

Performance scores and Strategic Choices of Kidney Transplant Centers*

Han Ng[†]

this version: November 10, 2023

[Click here for the latest draft.](#)

Abstract

Medicare introduced the Conditions of Participation (CoP) in 2007. This policy examines a transplant center's performance every six months. If the observed-expected (OE) 1-year death ratio exceeds 1.5, Medicare flags the center for poor performance and threatens decertification. I analyze the effect of this policy using a difference-in-differences design. The key assumption is that centers with a low observed-expected (OE) 1-year death ratio are unaffected by CoP's introduction. I show that transplant centers respond to the threat of punishment by declining transplants as the OE ratio approaches the CoP threshold, 1.5. This rejection pattern occurs more often for medium and high-risk transplants and low-volume transplant centers. Secondly, I found no evidence that CoP reduced the post-transplant death rate across low/medium/high-risk transplants. My results suggest that strategic behavior among transplant centers is more widespread than previously thought.

JEL codes: I11, I18, L38

*I am very grateful for the guidance and encouragement from Paul Grieco. I thank Marc Henry, Karl Schurter, Mark Roberts, and participants of Econ 555 at Penn State University for their valuable feedback. All errors are my own.

[†]Department of Economics, Pennsylvania State University, Kern Building, University Park, PA 16802, hln14@psu.edu.

1 Introduction

Centers for Medicare and Medicaid Services (CMS) strives to enact policies that protect patient health and safety. One prominent effort to achieve this is the Conditions of Participation (CoP) program, which requires hospitals to report outcomes of patient operations every six months. Medicare uses this report to monitor and penalize poor performers after suitable risk adjustments. An advantage of this program is that it incentivizes improving service quality and patient outcomes. However, one drawback is hospitals cherry-pick and avoid risky patients for performance gains.

This paper examines a section of the U.S. healthcare market influenced by CMS CoP policy, the deceased donor kidney transplant program. Under CoP, transplant centers submit the 1-year outcome of past transplant patients to CMS every six months. CMS flags the center for poor performance if the observed-expected (OE) 1-year death ratio exceeds 1.5, the CoP threshold. Since CMS is the primary insurer for kidney transplants, such a ruling can close a transplant program (Hamilton, 2013). Previous literature by (Schold et al., 2013; White et al., 2015) examined how transplant centers change their behavior after submitting their reports, and CMS flagged them for crossing the threshold. In this paper, I show that CoP affects the behavior of all transplant centers, even for those below the threshold before the submission to CMS, because they can track the prevailing OE ratio and are worried about potential punishment.

When a kidney offer arrives, transplant centers only know the survival of past transplants exceeding the 1-year mark. The uncertainty of the remaining transplant's 1-year survival creates a strategic problem for transplant centers, especially those with OE ratio approaching the threshold. If they accept the kidney offer for a patient, they risk being flagged for poor performance if either the patient or past transplants die within a year. But if they reject the kidney offer, they risk losing the opportunity to transplant the patient¹. In a 2012 *New York*

¹CMS does not penalize centers if a patient does not get a transplant and dies on the waitlist.

Times article, Dr. Lloyd E. Ratner, the director of Columbia Hospital, described how the threat of government penalties made doctors more selective about the organs and patients they accepted. *"... if you have had a couple of bad outcomes recently you say, 'Well, why should I do this?'... You can always find a reason to turn organs down. It is this whole cascade that winds up with people being denied care or with reduced access to care."*(Sack, 2012).

Based on the anecdotal evidence referenced above, I investigate whether the threat of punishment induces transplant centers to reject kidney offers for patients. To cleanly link it to Medicare's CoP policy, however, we must first overcome several empirical challenges. First, we do not observe the many unobservable factors influencing a center's behavior, such as center quality, which plausibly correlates with the OE ratio. I use a difference-in-difference research design to account for unobserved quality. I compare transplant centers of the same OE score before and after CoP's introduction. The comparison produces a causal effect on the policy because I can use the centers with low OE scores as a control group. The policy plausibly did not affect their behavior since they were far from the threshold. Additionally, in many cases, I can estimate specifications with center-fixed effects that control for any time-invariant center characteristics, meaning that the main effects we estimate come only from the changes induced by the OE ratio. Unlike previous literature that uses center-level observation, I have patient-kidney offer data and precise patient and kidney characteristics. This feature allows me to examine how CoP affects the behavior of different patient-kidney pairs.

OE scores are noisy estimates of center performance. Some centers have higher standard errors due to the low transplant volume. Small transplant centers are more likely to reject a kidney offer for a patient before the OE ratio approaches the contract. Once the OE ratio crosses the threshold, I do not find any evidence that small centers differ from large centers.

After establishing these results, I examine whether patient outcomes improved after CoP. I

show that the post-transplant 1-year death rate of transplant centers decreased after CMS introduced CoP. However, I did not find evidence of excess survival gains within a risk group. This finding suggests that the improvement in death rates of transplant centers is due to a shift in transplant composition from high-risk transplants to low-risk transplants rather than an improvement in quality of care. My paper contributes to the literature on the effect of CoP on transplant center behavior. Hamilton (2013), Schold et al. (2013), and White et al. (2015) show that transplant centers change their behavior after they submit their reports and CMS flagged them for poor performance. I show that all transplant centers exhibit strategic behavior, even for those below the threshold before the submission to CMS, because

2 Background on Deceased Donor Kidney Transplant

A patient diagnosed with end-stage renal disease (ESRD) has two options: dialysis or kidney transplant ². Dialysis requires two to three treatments a week. Sessions are time-consuming; patients can be infected if nurses do not disinfect stations appropriately after use. These disadvantages make kidney transplants the cheaper and preferred option (Matas and Schnitzler, 2004). In this study, I focus exclusively on deceased donor kidney transplants that account for 60% of all kidney transplants in the U.S. (AKF, 2003)³. This section describes how patients get on the waitlist, how kidneys are allocated, details of Conditions of Participation (CoP), and trends in kidney transplants.

2.1 Registration at Transplant Centers

The physician refers patients to a local transplant center when they have kidney failure. The center’s selection committee will evaluate if the patient is eligible for a kidney transplant (i.e., started dialysis or had a glomerular filtration rate (GFR) below 20mL per minute). The transplant center will then register accepted patients on the deceased donor waitlist and upload important information such as immunological profile, health conditions, and factors to compute priority into the UNet system(AKF, 2003).

2.2 Kidney Allocation and Transplant Process

The Organ Procurement and Transplantation Network (OPTN) designs and administers the deceased donor kidney allocation process. Hospitals upload a deceased donor’s medical history and organ condition into UNet when brain or cardiac death is imminent. The system

²Dialysis is a treatment that removes waste and excess water from the blood. There are two types of dialysis: hemodialysis and peritoneal dialysis.

³Kidney exchange is an alternative way of getting a kidney transplant (Roth et al., 2004). However, patients need a willing living donor, which can be logistically cumbersome. Hence, kidney exchange is considered a different program to deceased donor kidney transplant.

identifies biologically compatible patients and ranks them according to their priority order. Many factors contribute to the order, including, but not limited to, blood type, duration on the waitlist, where the patient lives, and, in some instances, weight and size compared to the donor⁴.

Recovered kidneys become unsuitable for transplants after 24 - 36 hours. UNet simultaneously contacts multiple transplant centers about their compatible patients to speed up the matching process. When contacted, a transplant center has 1 hour to decide which patient receives the kidney offer. During this hour, surgeons receive information about the donor’s medical history and can request additional information from the donor’s hospital. At the same time, surgeons also evaluate the patient’s health condition and decide if the patient is available or suitable for the transplant. For example, the patient’s condition might have deteriorated since the last evaluation, or the patient might be unavailable due to a family emergency. The transplant center does not contact every compatible patient because of the tight deadline⁵. It usually informs the patient after UNet confirms the center’s acceptance (Husain et al., 2019; King et al., 2023).

If UNet receives multiple acceptances, the center with the highest-priority patient will receive the kidney. After receiving the kidney, the center conducts a final blood test using samples from the patient and donor⁶. If the test results are satisfactory, the center proceeds with the transplant. Otherwise, the center declines the kidney offer, and UNet contacts the next center.

