluating NKR

The first-order concern is the transplant rate from a living donor as it is the main objective of NKR. From the perspective of patients, a living transplant is better for ESRD patients because the graft survival rate is higher compared to that of a transplant from a deceased donor (The Alliance for Paired Donation, 2018).[7]

We then look at whether NKR changes the characteristics of donors, which could affect the quality of a transplant. First, blood-type and tissue-type of the donor matter for finding a compatible match. Donors can engage in a transplant when they are blood-type and

---

[5]The United Network for Organ Sharing is the private, non-profit organization that manages the organ transplant system of the United States under contract with the federal government.

[6]A large number of hospitals participating in the other two national networks are mainly because these networks are free of charges. Hospitals' participation in these three networks are not mutually exclusive, but the other two networks facilitated a very small number of across hospital exchange transplants, overall.

[7]Graft survival is an estimate of the probability of the transplant functioning at a finite time after transplantation.

tissue-type compatible with the recipient of their organ. Blood-type compatibility happens when the donor does not have a blood protein that the patient lacks. There are two types of blood proteins, A and B, and the blood type of each individual is an indicator of proteins found in that person's blood. An O blood-type stands for the absence of any of these proteins. This makes O blood-type donors very valuable since they are blood-type compatible with a broader range of patients (Danovitch, 2009). Tissue-type compatibility happens when the donor lacks a protein that the patient has an immune response against. The Panel Reactive Antibody (PRA) is a measure indicating patients' tissue-type sensitivity. It reveals the percentage of the general population with whom the patient is likely tissue incompatible. For example, a patient with a PRA of 20 is incompatible with 20% of the population.

Besides, the donor's general health status affects the expected length of the graft survival of a transplant. There is a negative relationship between the donor's age and the graft survival after transplant, where about 21% of kidney transplant failures are the result of insufficient renal mass due to the higher age of donors (Terasaki et al., 1997). The donor's weight is another factor that impacts the length of graft survival. Medical studies have shown that transplants from obese donors have a higher percentage of acute rejection episodes, malfunction of the primary allograft, and surgical complications (Espinoza et al., 2006).[8] Finally, the risk of kidney failure is higher for patients whose donors had a history of hypertension, smoking, or diabetes (Grams et al., 2016).

Therefore, we are going to evaluate NKR from three aspects: transplant rate, donors' characteristics, and patients' health status after a transplant.

---

[8]An allograft is a tissue graft from a donor of the same species as the recipient but not genetically identical.

3 NKR and the Probability of Transplant

3.1 Data Description

This study used data from the Scientific Registry of Transplant Recipients (SRTR). The SRTR data system includes data on all donors, wait-listed candidates, and transplant recipients in the U.S., submitted by the members of the Organ Procurement and Transplantation Network (OPTN). The Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services provides oversight to the activities of the OPTN and SRTR contractors. Each registration provides detailed medical and demographic characteristics of the individual patients, their registration date, their status on the wait list, including removal from the waiting list if applicable, the removal reason, and the removal date. We merge this patient dataset with the information from an external dataset that contains information on the NKR participation status and participation date of hospitals. This combination allows us to look at the change in the likelihood of a transplant, characteristics of donors, and patients' outcomes before and after NKR participation.

The sample period is restricted to the patients registered beginning from January 2000 through September 2017. Observations before 2000 are omitted due to the poor reporting quality of the variables of interest. The full list in this period consists of 600,880 patients. We analyze individuals whose removal is either due to transplants, death or sickness, transferred to other centers, or received a transplant elsewhere. This leaves us 556,474 patients remain.[9]

Additionally, we drop patients who have a removal code but do not have a registration or removal date and patients with zero survival period. Together with also dropping the patients who have missing information in some characteristics, we have 498,544 patients in the final analysis. In total, there are 287 hospitals in the sample. However, some hospitals (including 15 out of 98 NKR hospitals) do not have information on transplants in all the sample years.

---

[9]Other reasons include improvement in the candidate's condition, refusing transplant, closure of their hospital, transfer to kidney and pancreas waiting list, receiving a transplant in another country, or removal by error or other unknown reasons.

Therefore, we have an imbalanced panel. Table 1 compares the reasons for removal by NKR participation of hospitals after 2007, the year NKR starts to operate, in terms of the number of patients. For a transplant, it can further be classified as a transplant from a living donor or a deceased donor. Patients can receive a living transplant either directly from their own donors or through the exchange from NKR, other networks, or within center matching. Throughout the paper, we call a living transplant from direct donors 'Direct', and through the exchange in an NKR hospital 'Exchange'.[10] The table suggests that NKR hospitals have more living transplants through an exchange than non-NKR hospitals do (t-stat: 6.08).

Table 1: Comparing the Reasons of Removal from Wait-list by NKR Participation After 2007

	NKR Hospitals	Non-NKR Hospitals	|t-stat.|
Reason of Removal			
Transplants	3093.06	1309.50	6.83
	(2357.46)	(1389.42)	
Living Transplants	1210.20	383.90	6.90
	(1117.20)	(514.30)	
Through Exchange	187.02	45.18	6.08
	(218.34)	(97.85)	
Direct	1007.59	335.66	6.88
	(0.0597)	(0.0754)	
Deceased Transplants	1882.85	925.60	5.69
	(1485.87)	(976.89)	
Die or Sick	1243.79	500.46	5.80
	(1181.69)	(602.49)	
Transferred	214.04	137.96	3.20
	(157.85)	(226.95)	
Transplants Elsewhere	420.12	179.97	4.16
	(6193.32)	(3319.89)	
No. of Hospitals	98	189	

 Notes: This table compares the difference in the reasons for removing a listed patient between NKR and non-NKR hospitals. The statistics are generated from the whole sample period (2008 - 2017). The number of observations (listed candidates) in the final analysis is 498,544.

However, we cannot conclude that NKR increases the living transplant rate through the exchange. First, subjects are often at risk of more than one mutually exclusive event. The incident of one of many events may prevent the event of interest from ever happening. For example, the probability of being removed from the wait list due to receiving a transplant

---

[10]The source of the exchange transplants is not distinguished in the data.

depends upon the likelihood of death while waiting and removal due to other reasons, as presented in Table 1. Second, the patients in NKR and non-NKR hospitals can be very different, as shown in Table 2. For example, patients in NKR hospitals tend to be older and more educated. The participation of NKR is possible to be endogenous to the characteristics of a hospital, hence hinders any causal inference. For example, the kind of hospitals enrolling in NKR could be inherently more/less likely to arrange transplant opportunities because of their profile of patients or donors. The following section introduces the empirical strategy to address the concern of competing risks and endogeneity of NKR adoption.

3.2 Competing-Risk Survival Analysis

In many standard survival models like the Cox model introduced by Royston and Parmar (2002), the estimation is based on the assumption of a single risk. That is, subjects are followed until either they experience the event of interest or are censored. In other words, we observe only the minimum of T and C, where C is the censoring time. This setup assumes that the censoring mechanism is non-informative, and anyone in the censored population should have the same likelihood function. More specifically, the censored individual should acquire no information regarding the expectation of survival of that observation (Rodriguez, 2007).[11]

However, as shown in Table 1, the listed patients face multiple events in which the occurrence of one of them might preclude the incidence of others. The existence of competing-risk violates the assumption of a single-risk survival model because some of the censored populations that experience a competing-risk convey information regarding the survival of those observations. Estimating the difference in the transplant rate without accounting for competing-risk can result in a biased estimate, and the direction of bias depends on the frequency of the different types of competing events in the treatment group.

