Performance Scores and Strategic Choices of Kidney Transplant Centers*

Han Ng†

this version: August 2024

Click here for the latest draft.

Abstract

Does the threat of punishment make hospitals reduce treatments and increase patient mortality? To answer this question, I analyze the decisions of U.S. kidney transplant centers, where a 2007 reform threatened to decertify a center if the observed-expected (OE) death ratio of transplanted patients exceeds a given threshold. Using variation in distance to the OE threshold, I show that the reform reduced the probability of transplanting a patient-kidney pair. The reduction is more significant for riskier operations and low-volume centers. I find little change in aggregate patient mortality. My results suggest centers avoid transplants for performance gains, and the reform's effect is mainly redistributive.

JEL codes: I11, I18, L38

Keywords: quality regulation, kidney transplant, mortality rates, nonprofit policy, medicare

*I am very grateful for the guidance and encouragement from Paul Grieco and Conor Ryan. I thank Marc Henry, Karl Schurter, Mark Roberts, and the participants of Applied Micro Brownbag 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 the outcomes of patient operations. 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 operations 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 are worried about potential punishment.

Three details in my setting motivate my findings. First, transplant centers can track their prevailing OE ratio before the deadline. CMS publishes the risk model for calculating expected death. Even though CMS updates the model in six-month intervals, the latest model serves as a reasonable proxy for centers to estimate and monitor their current performance[1]. Secondly, transplant centers have a lot of discretion in accepting or declining a kidney offer for a patient. When the allocation system identifies a compatible match between a patient

[1]During my conversation with surgeons, they shared that when Medicare first introduced CoP, transplant centers used this approach to track their performance.

and a kidney, it contacts the transplant center directly. However, the transplant center does not have to inform the patient if it declines the kidney offer on the patient's behalf(Husain et al., 2019; King et al., 2023). Thirdly, transplant centers do not immediately observe the deaths of past transplants preceding the 1-year mark. This uncertainty is challenging for centers approaching the threshold. When they accept a kidney offer for a patient, they risk crossing the threshold and being flagged for poor performance if either the patient or past transplants die within a year[2].

These details suggest transplant centers can and will decline kidney offers for patients whenever they think their OE ratio may exceed the CoP threshold. Previous reports have described how the threat of government penalties made doctors more selective about the organs and patients they accepted. For instance, in a 2012 *New York Times* article, Dr. Lloyd E. Ratner, the director of Columbia Hospital, said: *"... 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 above anecdotal evidence, I collected data on patient-kidney offers and CMS's risk model from 2003 - 2012 to construct a transplant center's prevailing OE ratio. I use this information to investigate whether the threat of punishment induces transplant centers to decline kidney offers for patients. To cleanly link it to Medicare's CoP policy, I must overcome some empirical challenges. I do not observe the many unobservable factors influencing a center's behavior, such as center quality, which plausibly correlates with the OE ratio (i.e., Better centers have a lower OE ratio.). I use a difference-in-difference research design. I compare the behavior of transplant centers in the same OE ratio before and after CoP's introduction. The comparison produces a causal effect on the policy because I can use the centers with a low OE ratio as the control group. The policy plausibly did not affect their behavior since they were far from the threshold. In many cases, I also estimate specifications

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

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. My patient-kidney offers data with precise patient and kidney characteristics, unlike previous literature (Schold et al., 2013; White et al., 2015; Stith and Hirth, 2016) using center-level observation. This feature allows me to examine how CoP affects the behavior of different patient-kidney pairs.

I find transplant centers declining patient-kidney offers even when the prevailing OE ratio is below 1.5. The probability of accepting a kidney offer drops by 22% at the threshold. Perhaps reflecting the threats at stake, transplant centers decline more medium and high-risk patient-kidney pairs than low-risk pairs. I interpret these results as imperfect risk adjustment in the CoP statistical model. Even though the model includes an extensive list of patient kidney characteristics, these calculations may not sufficiently compensate transplant centers for the downside of taking on riskier patient-kidney pairs. For example, I expect the variance of the unexplained portion of post-transplant death to be higher for the riskier patient-kidney pairs (Volk et al., 2017). A transplant center approaching the threshold would not take on these riskier transplants. Secondly, sample size matters. Low-volume centers are less likely to accept a patient-kidney pair than high-volume centers. OE ratio is a noisy estimate of center quality. A small sample size increases the standard errors (Lunsford, Prakash and Guarrera, 2022) and exacerbates the center's selection incentive when the OE ratio precedes 1.5.