UNet removes the patient from the waitlist 24 hours after a successful transplant. In the case of a failed transplant or declining a kidney offer, UNet returns the patient to the waitlist

⁴In 2014, OPTN introduced "Longevity matching" to the kidney allocation system, by adding closer matching based on the age of donor and recipient. For example, a kidney from a 30-year-old donor is more likely to go to someone in the age range (OPTN, 2023).

⁵Furthermore, no regulations mandate transplant centers to notify patients of their kidney offers (OPTN, 2023)

⁶The blood test is called a serum crossmatch. It mixes the donor cells with the patient’s blood to determine if the antibodies will bind to the donor cell and destroy the kidney. Source: <https://www.kidney.org/atoz/content/BloodTests-for-Transplant>

without any penalty on their priority for the next kidney offer (OPTN, 2023).

Centers discharge transplant patients within 3 - 5 days and offer patients immunosuppressive drugs to prevent organ rejection. After the discharge, patients will visit the transplant center for regular check-ups. The transplant center informs UNet if the patient dies within 365 days after the transplant (OPTN, 2023).

2.3 Conditions of Participation (CoP)

Before July 2007, OPTN was the primary organization responsible for monitoring a transplant center’s number of post-transplant survival but only twice recommended to the Department of Health and Human Services to remove a transplant center’s certification (Stith and Hirth, 2016). Center for Medicare and Medicaid Services (CMS) became concerned that the lack of severe penalties for poor performance may have led to a decline in the quality of kidney transplants. As stated in the Final Rule establishing the increase in CMS oversight:

“ The OPTN generally takes a collegial approach and assists the center in improving their performance, while we generally take a regulatory approach which sometimes may lead to termination ...” (CMS, 2007)

CMS introduced CoP in May 2007 to provide a foundation for improving quality and protecting the health and safety of transplant patients (CMS, 2007). Transplant centers submit the 1-year post-transplant outcomes of a rolling 2.5-year cohort to the Scientific Registry of Transplant Recipients (SRTR) on the first week of every January and July. Figure 1 illustrates an example of a rolling 2.5-year cohort. The January 2011 submission (black box) consists of transplants from July 1, 2007, to December 21, 2009 (black line). Similarly, the July 2011 submission (red box) contains transplants from January 1, 2008, to June 31, 2010 (red line).

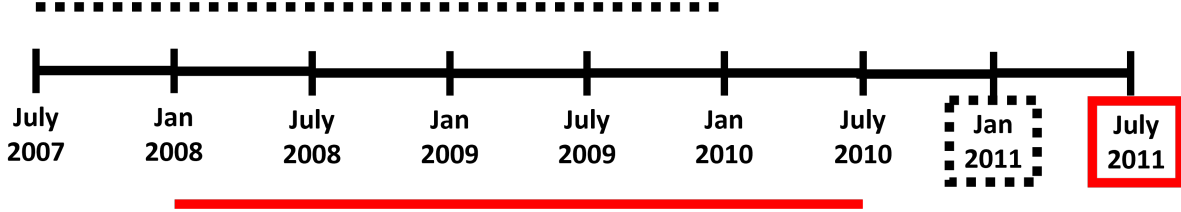


Figure 1: An illustration of the rolling 2.5-year cohort for CoP

SRTR measures a center’s performance by calculating the observed-expected (OE) 1-year death ratio. SRTR calculates expected deaths (E) by estimating a Cox regression model (Cox, 1972), using all the rolling 2.5-year cohorts submitted by each transplant center. The model uses extensive patient, donor, and match characteristics, including, but not limited to, age, race, diabetic status, donor cause of death, human leukocyte antigens (HLA) matching, etc. However, the model does not include center characteristics because *“center characteristics and practices may be associated with the differences that we are trying to identify and therefore should not be risk-adjusted away”* (Dickinson et al., 2008). Figure A2 in the Appendix shows a subset of variables used in estimating the model. SRTR updates the list of variables in the model every six months.

SRTR uses the estimated model to calculate a transplant center’s expected death, the sum of the 1-year expected death of each submitted patient-kidney pair in its rolling 2.5-year cohort, and obtain the final observed-expected (OE) death ratio. A transplant center has poor performance if all of the following criteria are satisfied:

1. OE ratio = $\frac{\text{Observed death (O)}}{\text{Expected death (E)}} > 1.5$
2. 1 sided p-value = $\Pr(\text{O-E} \geq 0) < 0.05$

The 1-sided p-value describes the probability that the observed difference is due to chance⁷. SRTR calculates the p-value by comparing the differences across all transplant centers in the U.S., accounting for the number of transplants by each center. The 5% critical value highlights Medicare’s tolerance of misclassifying a center as underperforming.

⁷The p-value calculation does not consider cases when expected failures exceed observed failures.

Medicare flags a transplant center for poor performance if it meets all the conditions above. Medicare then implements a data-driven quality assessment and performance improvement (QAPI) system. If the transplant center is flagged again within the next 30 months, it risks losing its program certification and Medicare funding.⁸.

2.4 Trends in Kidney Transplant

Figure 2 illustrates the significant change in the deceased donor market from 2003 - 2012. First, the post-transplant death rate (solid lines) has dropped by five percentage points. Second, total transplants (dash lines) increased from 2003 to 2007 and stagnated before rising again in 2010. Although these patterns coincide with the CoP's implementation, many factors can also explain the trends. For instance, medical technology is improving over time. Surgeons are improving at treating and identifying bad transplants (Thongprayoon et al., 2020; Hariharan, Israni and Danovitch, 2021). A routine before and after CoP comparison of the deceased donor market is insufficient to determine the causal effect of CoP. Hence, it motivates the difference-in-differences research design.

⁸However, most transplant centers have 210 days to appeal that their poor performances are due to mitigating circumstances.

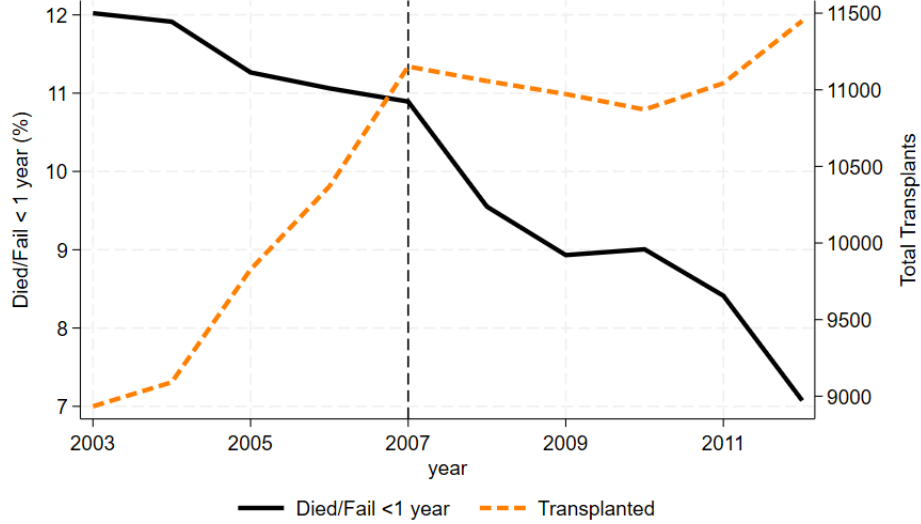


Figure 2: Post-transplant death rate and total transplants over 2003 - 2012. The vertical dash line indicates the introduction of CoP.

3 Data Sources and Prevailing OE ratio

I draw information on the universe of patient-kidney offers from the OPTN database from 2003 to 2012. My primary data is the Potential Transplant Recipient (PTR) file. It contains all kidney offers made to a patient on the waitlist and documents when the kidney offer arrived, the final decision, reasons for rejection (if applicable)⁹, and the patient's ranking on the kidney match. This data differs from the aggregate center-level data of the previous literature (Schold et al., 2013; White et al., 2015) that study behavior after CMS examination. My data lets me analyze how center behavior changes between intervening kidney offers as the prevailing OE ratio updates.

My second dataset is the Standard Transplant Analysis and Research (STAR) file managed by OPTN. It contains information on the patients' and deceased donors' demographics, health conditions, and immunological profiles. The STAR file has three advantages. First, it lets me control for important patient and kidney characteristics that affect a center's decision, which was missing from the previous literature that used aggregate center-level

⁹If a center declined the kidney due to a failed blood test. The data records it as a decline.

data. Secondly, STAR contains all the variables that CMS uses to calculate the expected death rate of a patient-kidney. This information helps construct the prevailing OE ratio between intervening offers. Thirdly, the STAR file contains measures for me to categorize patient-kidney pairs into risk groups¹⁰. I use this information in Section 5.0.1 to examine how center behavior changes for different risk groups.