---

[11]In Table A1 of the Appendix A, we also estimate a cox model and a spline-based survival model, which assumes individuals face a single risk. The result on the probability of receiving a transplant of any type is similar to the competing risk model.

Table 2: Characteristics of Hospitals (before 2008) Used in Entropy Balancing

	NKR Hospitals	Non-NKR Hospitals	|t-stat.|
Patients' Characteristics			
Age	42.66	37.14	3.741
	(8.654)	(14.761)	
Body Mass Index	26.06	25.12	2.849
	(1.966)	(3.27)	
O Blood-Type(%)	44.22	46.48	1.905
	(8.268)	(10.387)	
Panel Reactive Antibody (%)	10.95	9.44	1.935
	(4.977)	(7.396)	
Female(%)	41.84	38.22	2.800
	(8.157)	(12.341)	
Black(%)	14.69	15.45	0.431
	(12.133)	(15.678)	
Other Ethnicity(%)	5.587	4.821	0.638
	(9.748)	(8.309)	
College Education(%)	41.32	33.09	3.496
	(15.668)	(21.552)	
Medicare(%)	33.94	36.41	1.270
	(13.031)	(17.686)	
Donors' Characteristics			
Age	39.93	38.91	2.930
	(2.355)	(3.140)	
Body Mass Index	26.67	27.16	3.695
	(0.916)	(1.173)	
O Blood-Type(%)	64.95	66.15	1.039
	(8.694)	(9.197)	
Hypertension History(%)	1.594	1.456	0.305
	(2.703)	(4..478)	
Smoking History(%)	23.38	23.2	0.098
	(12.49)	(16.386)	
Diabetes History(%)	0.082	0.319	1.048
	(0.386)	(2.789)	
Other			
State Population (per 1,000)	13,053	11,704	1.006
	(10770.33)	(9518.403)	

Notes: This table compares the difference in the observed characteristics between NKR and non-NKR hospitals. The variables shown are used to carry out the entropy balancing procedure. The matching is based on the period before 2008 to avoid changes in the characteristics due to NKR participation.

For example, if the censored observations in NKR hospitals have a higher rate of death while waiting, single-risk survival estimation would result in an upward bias in the probability of receiving a transplant in NKR affiliated hospitals.

When competing risks are present, one should consider both the survival time and the event type $(k = 1, ..., K)$. The cause-specific cumulative incidence function for cause k is defined as $F_k(t) = Pr(T \leq t, cause = k)$, which describes the probability of receiving a transplant by time t accounting for the competing events. The cause-specific hazard function for cause k, $h_i^k(t)$, is defined as the conditional probability that a patient, who has not experienced any of the competing events in the period of 0 to t, gets a transplant in the interval of $(t, t + dt)$. Thus, the cause-specific hazard is defined by

$$h_i^k(t) = \lim_{dt \to 0} \frac{Pr(t \leq T_i < t + dt, cause = k | T_i \geq t)}{dt}. \tag{1}$$

Based on equation (1), the cause-specific hazard for event k depends on the cause-specific hazard of all the other $K - 1$ events and requires separate models for each event. The problem with fitting separate models for each cause is that it does not allow the inclusion of shared parameters. Hinchliffe and Lambert (2013) use the flexible parametric model introduced by Royston and Parmar (2002) to fit one model for all events simultaneously, which enables one to estimate the direct effect of covariates on the cause-specific cumulative incidence function of event k without the need to model the other events. They transform the data set by stacking it in a format where each individual has k rows of data, one for each event, so each patient is at risk of experiencing any of the events. This will allow one to incorporate both covariates that are shared across events, and covariates that vary for each event. To test the difference in the probability of receiving a transplant in NKR and non-NKR hospitals in the presence of competing risks, we specify the hazard function as

$$lnH^k_{ihy}(t|\mathbf{X}) = \sum_{j=1}^{J} C_j(ln(t)) + \sum_{j=1}^{J} \beta_j C_j(ln(t)) \times NKR_{hy} +$$

$$\sum_{j=1}^{J} \delta_j C_j(ln(t)) \times X_{ihy} + \sum_{j=1}^{J} \theta_j C_j(ln(t)) \times POP_{hy} + \tau_y + \sigma_h, \tag{2}$$

where $\sum_{j=1}^{J} C_j(ln(t))$ are the J cause indicators depicted in Table 1. We fit a stratified model with J baselines by including the J cause indicators as both main effects and time-dependent effects. The interaction between NKR and the cause indicators captures the distinct treatment effect for each of the J causes. More specifically, it captures the effect of NKR on the probability of experiencing any of the four events. Also, by including the interaction of each of the cause indicators with covariates, we allow for each covariate to vary by each cause separately. Hospital-specific linear trend, year and hospital fixed effects are shared across the J events.[12]

3.3 Entropy Balancing

We offer two alternatives to obtain a better causal estimate of the NKR effect. We first match NKR and non-NKR hospitals using the entropy balancing (E-B) technique proposed by Hainmueller (2012) to obtain more comparable treatment and control groups based on observables. This method allows smooth weighting across units and avoids dropping observations, which is a major defect in other methods like the nearest neighbor and coarsened exact matching.[13] Figure 2 shows the matching result from E-B technique based on the variables in Table 2 which shows that the method vastly reduces the standardized percent bias of the covariates.

---

[12]We adopt **stpm2** package in Stata to perform competing-risk analysis (Lambert and Royston, 2009).

[13]We thank an anonymous referee for this suggestion. When we employ coarsened exact matching using the same set of covariates, no hospitals can be matched.

Figure 2: Entropy Balancing Reduces Differences Between NKR and Non-NKR Hospitals
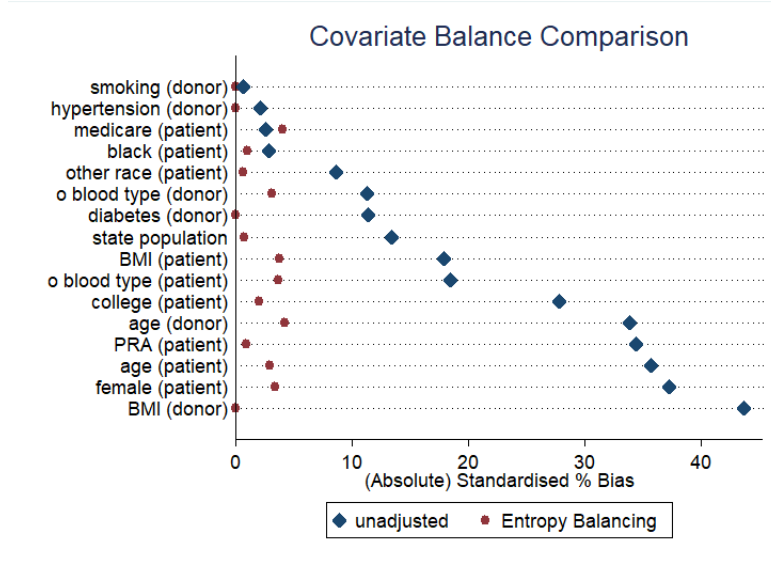
3.4   Instrumental Variable