Finally, there are two potential mechanisms where CoP affects patient mortality. First, CoP incentivizes centers to improve service quality, reducing post-transplant deaths. Second, CoP cherry-picks patients for performance gains, leaving behind patients and increasing deaths off the waitlist. I examine the net effect of these two potential countervailing mechanisms by estimating the effect of CoP on the 1-year mortality of all patients who received a kidney offer. I find little evidence that aggregate patient mortality decreased.

Related Literature: My paper contributes to the literature on quality disclosure and certification in the healthcare sector. (Bundorf et al., 2009; Ramanarayanan, 2011; Vatter, 2023) discussed how performance scores address the imperfect information in healthcare and change patient demand for services. My results differ because I focused on how healthcare providers' selective incentives respond to performance scores. (Dranove et al., 2003) find that cardiac surgery report cards in New York and Pennsylvania led both to provider selection behaviors, leading to higher levels of resource use and worse health outcomes. My results differ in two ways. First, I show the intensity of a provider's selective behavior depends on the proximity to the performance threshold. Second, I also demonstrate that selective behavior depends on the size of the provider. Providers with smaller sample sizes are more selective due to the high standard errors in their performance measures.

Second, my paper contributes to the literature on mechanism design in deceased donor kidney transplants. Zhang (2010); Agarwal et al. (2021) assumes the incentives of transplant centers are aligned with the patients when designing counterfactual allocation mechanisms. My paper provides evidence that suggests otherwise.

Finally, 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 deadline, because they are worried about potential punishments. On the other hand, Stith and Hirth (2016) uses a difference-in-differences approach and found no reduction in the post-transplant death rate after CMS flagged the centers for poor performance. I extend this result by considering mortality off the waitlist and show the CoP policy had little effect on aggregate patient mortality.

The rest of the paper proceeds as follows. Section 2 summarizes important institutional details about the U.S. deceased donor kidney transplant program. Section 3 describes my

data. Section 4 presents my research design and empirical strategy. Section 5 presents results on the transplant center's acceptance behavior. Section 6 discusses the effect of CoP on patient mortality. Section 7 concludes.

2 Background on Deceased Donor Kidney Transplant

A patient diagnosed with end-stage renal disease (ESRD) has two options: dialysis or kidney transplant [3]. 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)[4]. This section describes how patients get on the waitlist, how the centralized system allocates kidneys, 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).

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

[4]Kidney 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.

2.2 Kidney Allocation and Transplant Process

The Organ Procurement and Transplantation Network (OPTN) designs and administers the centralized 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 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 donor[5].

Recovered kidneys become unsuitable for transplants after 24 - 36 hours. So, 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 deadline[6]. 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 donor[7]. If the test results are satisfactory, the center proceeds with

[5]In 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).

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

[7]The 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:

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 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

https://www.kidney.org/atoz/content/BloodTests-for-Transplant

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).

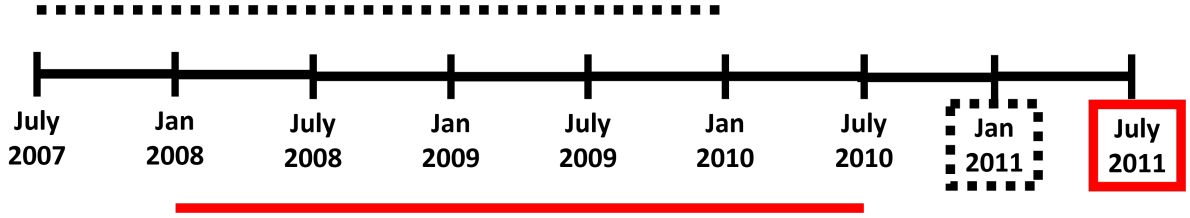


Figure 1: An illustration of the rolling 2.5-year cohort for CoP

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 $= \Pr(\text{O-E} \geq 0) < 0.05$

1.5 is the CoP threshold. It is when observed death exceeds expected death by 50%. The 1-sided p-value describes the probability that the observed difference is due to chance[8]. 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. My primary analysis focuses on constructing the OE ratio because it is the primary metric used to flag transplant centers for poor performance. Even though the p-value sometimes exempts centers crossing the threshold, surgeons do not monitor it closely because they do not have information on the performance of other centers.

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.[9].

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.

[8]The p-value calculation does not consider cases when expected failures exceed observed failures.