My third dataset is the center-specific report (CSR) file. SRTR publishes CSR every six months, in the first week of January and July. Each CSR details the transplant center's performance and activity within six months¹¹. CSR also contains the statistical model and variables Medicare uses to calculate a transplant's expected death rate. Crucially, I combine the information from CSR and STAR in Section 3.1 to construct the OE ratio between intervening kidney offers of transplant centers. In the Appendix, I provide examples of CSR in Figure A1 and A2. I merged the three datasets to conduct my analysis at the patient-kidney offer level from 2003 - 2012.

3.1 Constructing prevailing OE ratio, $OE_{ct(k)}$

CSR reports the transplant center's final OE ratio in the first week of January and July. I construct a measure of the OE ratio between intervening kidney offers to analyze the threat of punishment on center behavior. I assume transplant centers use the latest CoP statistical model and the relevant 2.5-year rolling cohort to construct their OE ratio whenever they receive a kidney offer. I call this the prevailing OE ratio, $OE_{ct(k)}$ for center c when kidney k arrives on day t . This measure was motivated by conversations with surgeons who shared how their transplant center monitors performance after CMS introduced CoP¹².

Figure 3 illustrates how I construct $OE_{ct(k)}$ for a kidney k arriving at center c on $t = \text{April}$

¹⁰I discuss this in Appendix C

¹¹For example, CSR reports the number of transplants, the number of patients on the waitlist, and the number of patients removed from the waitlist due to death.

¹²CSR was only available after January 2007. So I use the January 2007 CSR to construct $OE_{ct(k)}$ for the pre-CoP sample.

28, 2010. If accepted, this offer joins the rolling 2.5-year cohort for center c 's July 2011 CSR. The rolling 2.5-year cohort in July 2011 consists of all transplants from January 2008 to July 2010. Thus, I use the January 2010 CSR (white arrow) to calculate the expected death rate for all the transplant outcomes from January 1, 2008, to April 27, 2010 (red solid line).

Next, I calculate the observed deaths for the same group (red solid line). If a transplanted patient is alive and has not met the 1-year mark (red dashed line), I assume centers regard this as a successful transplant and do not consider the patient in observed deaths. This assumption is motivated by conversations with surgeons who shared that they do not count alive patients who have not met the 1-year mark as deaths. In my data, approximately 10% of transplants die within one year, and 40% of these deaths happen within two months after the transplants¹³. Finally, we calculate $OE_{ct(k)}$ by taking the prevailing observed and expected-death ratio.

Figure 4 demonstrates a sample path of how OE_{ct} evolves over six months for a center c . The center c starts with $OE_{ct} = 1.40$ until the first event. A past transplanted patient died before the 1-year mark. Hence, OE_{ct} increases and jumps up in the graph. Conversely, if a transplanted patient dies after the 1-year mark, OE_{ct} does not change. In the second event, a kidney offer arrives, and the transplant center accepts the kidney for a patient. The transplant center performs the transplant and calculates the expected death rate according to CMS's latest model. As mentioned above, I assume the transplant center does not count this towards observed death. Hence, the expected death rate increases, and $OE_{ct(k)}$ drops in the figure¹⁴.

¹³I plot the hazard rate for deaths in Figure A3

¹⁴One concern about my method of calculating prevailing OE score is that it incentivizes transplant centers to perform a lot of transplants to artificially lower OE scores, especially when only 10% of transplants die within 1-year post-transplant. In my conversations with transplant surgeons,

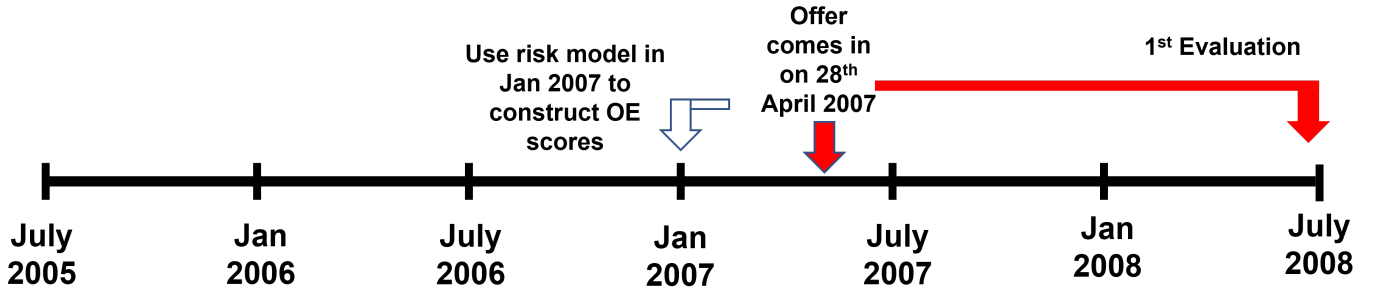


Figure 3: An illustration of how $OE_{ct(k)}$ is constructed

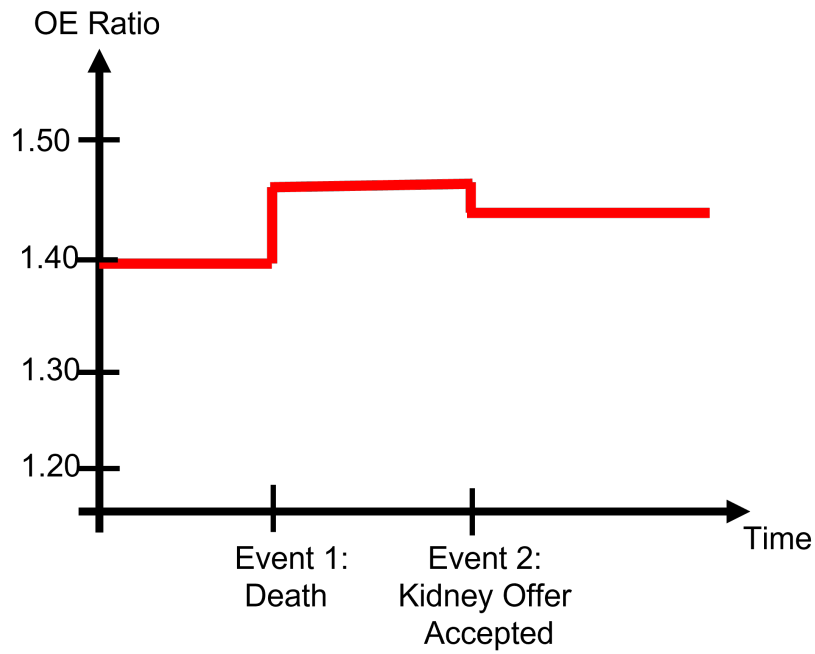


Figure 4: An sample path of OE_{ct} over a six months period

4 Research Design

My research design is a difference-in-differences framework. I compare the change in acceptance behavior of transplant centers pre and post-CoP according to their prevailing OE ratio, $OE_{ct(k)}$. The control group is $T_0 = \mathbb{1}\{OE_{ct(k)} < 0.5\}$. I assume transplant centers in T_0 are far from 1.5 and unlikely to exceed 1.5. Therefore, their incentives are the same before and after CoP because they do not face the threat of punishment. In my data, I have 1137 center-window observations with $OE_{ct(k)} < 0.5$ at the start of January or July. 2% of these centers eventually had $OE_{ct(k)} \geq 1.5$ at the end of June or December¹⁵.

Conversely, transplant centers with $OE_{ct(k)}$ close to 1.5 face different incentives before and after CoP. Previously, when past transplants died before the 1-year mark, it did not affect the center incentives because there was no threat of punishment. However, after CMS introduced CoP, these adverse outcomes pushed the $OE_{ct(k)}$ closer to 1.5, making it more likely for CMS to flag these centers for poor performance. In my main empirical specification, I divide the remaining $OE_{ct(k)}$ into groups of 0.1, forming 14 treatment groups. For example, a treatment group m is $T_{m(ct)} = \mathbb{1}\{OE_{ct(k)} \in [m - 0.1, m)\}$ for all $m \in \{0.6, 0.7, \dots, 1.9\}$.