Entropy balancing relies on the assumption that the covariates used for the matching process will explain policy endogeneity. However, we also worry about selection on unobserved heterogeneity. Therefore, we offer a second way to check our results by instrumenting for the NKR participation of a hospital. We borrow the idea of technology spillover in the health sector (Agha and Molitor, 2018). We utilize the unique information on the surgeon of each transplant in the data and construct a network of hospitals based on the presence of common surgeons. Two hospitals are connected as long as there is one surgeon who performs transplants in both hospitals. We restrict the connections to be static and based on the information before 2008 to avoid any surgeon movement due to NKR adoption. We then track the number of connected hospitals that joined the NKR since 2008. For example, Hospital A is connected with 10 hospitals based on the information before 2008. Some connections started to join NKR at different points of time after 2008. The pattern also differs for Hospital A and B, which creates year-by-hospital variations. According to the spillover idea, a hospital is more likely to participate in NKR when more of its connected hospitals participate (more NKR connections). This is the relevant condition. The identifying assumption is that the

participation of the connected hospitals is not related directly to the transplant arrangement of the focal hospital. As hospital fixed effect and hospital-specific time trends are already included, the exogenous variation then comes from year-by-hospital idiosyncratic differences in the number of NKR connections.[14]

Figure 3 takes a snapshot in 2017 (the end year of the sample period) and shows the distribution of NKR participation among hospitals. A hospital, on average, has around four connected hospitals and 1 NKR connection over the sample period.[15]

Figure 3: Author-Constructed Network Of Hospitals To Identify Exogenous Variations
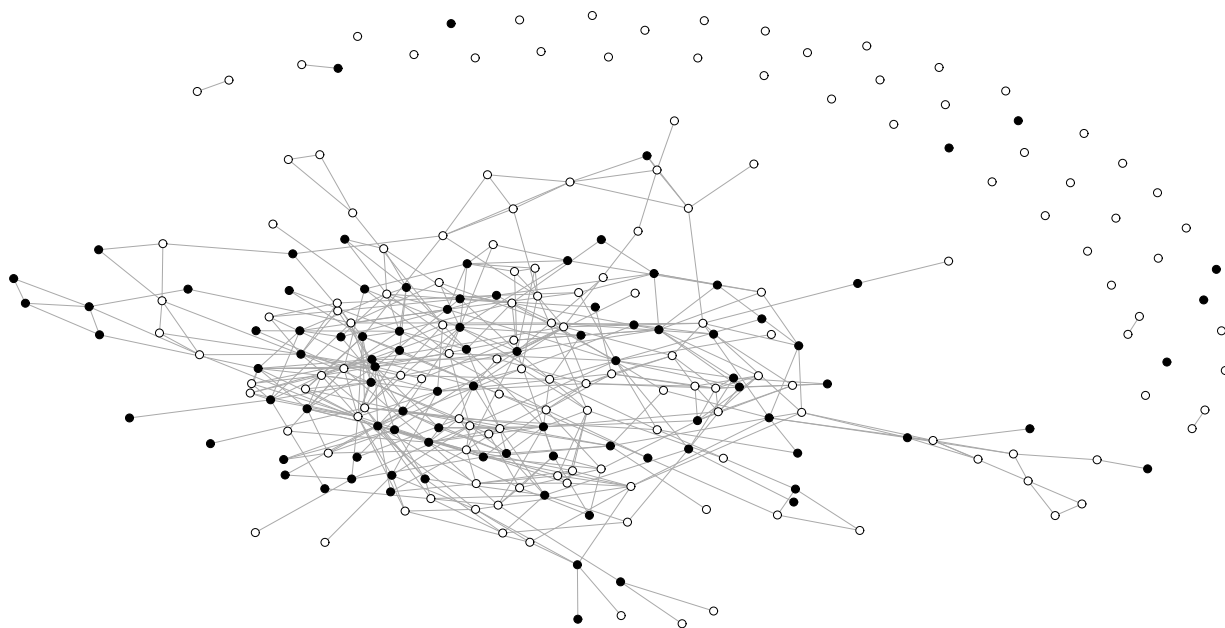
Notes: A static network of hospitals is defined by the presence of common surgeons between hospitals before 2008. The colored dots indicate the hospitals having enrolled in NKR by 2017. We exploit year-by-hospital variation in the number of NKR connections to identify the exogenous variation of NKR participation of a hospital.

Formally, the first stage equation is:

$$NKR_{t,h} = \beta_0 + \beta_1(\text{total NKR connections by time t})_h + \mathbf{X}_{t,h}B + \epsilon_{t,h}, \tag{3}$$

---

[14]We performed a balancing test. The characteristics of donors and patients of a hospital cannot jointly predict NKR adoption of its connected hospitals (F stat: 1.4).

[15]For the number of connected hospitals, the minimum is 0, and the maximum is 20. For the number of NKR connections, the minimum is 0, and the maximum is 6.

where $\mathbf{X}$ is a vector of all covariates (except the NKR indicator) in Equation 2. Since the outcome equation is a nonlinear model, we employ two-stage residual inclusion estimation (Terza, 2018). We first estimate Equation 3. Then, we put $\hat{\epsilon}_{t,h}$ into the second stage as a control.

3.5 Results

Table 3: NKR Increases Living Transplants Through Exchange (Competing Risks)

	(1)		(2)		(3)	
	Without Matching	E-B Matching	Without Matching	E-B Matching	Without Matching	E-B Matching
Competing Events:						
Transplants	1.0553	1.0527				
	(0.035)	(0.035)				
Living Transplants			1.2862***	1.2502***		
			(0.075)	(0.079)		
Direct					1.1372**	1.1027
					(0.065)	(0.069)
Through Exchange					2.9673***	2.957***
					(0.305)	(0.327)
Deceased Transplants			0.9487	.9574	0.9499	0.9591
			(0.048)	(.050)	(0.048)	(0.050)
Died or too Sick	1.0383	1.0598	1.0451	1.0661	1.0456	1.0669
	(0.059)	(0.056)	(0.051)	(.056)	(0.051)	(.0566)
Transferred Elsewhere	0.6536***	0.6547***	0.6589***	0.6597***	0.6586***	0.6595***
	(0.075)	(0.091)	(0.075)	(0.092)	(0.075)	(0.092)
Transplant Elsewhere	1.1012	1.1200	1.1093	1.1274	1.1098	1.1284
	(0.114)	(0.108)	(0.114)	(0.109)	(0.115)	(0.109)
Number of Patients	498,544	493,117	498,544	493,117	498,544	493,117

Notes: Independent variable is a dummy equal to 1 if the patient was listed in an NKR participant hospital. Exponentiated coefficients (hazard ratios) are reported with corresponding (pre-exponentiated) standard errors clustered at the hospital level. Coefficient estimates that are significant at 1%, 5%, and 10% level are denoted with ***, **, and *, respectively. The results are reported by expanding the data and fitting one model for all k events simultaneously. The full set of controls include patient's age, BMI, PRA, dummies for O blood type, gender, race, Medicare participation, college education, and state population of the patient's hospital. All regressions include year and hospital fixed effects, and hospital-specific linear time trend.