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

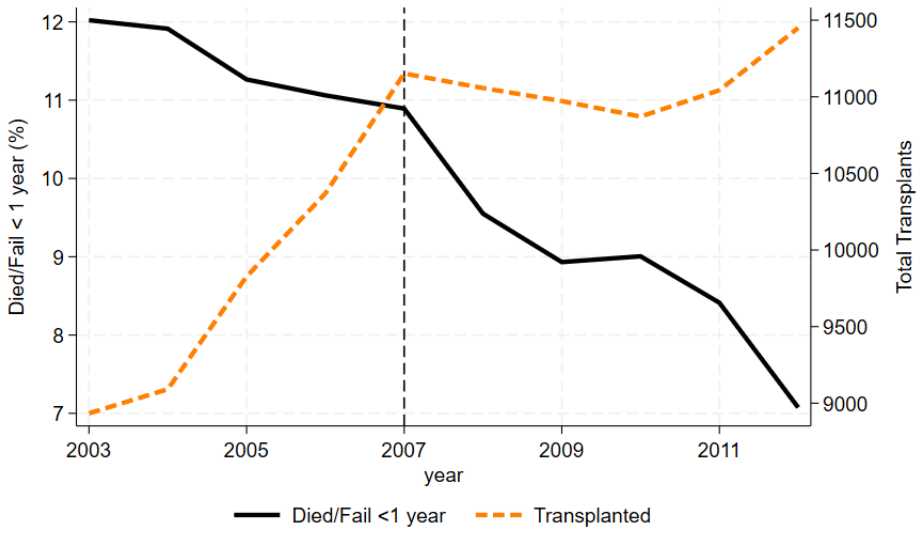


Figure 2: Post-transplant death rate and total transplants over 2003 - 2012. The vertical dash line indicates the introduction of CoP.

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 declining (if applicable)[10], 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

[10]If 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 groups[11]. I use this information in Section 5.1 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 months[12]. 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 rolling 2.5-year 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 CoP[13].

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

[11] I discuss this in Appendix C

[12] For 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.

[13] CSR 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, as depicted in Figure 1. 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.

Next, I calculate the observed deaths for the same group. If a transplanted patient is alive and has not met the 1-year mark, 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 transplants[14]. 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 figure [15].

[14] I plot the hazard rate for deaths in Figure A3

[15] One concern about my method of calculating prevailing OE score is that it incentivizes transplant centers to perform a lot of transplants in the short run to artificially lower OE ratio. This approach is not optimal. Accepting multiple transplants in the short run can hurt the future OE ratio if many of these transplants fail within a year and stay on multiple rolling 2.5-year cohorts.

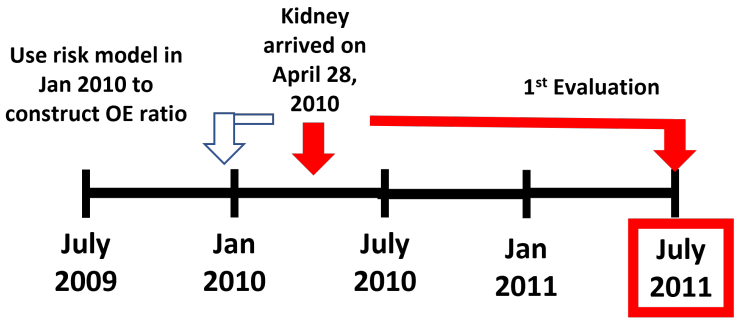


Figure 3: An illustration of how $OE_{ct(k)}$ is constructed

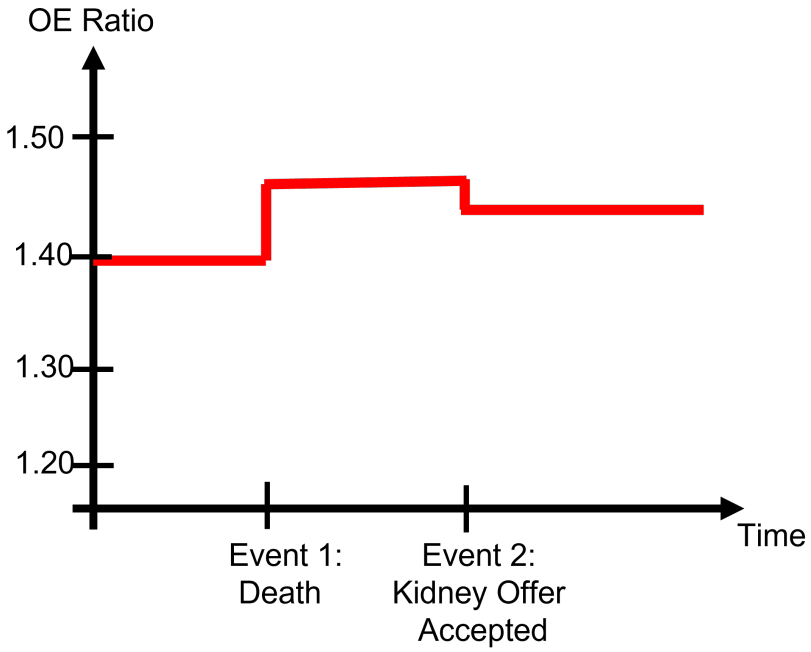


Figure 4: A sample path of OE_{ct} over a six months period

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, $OE_{ct(k)}$. My 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 December[16].