4.1 Subsample

In my analysis, I do not use all the patient-kidney offers from 2003-2012. First, I drop observations in 2007 to avoid any anti In my analysis, I use the top 2 patients of every deceased donor. First, UNet only records the acceptance of the final recipient. UNet records the remaining acceptances as rejections. This misclassification is problematic because we might misinterpret some genuine acceptance as declining the kidney offer. I overcome this issue by relying on the top 2 patients of every kidney offer whose final response is independent of other patient's decision. The point is that kidneys are offered multiple times so the policy

¹⁵I provide the breakdown in the Appendix B1.

itself will impact how a kidney of a given quality is treated. .

4.2 Empirical Specification

My main empirical specification is:

$$\begin{aligned}
 Accept_{ickt} = & \alpha_0 + \sum_{m=0.6}^{1.9} \alpha_m \times T_{m(ct)} + \sum_{m=0.6}^{1.9} \beta_m \times T_{m(ct)} \times CoP_k \\
 & + \delta_{w(t)} + \gamma_c + \gamma_1 X_i + \gamma_2 Z_k + \varepsilon_{ickt}
 \end{aligned} \tag{1}$$

where $Accept_{ickt}$ indicates if center c accepts and transplant kidney k for patient i on day t . $\delta_{w(t)}$ is a six-month fixed effect that captures how acceptance pattern changes over time. γ_c is center fixed effects that capture time-invariant characteristics within a center. X_i and Z_k are characteristics of patient i and kidney k respectively. ε_{ickt} is the idiosyncratic error term, capturing exogenous logistical shocks that affect the center's acceptance decision. For example, the center may have a shortage of surgeons on a particular day, or the center could not contact the patient within the 1-hour deadline, or the patient-kidney pair did not pass the blood test and could not proceed with the transplant.

My parameter of interest is β_m . It measures the differential impact of CoP on the acceptance behavior of the treatment group m relative to the control group. We expect β_m to be negative if CoP affects the acceptance behavior of transplant centers.

There are two potential threats to identification. First, UNOS offers the kidney to multiple patients simultaneously to avoid organ spoilage and only records the acceptance of the final recipient. The remaining acceptances are recorded as rejections. This misclassification is problematic because we might misinterpret some genuine acceptance as declining the kidney offer. I address this issue by relying on the top 2 patients of every deceased donor because humans have two kidneys. Hence, the decisions of the top 2 patients are independent of the decisions of other patients.

Secondly, there might be concerns about anticipatory behavior before CoP. For instance, transplant centers may reject more transplants before CoP to give themselves a good head-start. I address this by dropping observations from 2007 in my pre-CoP sample.

Table 1 compares patient covariates for the top 100 patients of each deceased donor and my subsample. From Panel A, my subsample spent more time on the waitlist. This is unsurprising because time on the waitlist is a criterion for getting a higher priority in a kidney match. Since they spend more time on the waitlist, they are also less likely to have a working income or education. Furthermore, Panel B suggests patients in my subsample are healthier than the full sample.

All in all, there are differences between my subsample and the full sample. However, the richness of the STAR data lets me control for clinically relevant covariates in my regressions.

4.2.1 Summary statistics for treatment groups $T_{m(ct)}$ and $T_{0(ct)}$

Next, I describe the treatment $T_{m(ct)}$ and control $T_{0(ct)}$ groups. OE_{ctk} increases whenever a past transplant fails between intervening kidney offers for the center. In Figure 5, 20% of centers do not switch between treatment groups in a six-month period. The remaining 80% of centers switch between treatment groups at least once.

Table 1: Patient covariates summary statistics

	Top 100 patients	Top 2 patients + no 2007
Panel A: Medical Information		
Years on Waitlist	3.177 (2.575)	4.602 (4.356)
Age	51.85 (13.60)	48.19 (15.07)
% White	41.54 (49.28)	45.25 (49.77)
% Have working income	18.51 (38.84)	15.87 (36.54)
% Only High School	66.28 (47.28)	64.06 (47.98)
% Completed Uni.	17.20 (37.74)	15.58 (36.26)
% Medicare as primary insurer	59.57 (49.08)	60.28 (48.93)
Panel B: Demographics		
Body Mass Index (B.M.I)	27.87 (5.800)	26.82 (5.940)
Expected Post Transplant Survival	65.99 (28.94)	69.85 (29.24)
% Diabetic	34.95 (47.68)	26.51 (44.14)
% On dialysis	73.34 (44.22)	69.99 (45.83)
Observations	2242824	113009

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

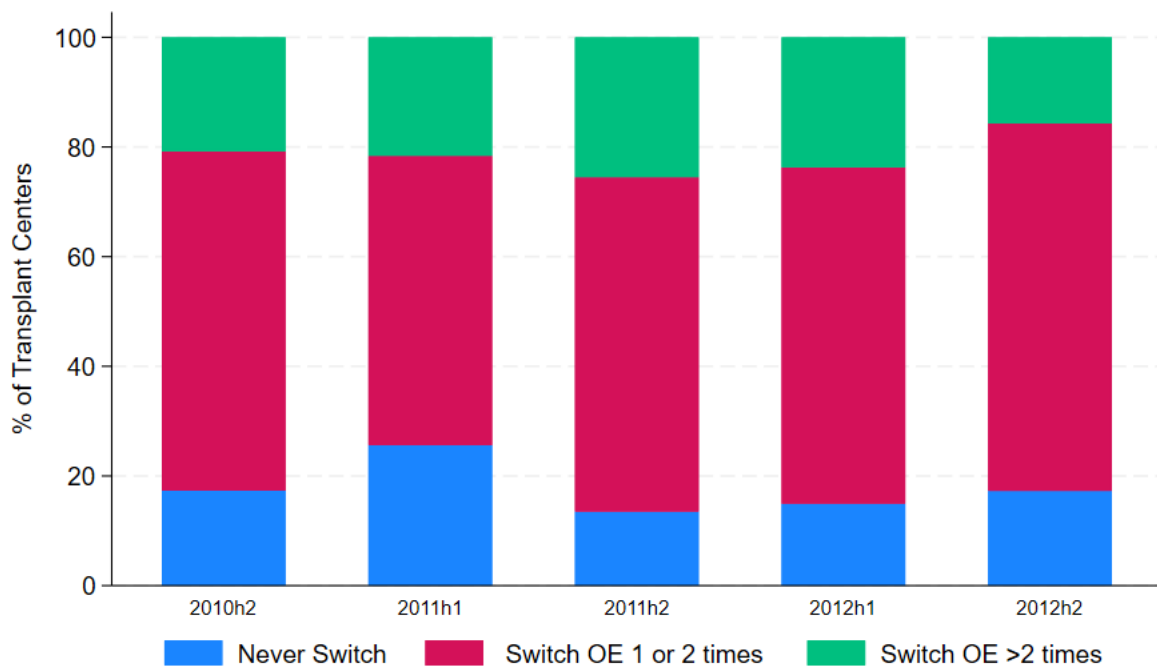
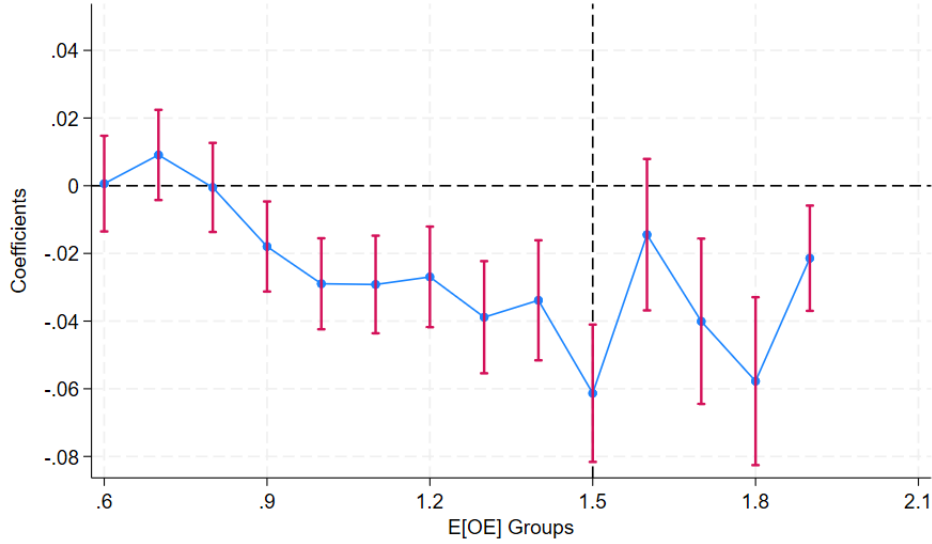


Figure 5: How often do transplant centers switch between treatment groups for selected six-months period.