We first present the hazard ratio of the NKR indicators from a competing-risk analysis with (adjusted) and without (unadjusted) matching in Table 3. As shown in Column (1), both the estimates from the unadjusted and adjusted sample have similar magnitude. For

example, NKR participation increases the probability of a patient to receive a transplant by 5.53%, though the effect is not statistically significant. Moving down to other events, we do find that NKR participation significantly reduces the probability of a patient to be transferred to other centers by 34.64%. We further differentiate the effect on living versus deceased donors in Column (2). NKR participation significantly increases the probability of a patient to receive a transplant from a living donor by 28.62%. The E-B balancing that addresses non-random selection does reduce the magnitude to 25.02%, but its effect remains significant. In Column (3), the increased rate for a living transplant comes mainly from the transplants through the exchange, and the statistical significance is robust to E-B balancing. The hazard ratio tells that a patient in an NKR hospital is 3 times more likely than a non-NKR patient to receive a living transplant through the exchange.

In Table 4, we check the robustness of the NKR effect by instrumenting the likelihood of NKR participation using the number of NKR connections. In the upper panel, the first stage has decent predictive power (F stat = 10.02) for the endogenous variable. When there is one more NKR connection, a focal hospital is 6.3% more likely to enroll in NKR. When we include the first–stage residuals in the outcome equation, most of the estimates decrease further in magnitude compared to the result in Table 3. For example, in Column (3), an NKR patient is 2.5 times more likely than a non-NKR patient to receive a living transplant through the exchange. However, the results from unadjusted E-B balancing and the IV approach are consistent qualitatively.

Figure 4 compares unadjusted and IV estimation in terms of cumulative incidence function of three competing events, namely living transplant through the exchange, direct living transplant, and deceased transplant. The estimated cumulative probability from the IV approach is smaller than that from the unadjusted estimation. The change in living transplants through the exchange is relatively more apparent than the other two.

In Table 5, we perform a placebo test to see if the NKR effect on exchange transplants is indeed due to a change in other hospital policies. We run the same competing-risk model

Table 4: NKR Increases Living Transplants From Exchange Donors - Robust to IV Approach

	(1)	(2)	(3)
First Stage			
Number of NKR connection		0.063***	
		(0.0199)	
Second Stage			
Competing Events:			
Transplants	0.8603		
	(0.2081)		
Living Transplants		1.0593	
		(0.2611)	
Direct			0.9336
			(0.2281)
Through Exchange			2.4363***
			(0.6612)
Deceased Transplants		0.7814	0.7799
		(0.1919)	(0.1913)
Died or too Sick	0.8464	0.8608	0.8585
	(0.2001)	(.2041)	(0.2032)
Transferred Elsewhere	0.5328**	0.5426**	0.5406**
	(0.1429)	(0.1459)	(0.1452)
Transplant Elsewhere	0.8977	0.9136	0.9112
	(0.2388)	(0.2432)	(0.2423)
Number of Patients		498,544	

 Notes: Estimation follows a 2-stage residual inclusion approach. The instrument in the first stage is the number of connected hospitals participating in NKR by time t. The independent variable, in the second stage, is a dummy equal to 1 if the patient was listed in an NKR participant hospital. Exponentiated coefficients (hazard ratios) are reported with corresponding (pre-exponentiated) standard errors clustered at the hospital level. Coefficient estimates that are significant at 1%, 5%, and 10% level are denoted with ***, **, and *, respectively. The results are reported by expanding the data and fitting one model for all k events simultaneously. The full set of controls includes the patient's age, BMI, PRA, dummies for O blood type, gender, race, Medicare participation, college education, and state population of the patient's hospital. All regressions include year and hospital fixed effects, and hospital-specific linear time trend.

on liver or intestine transplants and find no effects on all competing events. Table A3 of Appendix A validates the null effect using the IV strategy.

Figure 4: Comparing Cumulative Incidence - Unadjusted vs IV
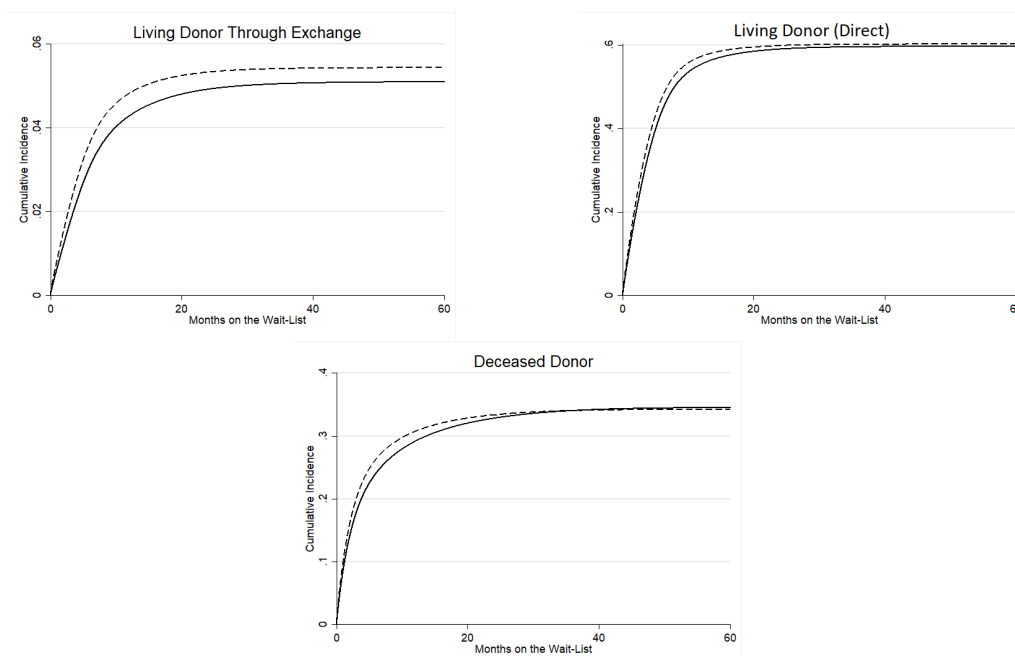
Notes: The three figures compare unadjusted (dashed) and IV (solid) estimation in terms of cumulative incidence function of three competing events, namely living transplant through the exchange, direct living transplant, and deceased transplant.

Table 5: Placebo Test - NKR Does Not Affect Transplant Rate Of Liver/Intestines

| | Competing Events | | | | |
	Living transplants	Deceased transplant	Die or Sick	Transferred Elsewhere	Transplant Elsewhere
NKR	1.1588	0.9650	0.9636	0.8107	0.7463
	(0.2077)	(0.0368)	(0.0551)	(0.1350)	(0.2325)

Notes: This table shows the result from the same competing-risk analysis on liver or intestine transplants (176,842 patients). Exponentiated coefficients (hazard ratios) are reported with corresponding (pre-exponentiated) standard errors clustered at the hospital level. Coefficient estimates that are significant at 1%, 5%, and 10% level are denoted with ***, **, and *, respectively. The full set of controls includes the patient's age, BMI, PRA, dummies for O blood type, gender, race, Medicare participation, college education, and state population of the patient's hospital. The regression includes year and hospital fixed effects, and hospital-specific linear time trend.

4   Donor's Characteristics and Patient's Outcomes

4.1   Difference-in-Differences

In this section, we further investigate the effect of NKR participation on the characteristics of donors and the outcomes of patients who received a kidney transplant. This is an important question because if NKR changes the profile of donors and causes adverse outcomes to patients, the welfare implication of the NKR program is then ambiguous given we find that NKR increases transplant rate through the exchange in the above section.