Conversely, transplant centers with $OE_{ct(k)}$ close to 1.5 face different incentives before and after CoP. Before CoP, 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 $m \in \{0.6, 0.7, ...1.9\}$. This division captures the extent of the transplant center's strategic behavior according to proximity to the CoP threshold.

My data from Section 3 runs from 2003 - 2012. CMS introduced CoP in May 2007 and published the first grade in July 2007. I define observations between January 1, 2003, and June 31, 2007, as the pre-CoP period. Similarly, observations between July 1, 2007, and December 31, 2012 forms the post-CoP period. I define $CoP_{t(k)}$ as an indicator for kidney k arriving on day t in the post-CoP period.

[16] I provide the breakdown in the Appendix B1.

4.1 Sample Restrictions

I do not use all the patient-kidney offers from 2003-2012. I restrict my sample in the following way. First, I dropped January 2007 to June 2007 observations to avoid concerns of anticipatory behavior. Second, I omit July 2007 to December 2007 observations to prevent problems of transplant centers adjusting to the new CoP policy. Third, I use the top 2 patients of every deceased donor. The centralized system offers kidneys to patients after rejections from preceeding patients. Transplant centers may reject kidneys if they perceive previous rejections as a poor quality signal. We do not know how the policy affects the perceived quality of kidneys in the post-CoP period, so the policy will impact how centers perceive kidneys of a given quality. Looking at the top 2 patients of every deceased donor ensures that the acceptance decisions are independent of the decisions of other centers. Finally, we look at the top 2 patients because every deceased donor has two kidneys.

Table 1 compares the average patient covariates for my subsample and the top 100 patients of each deceased donor. Both samples look similar on most covariates. There are differences in years on waitlist: time on the waitlist is a criterion for getting higher priority in a kidney match. Patients from my sample are healthier. I expect these differences because some patients do pre-emptive listing, queueing for a kidney before their kidney function deteriorates and needs dialysis (Kiberd, Tennankore and Vinson, 2022). About 31.2 % of kidney offers were accepted in the restricted sample

4.2 Summary Statistics of $OE_{ct(k)}$

I present summary statistics of the prevailing OE ratio, $OE_{ct(k)}$. Figure 5 presents the pre and post-CoP density plot for $OE_{ct(k)}$. I want to highlight two patterns in this figure. First, the post-CoP density (solid lines) has more weights in the $[1, 1.5)$ region than the pre-CoP density (dashed line). I interpret this as suggestive evidence that transplant centers are more

Table 1: Patient covariates summary statistics

	Top 2 patients + no 2007	Top 100 patients
Panel A: Demographic Information		
Years on Waitlist	4.193	3.177
	(3.662)	(2.575)
Age	48.62	51.85
	(15.16)	(13.60)
% White	43.19	41.54
	(49.53)	(49.28)
% Have working income	16.24	18.51
	(36.88)	(38.84)
% Only High School	65.35	66.28
	(47.59)	(47.28)
% Completed Uni.	15.73	17.20
	(36.41)	(37.74)
% Medicare as primary insurer	60.86	59.57
	(48.81)	(49.08)
Panel B: Medical Information		
Body Mass Index (B.M.I)	27.07	27.87
	(5.973)	(5.800)
Expected Post Transplant Survival (EPTS)	30.89	34.01
	(29.40)	(28.94)
% Diabetic	28.15	34.95
	(44.97)	(47.68)
% On dialysis	70.88	73.34
	(45.43)	(44.22)
Observations	285381	2242824

mean coefficients; sd in parentheses

Note: EPTS range from 0 to 100. Lower EPTS score indicates the patient is expected to experience more years of kidney function from a high-longevity kidney. Appendix C provides more details.

likely to reject transplants when $OE_{ck(t)}$ is close to 1.5, resulting in some form of "bunching" in the post-CoP period. Secondly, there is not a lot of weight in the $[2, 3)$ region for both periods. I interpret this as most transplant centers doing their best to ensure they do not have too many unexpected deaths.