5 Results for acceptance behavior

I plot the full range of β_m from equation 1 in Figure 6. Transplant centers are less likely to accept a kidney offer as their OE score approaches the CoP threshold. The results suggest that transplant centers consider their OE scores when accepting the kidney offer for the patient and are less likely to conduct transplants. A similar pattern persists even as we cross the threshold.

Figure 6: Acceptance behavior across different OE_{ck} groups



Note: The figure shows the acceptance behavior for each OE group m after CoP. We plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. We cluster standard errors at the center level.

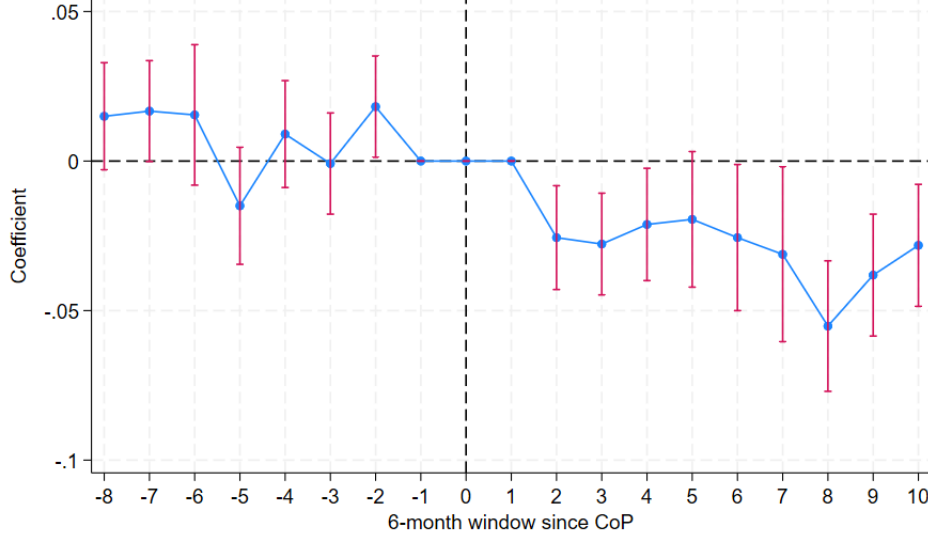
To interpret this estimate as the causal effect of CoP, we assume that CoP created a discontinuous change in center behavior and that any trends between the treatment and control groups are parallel in the absence of CoP. To support the parallel trends assumption, we

estimate the following events-study specification:

$$\begin{aligned}
Accept_{ick} = & \alpha_0 + \alpha \times T_{-0} + \sum_{s=-8, s \neq -1}^{10} \beta_{t-s} \times T_{-0} \times \delta_{t-s} \\
& + \sum_{s=-8, t \neq 1}^{10} \delta_{t-s} + \gamma_1 X_i + \gamma_2 X_k + \varepsilon_{ik}
\end{aligned} \tag{2}$$

We group all the treatment groups together and define $T_{-0} = \mathbb{1}\{OE_{ck} \geq 0.5\}$. The coefficient β_{t-s} measures the differential impact of CoP on the acceptance behavior of the treatment group relative to the control group. s measures the six-month periods relative to CoP implementation in July 2007. We plot the β_{t-s} coefficients in Figure 7 and find no evidence of pre-trends in acceptance behavior.

Figure 7: Acceptance behavior dynamic for T_{-0}



Note: The figure shows OLS estimates and 95% confidence intervals of the coefficients β_t from equation 2. We plot all coefficients relative to when CMS introduced CoP ($t=0$). We cluster standard errors at the center level.

5.0.1 Differences by risk profiles

Next, I extend the baseline specification to investigate if the effect of CoP on acceptance behavior differs by risk profiles. In my sample, I use information on patient risk to categorize patients into low and high risk. I follow a similar procedure for kidneys. Then, I group patient-kidney offers into low, medium, and high-risk transplants¹⁶. The regression equation in this case is:

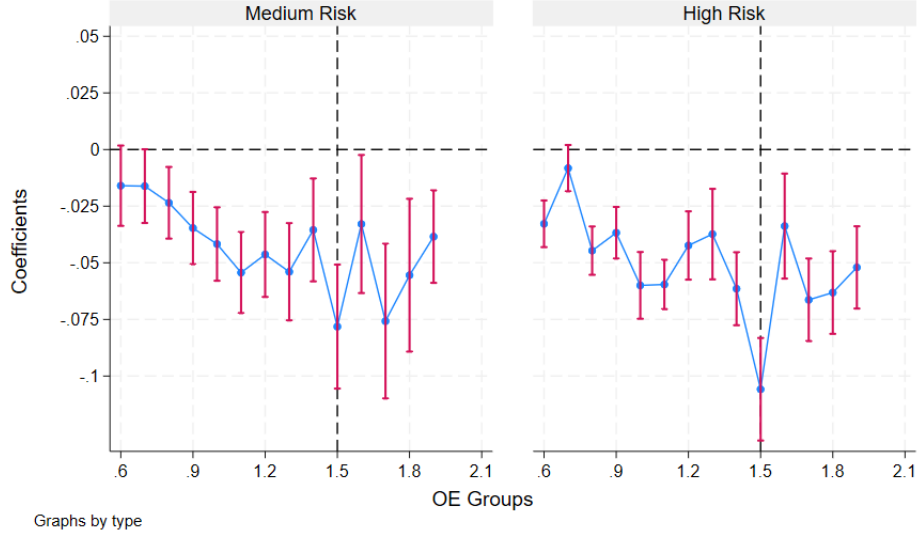
$$\begin{aligned}
Y_{ik} = & \sum_{r=med,high} \alpha_r \times R_{r(ik)} + \sum_{m=0.6}^{1.9} \alpha_m \times T_{m(ck)} \\
& + \sum_{r=med,high} \gamma_r \times R_{r(ik)} \times CoP_k + \sum_{m=0.6}^{1.9} \gamma_m \times CoP_k \times T_{m(ck)} \\
& + \sum_{r=med,high} \sum_{m=0.6}^{1.9} \gamma_{mr} \times T_{m(ck)} \times R_{r(ik)} \\
& + \sum_{r=med,high} \sum_{m=0.6}^{1.9} \beta_{mr} \times CoP_k \times T_{m(ck)} \times R_{r(ik)} \\
& + \sum_{t=1}^{20} \delta_t + \gamma_1 X_i + \gamma_2 Z_k + \varepsilon_{ick}
\end{aligned} \tag{3}$$

We have a triple difference-in-differences equation where we interact $R_{r(ik)}$, the risk indicator for patient i , kidney k with CoP_k and $T_{m(ck)}$. The omitted group for $R_{r(ik)}$ is low-risk transplants. Our parameter of interest is β_{mr} and measures the differential impact of CoP on the acceptance behavior of risk group r relative to the low-risk group. We plot the coefficients of β_{mr} in Figure 8.

We see that as OE scores approach the threshold, transplant centers are less likely to accept a medium/high-risk transplant than a low-risk one. The decrease is most significant at $m = 1.5$ for both risk types. The pattern persists even as we cross the threshold.

¹⁶Low-risk transplants are patient-kidney offers that are both low-risk. High-risk transplants are patient-kidney offers that are both high-risk. Medium-risk transplants are the remaining patient-kidney offers.

Figure 8: Acceptance behavior between different risk groups



Note: The figure shows the acceptance behavior for each risk group r after CoP. We plot the OLS estimates and 95% confidence intervals of the coefficients β_{mr} from equation 3. We cluster standard errors at the center level.

5.0.2 Differences by center size

Next, we use a similar specification to equation 3 to investigate if the effect of CoP on acceptance behavior differs by center size. The intuition is that OE scores are noisy estimates of transplant center performance. Some centers have a higher standard error due to the low volume of transplants.