We analyze the same data but a different specification. We begin by estimating a difference-in-differences model of the form

$$\bar{y}_{ht} = \alpha_0 + \alpha_1 NKR_{ht} + \mathbf{Z_{ht}}\gamma + \theta_h * t + \theta_h + \tau_t + u_{ht}, \tag{4}$$

where the dependent variable, $\bar{y}_{ht}$, measures the average outcome of interest (characteristic of donors/outcome of patients) in hospital h and year t. The variable of interest, NKR_{ht}, is a dummy for the adoption of NKR network in a hospital and year. $\mathbf{Z_{ht}}$ is a vector of time-varying characteristics of hospital h including the average characteristics of patients in Table 2 and the number of living and deceased transplants. Again, we also include hospital-specific linear trends, year and hospital fixed effects.[16]

To assess the impact of NKR on donor's quality, we focus on the characteristics that matter for transplant quality in Section 2.[17] We use principle component analysis to reduce the dimension of donor's characteristics. We adopt the first three components because of the relatively big eigenvalues, as shown in Table B1 of Appendix B. The three components together explain 64% of the total variations of the six characteristics. Table B2 of Appendix B reveals that almost five of the characteristics contribute to the first component. Smoking

---

[16]We do not find evidence in the violation of the parallel trend assumption. In Table A2, the interaction terms between NKR and pre-treatment period indicators are not statistically significant.

[17]The characteristics include age, history of hypertension, history of smoking, history of diabetes, BMI, O-blood type.

history makes the main contribution to the second component, whereas the O-blood type indicator contributes the most to the third component. We thereafter by order call the components 'Overall Desirability', 'Smoking-related,' and 'Blood-Type related' respectively.

Panel A of Table 6 shows the diff-in-diff coefficients of NKR on the three dependent variables, whereas Panel B employs entropy balancing to check the sensitivity. In reporting the estimate and standard error in the diff-in-diff setting, we follow a 2-step procedure by Bertrand et al. (2004) to correct for the serially-correlated outcome variable.[18] In each panel, we further differentiate the effects by transplant types. Panel A shows that, regardless of the transplant type, NKR adoption slightly decreases overall desirability of donors (desirability decreases with the index), reduces the smoking-related problem, and the likelihood of an O-blood type donor. However, the magnitude is small and statistically insignificant. The result by entropy balancing in Panel B is qualitatively the same, with both larger magnitude and standard error.

We then turn to the NKR effect on patient's outcomes with follow-up information between 1 to 3 years after receiving a transplant in Table 7.[19] The diff-in-diff strategy reveals that NKR has only marginally significant effects on patient's outcomes. For example, patients receiving a living transplant in an NKR hospital are less likely to request daily assistants, but the magnitude is significant at a 10% level. This is also the only coefficient that remains significant after matching hospitals based on observed characteristics.

In short, we have little evidence on NKR changing the donor pool and affecting patient's after-transplant health status.

---

[18]In a diff-in-diff setting where the outcome variable is usually serially correlated, the significance of the diff-in-diff coefficient is likely to be overestimated. We first regress the outcome on all covariates, excluding the NKR indicator. We then group the residuals by NKR affiliation and obtain the average for pre- and post-NKR period. In the second step, we regress the averaged residuals on the NKR indicator in a panel of length two. We do observe that without this correction, some of the estimates are significant at 10% level.

[19]Extending the follow-up outcome beyond 3 years will lose many patient samples from recently participating hospitals.

Table 6: NKR Adoption Does Not Change the Characteristics of Donors

	Dependent Variable: Characteristics of Donors[#]		
	(1) Overall Desirability	(2) Smoking-Related	(3) Blood-Type Related
Panel A: Simple Diff-in-Diff			
All			
NKR	0.0105 (0.00177)	-0.000275 (0.000046)	-0.00215 (0.000362)
Living			
NKR	0.00423 (0.00112)	-0.00523 (0.00138)	-0.00176 (0.000465)
Deceased			
NKR	0.0111 (0.00189)	-0.00154 (0.000263)	-0.000001 (0.000002)
Panel B: Entropy Balancing			
All			
NKR	0.0453 (0.0303)	-0.00353 (0.0218)	-0.00725 (0.0105)
Living			
NKR	0.0325 (0.0251)	-0.0402 (0.0274)	-0.0218 (0.0226)
Deceased			
NKR	0.0683* (0.0372)	-0.00808 (0.0238)	0.00169 (0.00947)

Notes: Each entry is a separate regression. Each regression includes the average characteristics of recipients (body mass index, PRA, indicators for age, O blood-type, female, black, other ethnicity, college education, and participation of Medicare), the total number of living and deceased transplant, hospital-specific linear time trends, year and hospital fixed effects. In Panel A, we follow a two-step procedure by Bertrand et al. (2004) to obtain the estimates and standard errors to correct for the overestimation of significant levels. All estimates are *not* significant at a 10% level after adjusting the degree of freedom. In Panel B, the entropy balancing procedure is done by matching the treated and control units based on the control variables in Panel A before 2008. The number of observations is not the same across regressions because of missing dependent variables.

The number of (hospital-by-year) observation for 'All Transplant', 'Living Transplant', and 'Deceased Transplant' is 4,361, 2,985, and 4,311 respectively.

[#] Three composite indexes are generated from age, the record of hypertension, smoking, diabetes, BMI, and O-blood type using principal component analysis.

Table 7: NKR Adoption Has Small Effect On Patient's Outcomes

	Dependent Variable: Average Patients' Outcome (1 to 3 Years Follow-Up)			
	(1)	(2)	(3)	(4)
	Death	Hospitalized	Acute Rejection	Need Daily Assist.
	Panel A: Simple Diff-in-Diff			
All				
NKR	-0.151*	0.00290*	-0.00156	0.000131*
	(0.0225)	(0.000438)	(0.000250)	(1.97e-05)
Observations	4,129	4,064	3,880	4,057
Living-Donor				
NKR	-0.0918	0.00196	-0.000719	-0.00151*
	(0.0146)	(0.000312)	(0.000123)	(0.000239)
Observations	3,830	3,783	3,586	3,775
Deceased-Donor				
NKR	-0.104*	0.00561*	-0.00118	0.00114*
	(0.0160)	(0.000878)	(0.000197)	(0.000178)
Observations	4,008	3,930	3,743	3,920
	Panel B: Entropy Balancing			
All				
NKR	-0.362	0.0127	-0.00701	-0.000221
	(0.490)	(0.0128)	(0.00683)	(0.00488)
Observations	4,129	4,064	3,880	4,057
Living-Donor				
NKR	-0.261	0.00756	0.00173	-0.0104**
	(0.367)	(0.0169)	(0.00954)	(0.00527)
Observations	3,830	3,783	3,586	3,775
Deceased-Donor				
NKR	-0.162	0.0254	-0.00771	0.00680
	(0.781)	(0.0162)	(0.00658)	(0.00509)
Observations	4,008	3,930	3,743	3,920

Notes: The dependent variable in each column is the hospital-year average of the corresponding patient's outcome (1 year to 3 years follow-up). Each entry is a separate regression. Each regression includes the average characteristics of recipients (body mass index, PRA, age, O blood-type, female, black, other ethnicity, college education, and participation of Medicare), the total number of living and deceased transplant, hospital-specific linear time trends, year and hospital fixed effects. In Panel A, we follow a two-step procedure by Bertrand et al. (2004) to obtain the estimates and standard errors to correct for the overestimation of significant levels. In Panel B, the entropy balancing procedure is done by matching the treated and control units based on the control variables in Panel A before 2008. The number of observations is not the same across regressions because of missing dependent variables. Coefficient estimates that are significant at 1%, 5%, and 10% level are denoted with ***, **, and *, respectively.