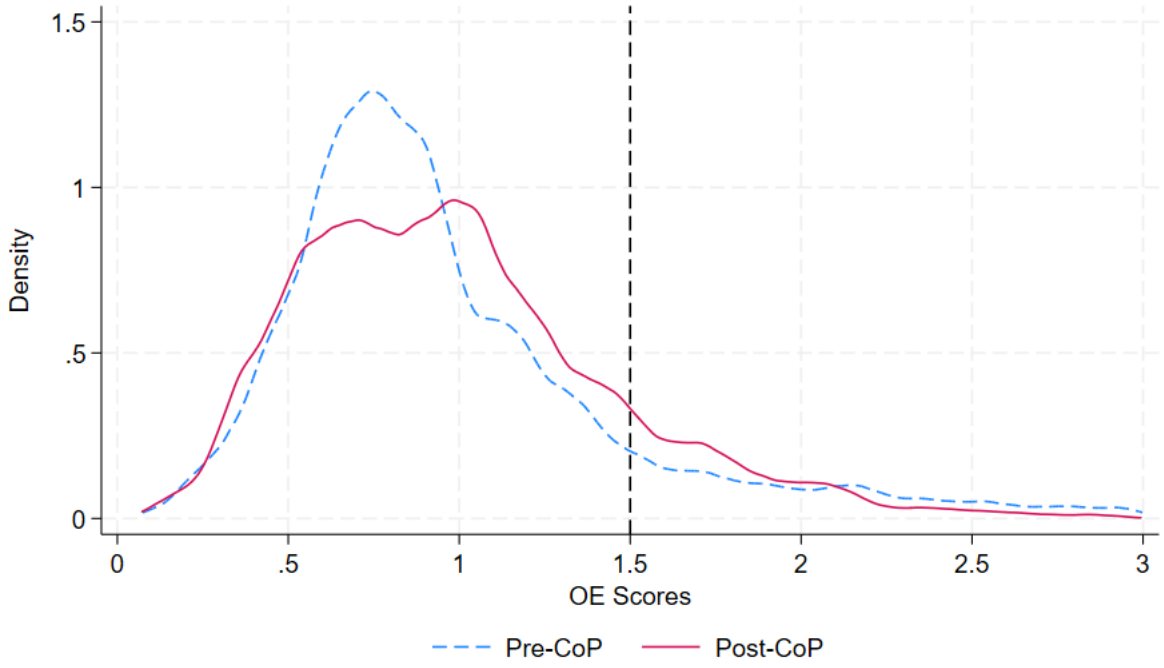


Figure 5: Kernel density estimate of $OE_{ct(k)}$. The vertical dash line is the CoP threshold.

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 6, 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. The lack of switchers is problematic in my setting because when I use center-fixed effects in equation 1, the within-center variation might be insufficient to identify the β_m coefficients. So, I present robustness checks where I estimate β_m from within-center and cross-centers variation.