We categorize transplant centers into low and high-volume centers based on the number of transplants they perform every six months from 2003 - 2012. We regard transplant centers below the median number of transplants as low-volume centers. The regression equation in

this case is:

$$\begin{aligned}
Y_{ick} = & \alpha \times LV_c + \sum_{m=0.6}^{1.9} \alpha_m \times T_{m(ck)} \\
& + \gamma \times LV_c \times CoP_k + \sum_{m=0.6}^{1.9} \gamma_m \times CoP_k \times T_{m(ck)} \\
& + \sum_{m=0.6}^{1.9} \delta_m \times T_{m(ck)} \times LV_c \\
& + \sum_{m=0.6}^{1.9} \beta_m \times CoP_k \times T_{m(ck)} \times LV_c \\
& + \sum_{t=1}^{20} \delta_t + \gamma_1 X_i + \gamma_2 Z_k + \varepsilon_{ick}
\end{aligned} \tag{4}$$

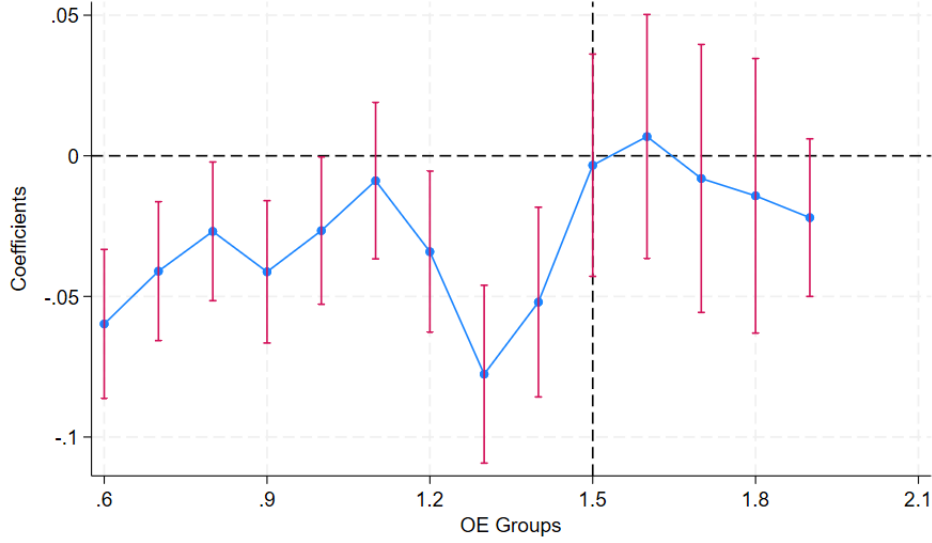
We have a triple difference-in-differences equation where we interact LV_c , the indicator for low-volume centers, with CoP_k and $T_{m(ck)}$. Our parameter of interest is β_m and measures the differential impact of CoP on the acceptance behavior of low-volume centers relative to high-volume centers. We plot the coefficients of β_m in Figure 9.

Before OE scores approach the threshold, low-volume centers are less likely to accept a patient-kidney offer than high-volume centers. However, the difference is not statistically significant once we reach or cross the threshold. The result differs from previous results where the effect of CoP is the greatest at the threshold. Low-volume centers seem more sensitive to CoP than high-volume centers before the threshold.

5.1 Patient Outcome

The richness of our data allows us to connect the changes in transplant center behavior to patient outcomes. We consider whether transplanted patients are more likely to die within 365 days after the transplant. We estimate equation 1 using only the set of transplanted patients as Section ?? but with Y_{ick} indicating whether patient i dies within 365 days after

Figure 9: Acceptance behavior for low-volume centers compared to high-volume centers



Note: The figure shows the acceptance behavior for low-volume centers after CoP. We plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 4. We cluster standard errors at the center level.

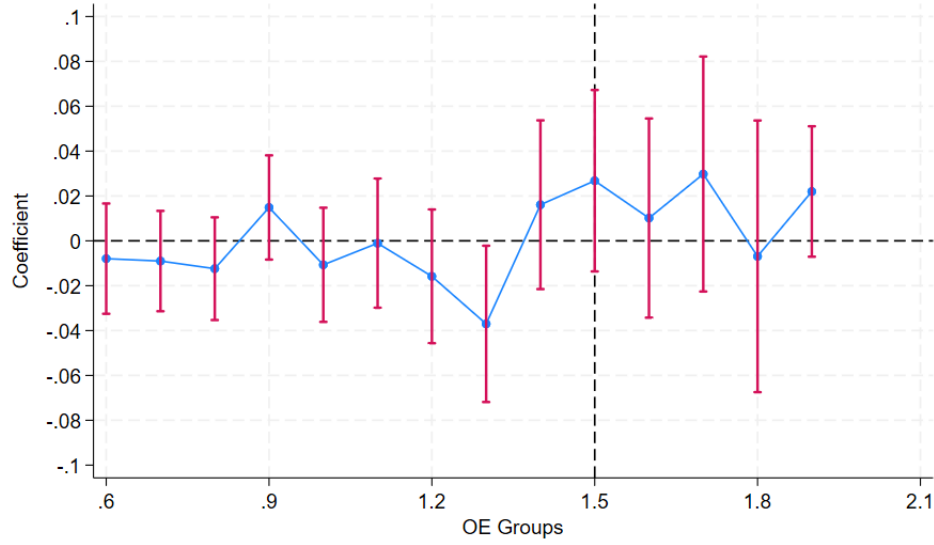
the transplant by center c with kidney k . We plot the full range of β_m in Figure 10.

Unlike the results above, we do not see a significant change in patient survival rates across the OE groups. There is no evidence to suggest that taking a kidney offer near or far from the threshold impacts patient survival.

Next, we examine if the effect of CoP on patient survival differs by risk profiles. We follow section 5.0.1 and categorize transplanted patient-kidney offers into low, medium, and high-risk. We use a similar specification to equation 3 to investigate if the effect of CoP on patient survival differs by risk profiles. We have Y_{ick} as an indicator for patient death within 365 days after the transplant for patient i with kidney k at center c . We estimate equation 3 for the subsample of transplanted patients and plot the coefficients of β_{mr} in Figure

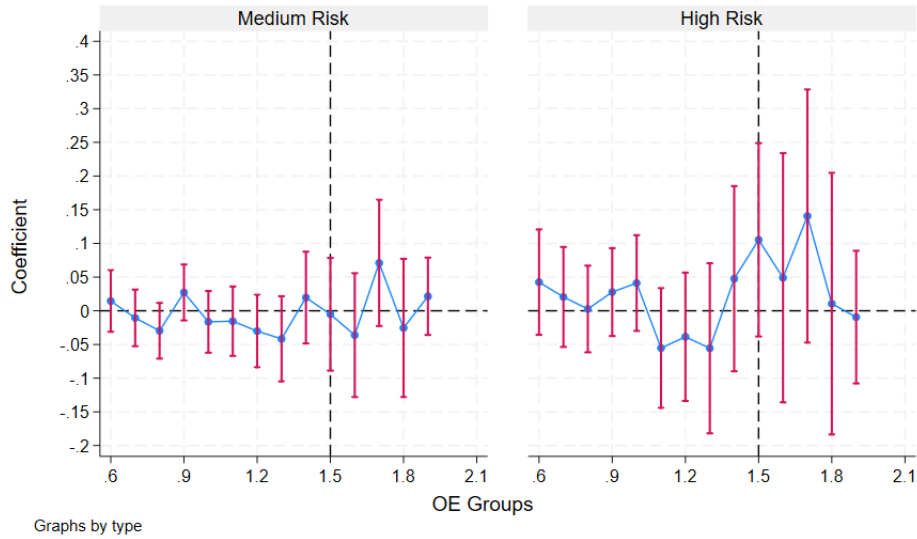
There is no evidence in Figure 11 to suggest accepting a medium/high-risk transplant at the threshold affects patient survival compared to a low-risk transplant

Figure 10: Transplanted patient survival across different OE_{ck} groups



Note: The figure shows the survival rate for each OE group m after CoP. We plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. We cluster standard errors at the center level.

Figure 11: Transplanted patient survival between different risk groups



Note: The figure shows the transplanted patient's survival for each risk group r after CoP. We plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 3. We cluster standard errors at the center level.