4.2 Synthetic Control

Non-random selection due to unobserved heterogeneity of patients/donors may contaminate the causal inference explained above. For example, a negative sorting of patients into NKR hospitals could nullify any positive effect of exchange transplants on patient's outcomes. A difference-in-differences approach or matching based on observed characteristics may not alleviate this concern if the unobserved heterogeneity of patients/donors is time-varying.[20]

To check the robustness of our result on the donor's characteristics and the patient's outcomes, we adopt the synthetic control method proposed by Abadie and Gardeazabal (2003) and later expanded by Abadie et al. (2010). Instead of comparing all NKR hospitals with all non-NKR hospitals, the synthetic control method finds a comparable counterfactual hospital for each NKR hospital by assigning weights on non-NKR hospitals.

In general, let $\bar{y}_{ht} = \bar{y}_{ht}^N + \alpha_{ht} NKR_{ht}$ be the observed outcome for hospital h in any year t. The effect of NKR participation for each post-treatment year will then be $\alpha_{ht} = \bar{y}_{ht}^{Post} - \bar{y}_{ht}^N$. We do observe $\bar{y}_{ht}^{Post}$ directly but not the counterfactual $\bar{y}_{ht}^N$. Therefore, Abadie and Gardeazabal (2003) propose to estimate

$$\hat{\alpha}_{ht} = \bar{y}_{ht}^{Post} - \sum_j^J w_j * \bar{y}_{jt}$$

where $\bar{y}_{jt}$ is the observed outcome of hospital j never enrolling into NKR in the sampling period in year t. w_j is the weight assigned to hospital j to minimize the difference between the treated and a weighted average of the untreated. $\sum_j w_j * \bar{y}_{jt}$ is then the synthetic control to approximate the counterfactual $\bar{y}_{ht}^N$. We follow Abadie et al. (2010) to determine w_j by using the pre-NKR values of the corresponding outcome and $\mathbf{Z}_{ht}$ in Equation 4.[21] We then obtain an aggregate effect of $\bar{\hat{\alpha}}_h$ for each NKR hospital by averaging $\hat{\alpha}_{ht}$ across the post-NKR years.

For clarity, we do not differentiate the transplant types. In Table 8, 'mean effect' is the

---

[20]We do not pursue the IV, which we construct for the transplant rate analysis because we confront a weak IV issue (F stat. between 7 and 9) in these outcomes.

[21]We adopt **synth** package in Stata to obtain the weights (Abadie et al., 2014).

average effect by participation year. As the computation requires a full panel, the estimates are derived from 83 out of 98 treated hospitals. To assess statistical significance, we compare the estimated effect to the placebo effect of the synthetic control (Abadie et al., 2010). For each treated hospital, we perform placebo studies for each of the corresponding control units. For example, in Figure 5, a hospital joined NKR in 2008, and the black line displays its outcome variable (quality index) over the sample period. Next, we estimate the NKR effect for each of its control hospitals, assuming they also join NKR in the same year. This forms an empirical distribution, as shown by the lighter-colored lines. At the 90% confidence level, we reject the null of zero effect if $\hat{\alpha}_{ht}$ ranks below the 5th percentile or above the 95th percentile in the placebo distribution.

Figure 5: Construct A Placebo Distribution For Statistical Inference - An Example
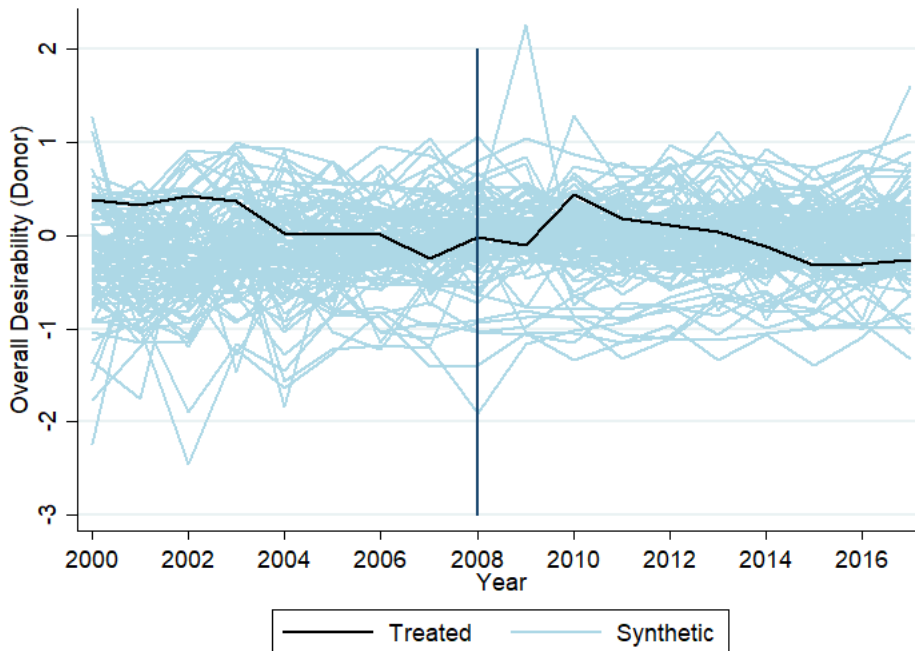
Notes: This graph compares the overall desirability index of a treated hospital (joining NKR in 2008) with the placebo distribution from its synthetic control. We fail to reject the null hypothesis that NKR has no effects on changing the quality index.

24

Because this paper involves multiple event studies, we report the mean rank of each NKR hospital by participation year (Dube and Zipperer, 2015).[22] Consistent with the finding using diff-in-diff and entropy balancing, Table 8 shows that NKR imposes insignificant changes on the three donor's characteristics. For patient's after-transplant health status, we focus on patients receiving a living transplant in Table 9. In contrast to Table 7 that NKR reduces the probability of requesting assistant on daily activity; we find an insignificant NKR effect on all the four health statuses using the synthetic control method.

Using three empirical strategies, we do not find strong evidence that NKR changes the donor's characteristics or the patient's health status after the transplant.

---

[22]Not all treated units have more than 20 synthetic control units. However, we still include them in calculating the mean rank. The result holds qualitatively when we drop those units.

Table 8: Synthetic Control - Insignificant Effect On Donor's Characteristics (All Transplants)

		Dependent Variable: Characteristics of Donors[#]		
		(1)	(2)	(3)
		Overall Desirability	Smoking-Related	Blood-Type Related
By Event Year				
2008	Mean Effect	0.1008	-0.0236	-0.0109
	Mean Rank	0.8131	0.5508	0.4700
2009	Mean Effect	0.0771	0.0055	-0.0344
	Mean Rank	0.7024	0.4721	0.4178
2010	Mean Effect	0.1059	-0.0369	-0.0197
	Mean Rank	0.7628	0.3631	0.5804
2011	Mean Effect	-0.0742	0.0230	-0.0186
	Mean Rank	0.4872	0.5986	0.4164
2012	Mean Effect	-0.0232	0.0942	-0.0577
	Mean Rank	0.5738	0.7617	0.4061
2013	Mean Effect	-0.00003	0.0611	-0.0344
	Mean Rank	0.5009	0.6321	0.4103
2014	Mean Effect	-0.0070	0.0344	-0.0049
	Mean Rank	0.5265	0.8809	0.4844
2015	Mean Effect	-0.0329	-0.0178	-0.0237
	Mean Rank	0.5556	0.4365	0.5417
2016	Mean Effect	-0.1508	0.1429	-0.0186
	Mean Rank	0.3149	0.7525	0.5441
Pooled				
	Mean Effect	-0.0005	0.0314	-0.0248
	Mean Rank	0.5819	0.6054	0.4746

Notes: 'Mean effect' and 'Mean Rank' represent the average effect and the average rank by participation year. The rank of each NKR hospital is obtained by comparing the relative position of its magnitude to the placebo distribution of its synthetic control. At the 90% confidence level, we reject the null of zero effect if $\hat{\alpha}_{ht}$ ranks below the 5th percentile or above the 95th percentile in the placebo distribution.
[#] Three composite indexes are generated from age, the record of hypertension, smoking, diabetes, BMI, and O-blood type using principal component analysis.