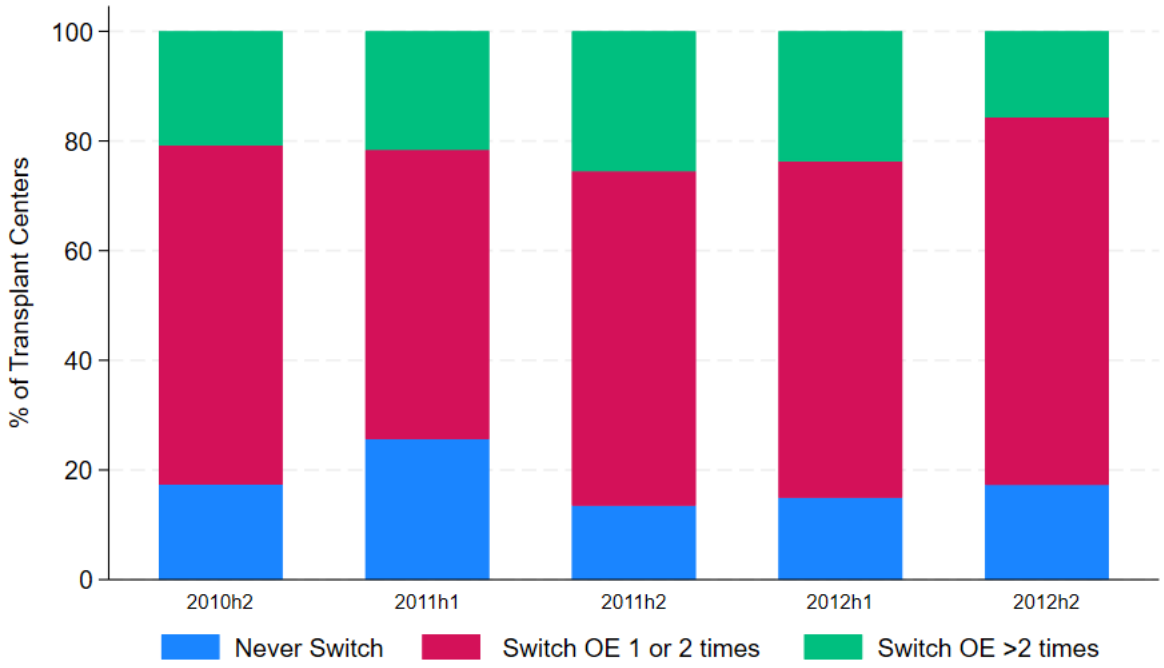


Figure 6: How often do transplant centers switch between treatment groups for selected six-months period.

4.3 Empirical Specification

My main empirical specification is:

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

where $Accept_{ickt}$ indicates if center c accepts and transplants kidney k for patient i on day t. $CoP_{t(k)}$ is an indicator for kidney k arriving in the post-CoP period; $T_m(ct)$ is an indicator treatment group m, all defined at the start of Section 4. $\delta_{w(t)}$ 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. For $m \leq 1.5$, I expect β_m to be negative because transplant centers are worried about potential punishment and do not want to exceed the threshold. For $m > 1.5$, I still expect $beta_m$ to be negative. Even though the center has exceeded the threshold, these $OE_{ct(k)}$ are just approximations of the true OE published later. So, the center still wants to decline transplants to avoid punishment.

One of the main threats to identification is unobserved confounders. I address this concern in two ways. First, I use center fixed effects γ_c to control for time-invariant center characteristics. Secondly, the richness of the STAR data lets me control for clinically relevant covariates in my regressions.

To interpret β_m as a causal effect, I assume that CoP created a discontinuous change in center behavior and that any trends between the treatment and control groups are parallel. To support the parallel trends assumption, I estimate the following events-study specification on my sample:

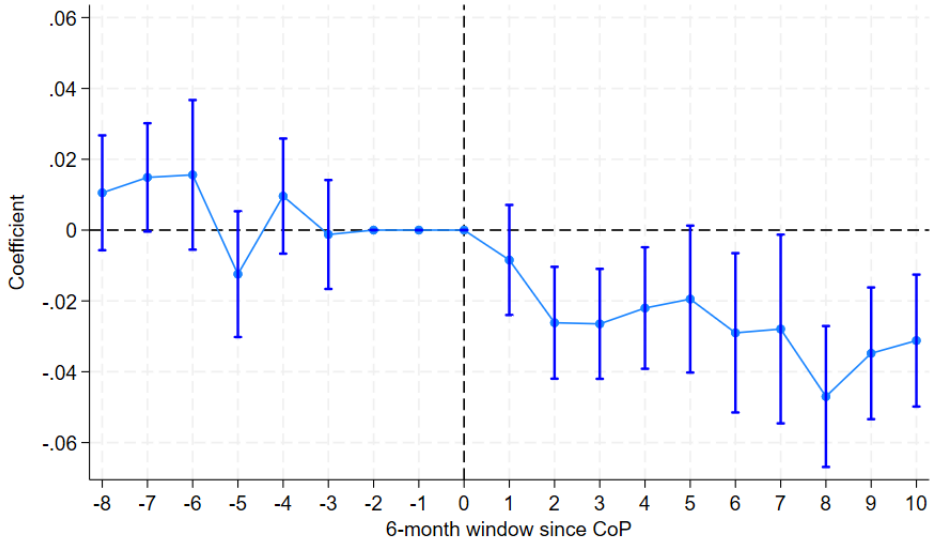
$$\begin{aligned} Accept_{ickt} = \alpha_0 + \alpha \times T_{-0(ct)} + \sum_{s \in [-8,10]} \mu_s \times T_{-0(ct)} \times \mathbb{1}\{w(t) - \mathbb{T}_{CoP} = s\} \\ + \delta_{w(t)} + \gamma_c + \gamma_1 X_i + \gamma_2 X_k + \varepsilon_{ickt} \end{aligned} \tag{2}$$

w denotes a six-month interval, and $\mathbb{T}_{CoP}$ is the first six-month interval Medicare implemented CoP, July 2007 - December 2007. I group all the treatment groups as $T_{-0(ct)} = \mathbb{1}\{OE_{ct(k)} \geq 0.5\}$ and examine the dynamic effect of CoP on center behavior. The rest of the variables are defined analogously as equation 1.