References

- Adler, Joel T., and David A. Axelrod.** 2016. “Regulations’ impact on donor and recipient selection for liver transplantation: how should outcomes be measured and MELD exception scores be considered?” *AMA Journal of Ethics*, 18(2): 133–142.
- American Kidney Fund (AKF).** 2008. “Deceased donor kidney transplants.”
- Centers for Medicare & Medicaid Services (CMS).** 2007. “Medicare program; hospital conditions of participation: requirements for approval and re-approval of transplant centers to perform organ transplants. Final rule.” *Federal Register*, 72(61): 15197–15280.
- Cox, D. R.** 1972. “Regression Models and Life-Tables.” *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2): 187–220.
- Dickinson, D. M., C. J. Arrington, G. Fant, G. N. Levine, D. E. Schaubel, T. L. Pruett, M. S. Roberts, and R. A. Wolfe.** 2008. “SRTR program-specific reports on outcomes: A guide for the new reader.” *American Journal of Transplantation*, 8(4 PART 2): 1012–1026.
- Hamilton, Thomas.** 2013. “Regulatory Oversight in Transplantation: Are the Patients Really Better Off?” *Current opinion in organ transplantation*, 18.
- Hariharan, Sundaram, Ajay K. Israni, and Gabriel Danovitch.** 2021. “Long-Term Survival after Kidney Transplantation.” *New England Journal of Medicine*, 385(8): 729–743. PMID: 34407344.
- Husain, S. Ali, Kristen L. King, Stephen Pastan, Rachel E. Patzer, David J. Cohen, Jai Radhakrishnan, and Sumit Mohan.** 2019. “Association Between Declined Offers of Deceased Donor Kidney Allograft and Outcomes in Kidney Transplant Candidates.” *JAMA Network Open*, 2(8): e1910312–e1910312.

- King, Kristen L., S. Ali Husain, Miko Yu, Joel T. Adler, Jesse Schold, and Sumit Mohan.** 2023. “Characterization of Transplant Center Decisions to Allocate Kidneys to Candidates With Lower Waiting List Priority.” *JAMA Network Open*, 6(6): e2316936–e2316936.
- Matas, Arthur J, and Mark Schnitzler.** 2004. “Payment for Living Donor (Vendor) Kidneys: A Cost-Effectiveness Analysis.” *American Journal of Transplantation*, 4(2): 216–221.
- Organ Procurement Transplantation and Network (OPTN).** 2023. “Policies.”
- Roth, Alvin, Tayfun Sönmez, Utku Ünver, and M. Ünver.** 2004. “Kidney Exchange.” *The Quarterly Journal of Economics*, 119: 457–488.
- Sack, Kevin.** 2012. “In Discarding of Kidneys, System Reveals Its Flaws.” *The New York Times*.
- Schold, J.D., L.D. Buccini, T.R. Srinivas, R.T. Srinivas, E.D. Poggio, S.M. Flechner, C. Soria, D.L. Segev, J. Fung, and D.A. Goldfarb.** 2013. “The Association of Center Performance Evaluations and Kidney Transplant Volume in the United States.” *American Journal of Transplantation*, 13(1): 67–75.
- Stith, Sarah S., and Richard A. Hirth.** 2016. “The Effect of Performance Standards on Health Care Provider Behavior: Evidence from Kidney Transplantation.” *Journal of Economics & Management Strategy*, 25(4): 789–825.
- Thongprayoon, Charat, Panupong Hansrivijit, Napat Leeaphorn, Prakrati Acharya, Aldo Torres-Ortiz, Wisit Kaewput, Karthik Kovvuru, Swetha Kanduri, Tarun Bathini, and Wisit Cheungpasitporn.** 2020. “Recent Advances and Clinical Outcomes of Kidney Transplantation.” *Journal of Clinical Medicine*, 9(4): 1193.

White, Sarah L., Dawn M. Zinsser, Matthew Paul, Gregory N. Levine, Tempie Shearon, Valarie B. Ashby, John C. Magee, Yi Li, and Alan B. Leichtman. 2015. “Patient Selection and Volume in the Era Surrounding Implementation of Medicare Conditions of Participation for Transplant Programs.” *Health Services Research*, 50(2): 330–350.

Appendices

A Supplementary Figures

<u>Line</u>		<u>Center</u> <u>1 Year</u>	<u>National</u> <u>1 Year</u>
	<u>Adult (Age 18+)</u>		
1	Transplants (n=number)	90	10,781
2	Percent (%) of Patients Surviving at End of Period		
3	Observed at this Center	87.78	86.26
4	Expected, based on national experience	89.41	
5	Deaths During Follow-up Period		
6	Observed at this center	11	1,392
7	Expected, based on national experience	8.48	1,392
8	Ratio: Observed to Expected (O/E)	1.30	1.00
9	(95% Confidence Interval)	(0.65-2.32)	
10	P-value (2-sided), observed v. expected	0.469	
11	How does this center's survival compare to what is expected for similar patients?	Not Significantly Different (a)	
12	Percent retransplanted	5.5	4.4
13	Follow-up days reported by center (%)	91.7	93.9
14	Maximum Days of Follow-up (n)	365	365

Figure A1: A page of a July 2007 CSR report from (Dickinson et al., 2008)

Deceased Donor Graft Survival Model Description
1 Year (and 1 Month) after Transplant
Organ: Kidney
Adult (Age 18+)

77.0% graft functioning at 1 Year when all covariates=0. 91.3% graft functioning at 1 month when all covariates=0.
The indexes of concordance are 66.7%, 66.6%, and 66.7%, respectively.

Characteristic Covariates	Reference Group	CSR Cohort Released 07/13/2010 Transplants between 01/01/2007 and 06/30/2009				Models Used 07/13/2010 for Transplants between 07/01/2006 and 12/31/2008 (see note)				Models Used 07/13/2010 for Transplants between 01/01/2006 and 06/30/2008 (see note)			
		beta	hazard ratio	standard error	p-value	beta	hazard ratio	standard error	p-value	beta	hazard ratio	standard error	p-value
Cold ischemia time: continuous (per 1 hour)		0.007696	1.01	0.0026	0.0017	0.009014	1.01	0.0026	0.0005	0.010633	1.01	0.0026	<0.0001
Cold ischemia time: missing		0.301176	1.35	0.1069	0.0046	0.370824	1.45	0.0979	0.0001	0.424076	1.53	0.0928	<0.0001
Deceased donor kidney was pumped: missing	no	0.372378	1.46	0.2605	0.1671	0.366577	1.44	0.2521	0.1459	0.470208	1.6	0.2907	0.1057
Deceased donor kidney was pumped: yes	no	-0.177109	0.84	0.0567	0.0025	-0.121486	0.89	0.0559	0.0299	-0.117816	0.89	0.0544	0.0320
Deceased donor with history of diabetes: missing	no	-0.560706	0.57	0.4069	0.1712	-0.758628	0.47	0.4485	0.0907	-0.385138	0.66	0.3553	0.2783
Deceased donor with history of diabetes: yes	no	0.208408	1.23	0.0779	0.0072	0.301400	1.35	0.0725	<0.0001	0.353804	1.42	0.0705	<0.0001
Deceased donor with history of hypertension: yes	no or missing	0.179440	1.2	0.0569	0.0023	0.133714	1.14	0.0567	0.0184	0.138068	1.15	0.0550	0.0120
Donation after cardiac death: yes	no	0.426267	1.53	0.0745	<0.0001	0.302956	1.35	0.0735	<0.0001	0.317205	1.37	0.0728	<0.0001
Donor Age (age minus 25): applies to >25		0.051956	1.05	0.0087	<0.0001	0.041346	1.04	0.0083	<0.0001	0.037606	1.04	0.0081	<0.0001
Donor Age: applies to all		-0.034400	0.97	0.0070	<0.0001	-0.023252	0.98	0.0067	0.0005	-0.019913	0.98	0.0065	0.0022
Donor DSA different from recipient DSA: yes	no	0.056520	1.06	0.0645	0.3811	0.007174	1.01	0.0626	0.9088	-0.008390	0.99	0.0617	0.8918
Donor meets expanded donor criteria for deceased donor kidney: yes	no	0.108870	1.12	0.0802	0.1744	0.020155	1.02	0.0765	0.7922	0.019764	1.02	0.0744	0.7905
Donor race: Asian	White	0.124338	1.13	0.1444	0.3892	0.161514	1.18	0.1360	0.2462	0.152541	1.16	0.1369	0.2638
Donor race: Black	White	0.137923	1.15	0.0670	0.0396	0.162123	1.16	0.0838	0.0111	0.123103	1.13	0.0835	0.0526
Donor race: Hispanic/Latino	White	-0.038753	0.98	0.0741	0.8199	-0.017490	0.98	0.0704	0.8940	-0.006520	0.98	0.0881	0.9239
Donor race: multi-racial, other, unknown or missing	White	-0.029859	0.97	0.2908	0.9182	0.247144	1.28	0.2450	0.3130	0.422332	1.53	0.2320	0.0687
Donor serum creat (centered at 1.3 mg/dL): applies to >1.3		-0.470761	0.62	0.2033	0.0205	-0.547042	0.58	0.1963	0.0053	-0.680845	0.51	0.1967	0.0005
Donor serum creatinine (per 1 mg/dL): applies to >0.9		0.652311	1.92	0.3277	0.0406	0.783702	2.19	0.3157	0.0131	0.837594	2.31	0.3111	0.0071
Donor serum creatinine (per 1 mg/dL): applies to all		-0.115782	0.89	0.1977	0.5581	-0.216504	0.81	0.1914	0.2577	-0.196728	0.82	0.1892	0.2983
Donor serum creatinine: missing		-0.001372	1	0.3599	0.9970	-0.085849	0.92	0.3583	0.8106	-0.007098	0.99	0.3418	0.9839
Donor to recipient weight ratio: continuous		-0.018267	0.98	0.0829	0.8256	-0.070278	0.93	0.0795	0.3769	-0.158743	0.88	0.0779	0.0415
Donor to recipient weight ratio: missing		0.288293	1.33	0.1414	0.0415	0.369775	1.45	0.1356	0.0064	0.383945	1.47	0.1282	0.0028
Donor: deceased, COD cerebrovascular/stroke	other	0.217328	1.24	0.0557	0.0001	0.191546	1.21	0.0531	0.0003	0.210852	1.23	0.0523	0.0001
HLA mismatch: 0 ABDR mismatch	non-0 ABDR, 2DR mismatch	0.081112	1.08	0.1128	0.4721	0.002800	1	0.1091	0.9795	-0.097304	0.91	0.1069	0.3629