Table 9: Synthetic Control - Insignificant Effect On Patients' Outcomes (All Transplants)

		Dependent Variable: Average Patients' Outcome (1 to 3 Years Follow-Up)			
		(1)	(2)	(3)	(4)
		Death	Hospitalized	Acute Rejection	Need Daily Assist.
By Event Year					
2008	Mean Effect	0.0099	0.0078	-0.0136	0.0189
	Mean Rank	0.7402	0.5102	0.3825	0.6430
2009	Mean Effect	0.0024	0.0143	-0.0056	0.03738
	Mean Rank	0.5283	0.5333	0.4864	0.7682
2010	Mean Effect	-0.0118	0.0274	0.0060	-0.0056
	Mean Rank	0.4697	0.5492	0.4619	0.5816
2011	Mean Effect	-0.0071	-0.0354	-0.0017	-0.0124
	Mean Rank	0.4751	0.4393	0.5818	0.5351
2012	Mean Effect	0.0149	-0.0089	-0.0019	-0.0058
	Mean Rank	0.6222	0.5705	0.4961	0.5851
2013	Mean Effect	-0.0307	0.0351	-0.0003	-0.0175
	Mean Rank	0.5533	0.6713	0.5347	0.8887
2014	Mean Effect	-0.0223	-0.0755	-0.0034	-0.0002
	Mean Rank	0.4811	0.5059	0.7500	0.7362
Pooled					
	Mean Effect	-0.0064	-0.0050	-0.0033	0.0021
	Mean Rank	0.5529	0.5569	0.5452	0.6775

 Notes: 'Mean effect' and 'Mean Rank' represent the average effect and the average rank by participation year. The rank of each NKR hospital is obtained by comparing the relative position of its magnitude to the placebo distribution of its synthetic control. At the 90% confidence level, we reject the null of zero effect if $\hat{\alpha}_{ht}$ ranks below the 5th percentile or above the 95th percentile in the placebo distribution.

5   Conclusion

The number of kidney-exchange transplants grew gradually after the introduction of the National Kidney Registry (NKR) in 2007, increasing the pool of patient-donor pairs willing to participate in an exchange. With the expansion of hospitals' use of NKR, the questions that arise are how effective this network is in expanding the number of transplants, and whether such expansions have an impact on the quality of patient-donor pairs in these hospitals.

We assemble a unique dataset that covers the universe of organ transplants in the U.S. back to 2007 and provides the first empirical evidence to the aforementioned inquiries. Overall, our evidence shows that the National Kidney Registry is welfare improving, at least for the outcomes of interest that this paper covers. NKR participation does speed up the kidney-transplant rate from a living donor. Although, in theory, there might be negative sorting either from the donor or the recipient side that potentially dampens the effectiveness of NKR, we do not find any empirical evidence on this uncertainty.

Our study offers a timely discussion about the use of kidney-exchange networks. As noted by Agarwal et al. (2019), the growth in NKR participation has slowed. The concerns over the efficiency of NKR include ethical and logistical issues (Ross et al., 2017), financial management problems (Rees et al., 2012), conflicting incentives (Ashlagi and Roth, 2014), and the risk to the health of living donors (Emara et al., 2008). Our evidence, to a large extent, supports the expansion of NKR adoption. This message is crucial because the size of the patient-donor pool determines the effectiveness of a matching network. Given the positive result in terms of the transplant rate in NKR hospitals, continuous effort is needed to identify the incentives behind the participation decision of hospitals to realize the full potential of the network. As both the practitioners and researchers acknowledged, better coordination among transplant centers can alleviate some of the uncertainty about the NKR adoption (Ashlagi and Roth, 2011; Mast et al., 2011; Melcher et al., 2012). The first stage of our IV strategy provides preliminary evidence on this possibility. We leave this inquiry to

future research.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program," *Journal of the American statistical Association*, 2010, *105* (490), 493–505.

_ , _ , **and** _ , "SYNTH: Stata module to implement synthetic control methods for comparative case studies," *Statistical Software Components*, 2014.

_ **and Javier Gardeazabal**, "The economic costs of conflict: A case study of the Basque Country," *American Economic Review*, 2003, *93* (1), 113–132.

Agarwal, Nikhil, Itai Ashlagi, Eduardo Azevedo, Clayton R Featherstone, and Ömer Karaduman, "Market failure in kidney exchange," *American Economic Review*, 2019, *109* (11), 4026–70.

Agha, Leila and David Molitor, "The local influence of pioneer investigators on technology adoption: evidence from new cancer drugs," *Review of Economics and Statistics*, 2018, *100* (1), 29–44.

Ashlagi, Itai and Alvin E Roth, "Individual rationality and participation in large scale, multi-hospital kidney exchange," Technical Report, National Bureau of Economic Research 2011.

_ **and** _ , "Free Riding and Participation in Large Scale, Multi-Hospital Kidney Exchange," *Theoretical Economics*, 2014, *9* (3), 817–863.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust differences-in-differences estimates?," *The Quarterly journal of economics*, 2004, *119* (1), 249–275.

Bilgel, Fırat and Brian Galle, "Financial incentives for kidney donation: a comparative case study using synthetic controls," *Journal of health economics*, 2015, *43*, 103–117.

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

Dube, Arindrajit and Ben Zipperer, "Pooling multiple case studies using synthetic controls: An application to minimum wage policies," 2015.

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

Emara, Mahmoud, Ahmed Ragheb, Abubaker Hassan, and Ahmed Shoker, "Evidence for a need to mandate kidney transplant living donor registries," *Clinical transplantation*, 2008, *22* (5), 525–531.

Espinoza, R, C Gracida, J Cancino, and A Ibarra, "Effect of Obese Living Donors on the Outcome and Metabolic Features in Recipients of Kidney Transplantation," *Transplantation Proceedings*, 2006, *38* (3), 888–889.

Grams, Morgan E, Yingying Sang, Andrew S Levey, Kunihiro Matsushita, Shoshana Ballew, Alex R Chang, Eric KH Chow, Bertram L Kasiske, Csaba P Kovesdy, Girish N Nadkarni et al., "Kidney-Failure Risk Projection for the Living Kidney-Donor Candidate," *New England Journal of Medicine*, 2016, *374* (5), 411–421.

Hainmueller, Jens, "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies," *Political Analysis*, 2012, *20* (1), 25–46.

Held, Philip J, Frank McCormick, Akinlolu Ojo, and John P Roberts, "A Cost-Benefit Analysis of Government Compensation of Kidney Donors," *American Journal of Transplantation*, 2016, *16* (3), 877–885.