I plot the dynamic effect coefficients μ_s in Figure 7. The results suggest no significant

evidence of differential pre-treatment trends. The coefficients on $s \in [-8, -3]$ are small and insignificant, meaning that the treatment and control groups do not have different acceptance trends pre-policy. There are no coefficients for $s = -1, 0$ because these correspond to the 2007 observations I dropped from my sample. Most of the μ coefficients for $s \in [1, 10]$ are negative and significant. The fact that the magnitude is persistent throughout the post-policy period suggests that transplant centers adjusted their acceptance behavior very quickly. In the Appendix, I define $T_{-0} = \mathbb{1}\{OE_{ct(k)} \geq 1.0\}$ and run equation 2 as a robustness check. I find a similar pattern but with more noise in Figure A4.

Figure 7: Impact of CoP on dynamic acceptance behavior



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

5 Results on Acceptance Behavior

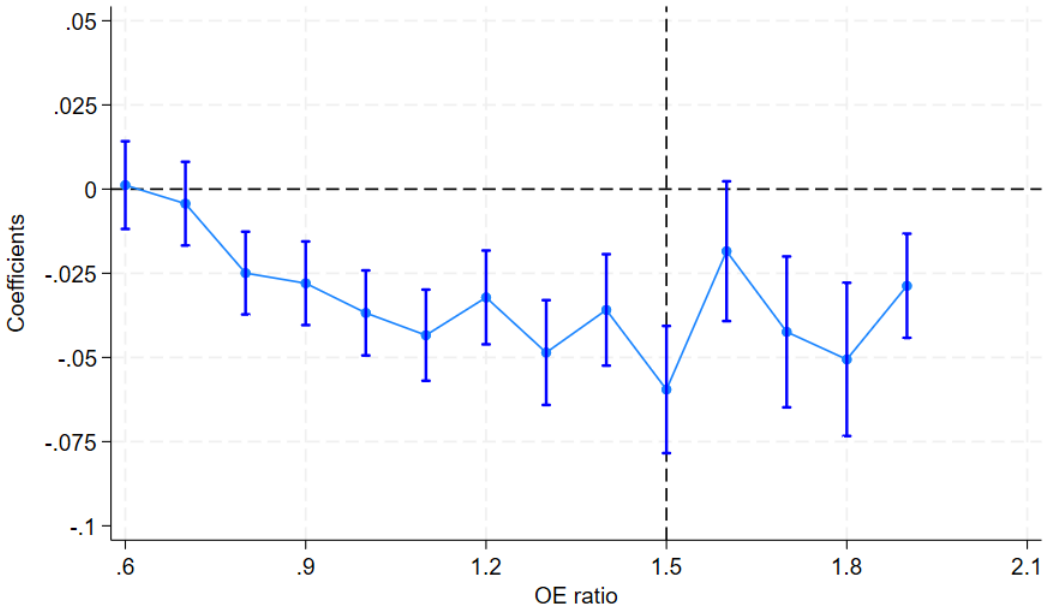
The results in Figure 7 suggest CoP affected transplant centers' acceptance behavior. But transplant centers were quick to learn, and the effect persisted. Next, I use equation 1 to investigate whether the proximity of a transplant center's prevailing OE ratio affected its

acceptance behavior differently. I plot the full range of β_m in Figure 8.

The probability of accepting a kidney offer decreases as a transplant center's prevailing OE score approaches the threshold. The decrease is most significant at the CoP threshold, with a decrease of 6.1 percentage points or a 22% drop relative to the mean acceptance rate. The previous literature focuses exclusively on the centers exceeding the threshold after the CoP deadline. My result highlights how CoP affects all transplant centers before the deadline, even for those below the threshold.

Interestingly, the pattern persists even as the prevailing OE ratio crosses the threshold. I interpret this as transplant centers being cautious about the accuracy of the prevailing OE ratio. The prevailing OE ratio estimates the actual OE ratio published later. The center may not know if it has exceeded the threshold until the following report. So, the center still wants to decline transplants to avoid punishment.

Figure 8: Acceptance behavior across different $OE_{ct(k)}$ groups



Note: The figure shows the acceptance behavior for each OE group m after CoP. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. I cluster standard errors at the center level. The point estimates are presented in Table B2 of the Appendix.