Note: Coefficients for these cohorts are shown to compare stability of the models across time. These coefficients differ from those actually used in previous CSRs because the data and models may have changed since the release of those CSRs.
* Reference for diagnosis group is glomerular diseases.

Figure A2: An example of the risk adjustment model in the CSR report

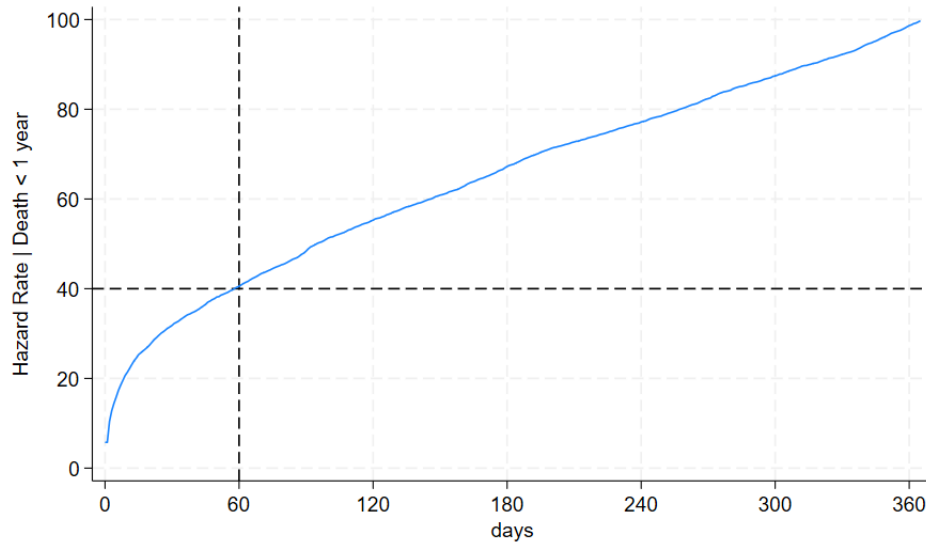


Figure A3: Hazard rate conditional on transplant death within 1-year. The intersection of the dash lines indicates 40% of transplants that die within 1-year die within 60 days/2 months after the transplant.

B Supplementary Tables

Table B1: Distribution of OE ratio in June or December if $OE < 0.5$ in January or July

	Frequency	Percent	CDF
$OE < 0.5$	869	76.43	76.43
$OE \in [0.5, 1.0)$	228	20.05	96.48
$OE \in [1.0, 1.5)$	7	0.616	97.10
$OE \geq 1.5$	33	2.902	100
N	1137		

Table B2: Distribution of OE ratio in June or December if $OE < 0.5$ in January or July

	(1)	(2)	(3)
$\beta_{0.6}$	0.00342 (0.00732)	-0.000886 (0.00706)	-0.00470 (0.00722)
$\beta_{0.7}$	0.0129 (0.00690)	0.00121 (0.00666)	-0.0133 (0.00688)
$\beta_{0.8}$	0.00487 (0.00683)	-0.0120 (0.00658)	-0.0323*** (0.00680)
$\beta_{0.9}$	-0.0126 (0.00691)	-0.0209** (0.00666)	-0.0353*** (0.00688)
$\beta_{1.0}$	-0.0237*** (0.00699)	-0.0318*** (0.00673)	-0.0443*** (0.00700)
$\beta_{1.1}$	-0.0223** (0.00750)	-0.0381*** (0.00722)	-0.0546*** (0.00751)
$\beta_{1.2}$	-0.0163* (0.00771)	-0.0236** (0.00744)	-0.0337*** (0.00773)
$\beta_{1.3}$	-0.0365*** (0.00861)	-0.0457*** (0.00829)	-0.0541*** (0.00862)
$\beta_{1.4}$	-0.0319*** (0.00921)	-0.0289** (0.00887)	-0.0363*** (0.00917)
$\beta_{1.5}$	-0.0661*** (0.0106)	-0.0654*** (0.0102)	-0.0709*** (0.0105)
$\beta_{1.6}$	-0.0275* (0.0116)	-0.0122 (0.0112)	-0.0233* (0.0115)
$\beta_{1.7}$	-0.0525*** (0.0127)	-0.0388** (0.0122)	-0.0522*** (0.0125)
$\beta_{1.8}$	-0.0424*** (0.0129)	-0.0464*** (0.0124)	-0.0586*** (0.0127)
$\beta_{1.9}$	-0.00793 (0.00805)	-0.0182* (0.00778)	-0.0312*** (0.00858)
Center FE			✓
6-months period FE	✓	✓	✓
Pat. and Kid. Controls		✓	✓
Observations	244051	244051	244050

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C Constructing Risk Profiles

The STAR file from Section 3 measures kidney risk with kidney donor profile index (KDPI) for all deceased donors. It combines a variety of donor factors into a single number that summarizes the likelihood of graft failure after a deceased donor kidney transplant. The KDPI runs from 0 to 100, with higher values indicating a higher risk of graft failure. I follow (Adler and Axelrod, 2016) and define high-risk kidneys as those with $KDPI > 50$ and low-risk kidneys as those with $KDPI \leq 50$ ¹⁷.

Similarly, I also use information on patient characteristics to measure patient risk with estimated post-transplant survival (EPTS). It is a numerical measure used to allocate kidneys in 2014 after "Longevity Matching" was introduced. EPTS scores range from 0 to 100. Candidates with lower EPTS scores are expected to experience more years of graft function from high-longevity kidneys. So, I define high-risk patients as those with $EPTS > 50$ and low-risk patients as those with $EPTS \leq 50$ ¹⁸.

I combine these two pieces of information to define the risk profile for a patient-kidney offer. I consider a patient-kidney offer low(high)-risk if both the patient and kidney are low(high)-risk. A patient-kidney offers medium risk if the patient is low-risk and the kidney is high-risk or vice versa.

¹⁷Source:<https://optn.transplant.hrsa.gov/data/allocation-calculators/kdpi-calculator/learn-about-kdpi/>

¹⁸Source:<https://optn.transplant.hrsa.gov/data/allocation-calculators/epts-calculator/learn-about-epts/>