Hinchliffe, Sally R and Paul C Lambert, "Flexible Parametric Modelling of Cause-Specific Hazards to Estimate Cumulative Incidence Functions," *BMC Medical Research Methodology*, 2013, *13* (1), 13.

Irwin, FD, AF Bonagura, SW Crawford, and M Foote, "Kidney Paired Donation: a Payer Perspective," *American Journal of Transplantation*, 2012, *12* (6), 1388–1391.

Lacetera, Nicola, Mario Macis, and Sarah S Stith, "Removing financial barriers to organ and bone marrow donation: The effect of leave and tax legislation in the US," *Journal of health economics*, 2014, *33*, 43–56.

Lambert, Paul and Patrick Royston, "Further development of flexible parametric models for survival analysis," *Stata Journal*, 2009, *9* (2), 265–290(26).

Mast, DA, W Vaughan, S Busque, JL Veale, JP Roberts, BM Straube, N Flores, C Canari, E Levy, A Tietjen et al., "Managing finances of shipping living donor kidneys for donor exchanges," *American Journal of Transplantation*, 2011, *11* (9), 1810–1814.

Melcher, ML, DB Leeser, HA Gritsch, J Milner, S Kapur, S Busque, JP Roberts, S Katznelson, W Bry, H Yang et al., "Chain transplantation: initial experience of a large multicenter program," *American Journal of Transplantation*, 2012, *12* (9), 2429–2436.

National Kidney Registry, https://www.kidneyregistry.org/ 2018. Information access on the web, last accessed on 2018-07-30.

Organ Procurement and Transplantation Network, https://optn.transplant.hrsa.gov/ 2018. Information access on the web, last accessed on 2018-07-30.

Rees, Michael A, 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.

Rodriguez, G., "Lecture Notes on Generalized Linear Models," http://data.princeton.edu/wws509/notes/ 2007.

Ross, Lainie Friedman, James R Rodrigue, and Robert M Veatch, "Ethical and logistical issues raised by the advanced donation program pay it forward scheme," in "The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine," Vol. 42 Oxford University Press 2017, pp. 518–536.

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

_ , _ **et al.**, "A Kidney Exchange Clearinghouse in New England," *American Economic Review*, 2005, *95* (2), 376–380.

Royston, Patrick and Mahesh KB Parmar, "Flexible Parametric Proportional-Hazards and Proportional-Odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects," *Statistics in Medicine*, 2002, *21* (15), 2175–2197.

Teltser, Keith, "Do Kidney Exchanges Improve Patient Outcomes?," Working Paper 2018.

Terasaki, Paul I, David W Gjertson, J Michael Cecka, Steve Takemoto, and Yong Won Cho, "Significance of the Donor Age Effect on Kidney Transplants.," *Clinical Transplantation*, 1997, *11* (5 Pt 1), 366–372.

Terza, Joseph V, "Two-stage residual inclusion estimation in health services research and health economics," *Health services research*, 2018, *53* (3), 1890–1899.

The Alliance for Paired Donation, https://paireddonation.org/ 2018. Information access on the web, last accessed on 2018-07-30.

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.

A Supplementary Tables

Table A1: Probability of Receiving a Transplant

	Cubic Spline Model	Cox Model
NKR	1.0653	1.0653
	(0.045)	(0.045)
Patient's Age	0.9895***	0.9895***
	(0.001)	(0.001)
Body Mass Index	0.9780***	0.9780***
	(0.001)	(0.001)
O Blood-Type	0.7115***	0.7116***
	(0.007)	(0.007)
Panel Reactive Antibody	0.9911***	0.9911***
	(0.000)	(0.000)
Female	1.0905***	1.0905***
	(0.007)	(0.007)
Black	0.6883***	0.6884***
	(0.012)	(0.012)
Other Ethnicity	07153.***	0.7154***
	(0.013)	(0.013)
College Education	1.0745***	1.0745***
	(0.011)	(0.011)
Medicare	0.9036***	0.9037***
	(0.011)	(0.011)
State Population/1000	0.9658	0.9658
	(0.022)	(0.022)
Observations	498,544	498,544
Year FE	Yes	Yes
Hospital FE	Yes	Yes

Notes: Exponentiated coefficients (hazard ratios) are reported with corresponding standard errors. The probability of receiving a transplant is calculated using a single-risk hazard function by fitting a cubic spline function for the baseline hazard. We also provide the estimation using a Cox model for the robustness check. Coefficient estimates that are significant at 1%, 5%, and 10% level are denoted with ***, **, and *, respectively. Standard errors are in parentheses and are clustered at the hospital level.

Table A2: Testing Parallel Trend Assumption

	Characteristics of Donors		
	(1) Overall Desirability	(2) Smoking-Related	(3) Blood-Type Related
1 Year Ahead	-0.154	0.0322	-0.0246
	(0.101)	(0.0390)	(0.0677)
2 Year Ahead	-0.108	-0.0290	0.0178
	(0.0926)	(0.0458)	(0.0699)
3 Years Ahead	-0.0864	0.0381	-0.0145
	(0.0925)	(0.0378)	(0.0596)

Notes: Each regression includes the average characteristics of recipients (body mass index, PRA, indicators for age, O blood-type, female, black, other ethnicities, college education, and participation of Medicare), total number of living and deceased transplant, hospital-specific line time trends, year and hospital fixed effects.

Table A3: Placebo Test - NKR Does Not Affect Transplant Rate Of Liver/Intestines (IV)

| | *Second Stage* | | | | |
| | Competing Events | | | | |
	Living transplants	Deceased transplant	Die or Sick	Transferred Elsewhere	Transplant Elsewhere
NKR	1.1762	0.9795	0.9782	0.8229	0.7576
	(0.3035)	(0.2065)	(0.2066)	(0.2140)	(0.2847)

First Stage IV: 0.073***(0.0175)

Notes: This table reaffirms the placebo result using the IV strategy (the number of hospital-by-year varying NKR connections to instrument for the likelihood of NKR adoption). Exponentiated coefficients (hazard ratios) are reported with corresponding (pre-exponentiated) standard errors clustered at the hospital level. Coefficient estimates that are significant at 1%, 5%, and 10% level are denoted with ***, **, and *, respectively. The full set of controls includes the patient's age, BMI, PRA, dummies for O blood type, gender, race, Medicare participation, college education, and state population of the patient's hospital. The regression includes year and hospital fixed effects, and hospital-specific linear time trend.

B Principal Component Analysis Of Donor's Characteristics

Table B1: Component Analysis

Component	Eigenvalue	Difference	Proportion	Cumulative
1	1.8019	0.7633	0.3003	0.3003
2	1.0387	0.0417	0.1731	0.4734
3	0.9969	0.1409	0.1662	0.6396
4	0.8560	0.1398	0.1427	0.7823
5	0.7163	0.1261	0.1194	0.9016
6	0.5901	–	0.0984	1.0000

Table B2: Contribution Of Each Variable

Donor's Characteristic	Component 1	Component 2	Component 3	Unexplained
age	0.5408	0.2727	-0.0361	0.3944
hypertension	0.5419	-0.1191	0.0046	0.4561
smoking	0.2997	0.7016	-0.3147	0.2281
diabetes	0.3729	-0.4601	0.1252	0.5139
BMI	0.4292	-0.2600	0.2082	0.5546
O Blood	-0.0269	0.3740	0.9169	0.0153