5.1 Differences by Risk Profiles

The previous section highlights transplant centers are declining patient-kidney pairs as performance worsens. So, I investigate if this behavior comes from a particular patient-kidney risk profile. If it is, does CoP threshold proximity matter? The STAR files provide information on patient and kidney risk, respectively. I use this information to group patient-kidney offers into low-, medium-, and high-risk transplants. Low(High)-risk transplants have both patients and kidneys as low(high)-risk. Medium-risk transplants are the remaining patient-kidney offers[17]. I refer interested readers to Appendix C for more details on the patient and kidney risk measures.

I estimate equation 1 with the new risk profile definitions. I have the same control group T_0 and break down patient-kidney offers in the treatment groups into low, medium, and high-risk. I compare the β_m coefficient for the low-risk and medium(high)-risk groups in Figure 9a (9b).

Figure 9 shows that transplant centers are generally less likely to accept a medium/high-risk patient-kidney offer than a low-risk patient-kidney offer. For $m \in [0.6, 1.2]$, there is a significant gap between the probability of accepting low-risk and medium/high-risk transplants. But the gap is not significant for $m \in [1.3, 1.9]$. The results suggest that when the prevailing OE ratio is close to the threshold, transplant centers have performance concerns and generally decline transplants regardless of patient-kidney risk profile.

The findings in $m \in [0.6, 1.2]$ are intriguing because risk adjustment was expected to make transplant centers indifferent across the risk profiles. Yet, I see centers preferring low-risk transplants. I interpret this as an imperfect risk adjustment in the CoP statistical model. Even though the model includes an extensive list of patient kidney characteristics, the calculations may not sufficiently compensate transplant centers for the downside of taking on

[17]This grouping is only for exposition purposes. In the Appendix, I present results for low-risk kidneys with high-risk patients and high-risk kidneys with low-risk patients.

riskier patient-kidney pairs. For example, the variance of the unexplained portion of the post-transplant death is higher for the riskier patient-kidney pairs. A transplant center approaching the threshold would not take on these riskier transplants.

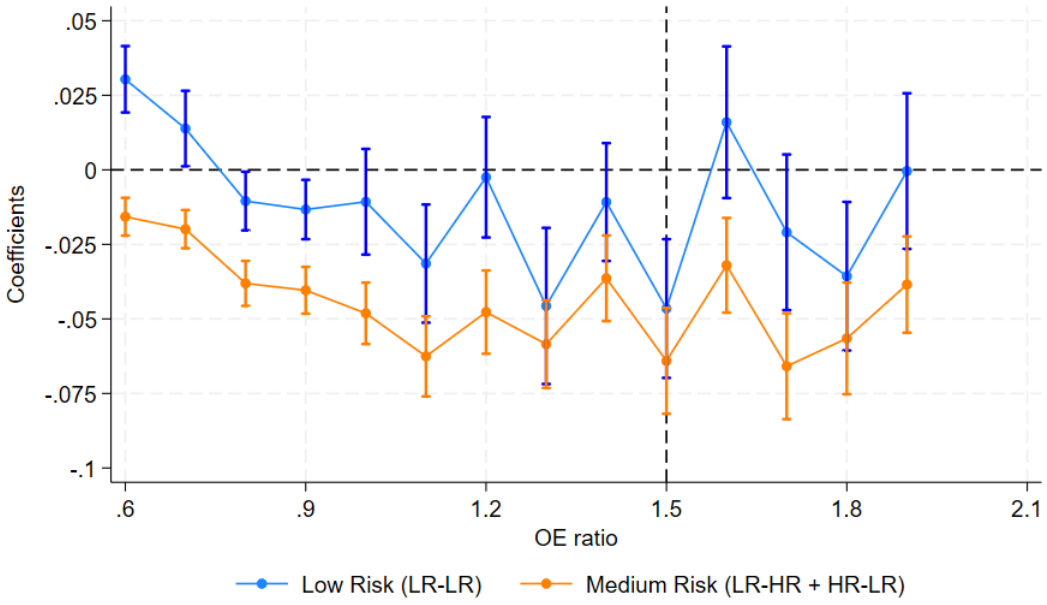
5.2 Differences by Center Size

OE scores are noisy estimates of transplant center performance. Some centers have a higher standard error due to the low volume of transplants. In this section, I investigate if the effect of CoP on acceptance behavior differs by center size. I categorize transplant centers into low- and high-volume centers based on the number of transplants performed every six months from 2003 to 2012. I regard transplant centers below the median number of transplants as low-volume centers. The regression equation in this case is:

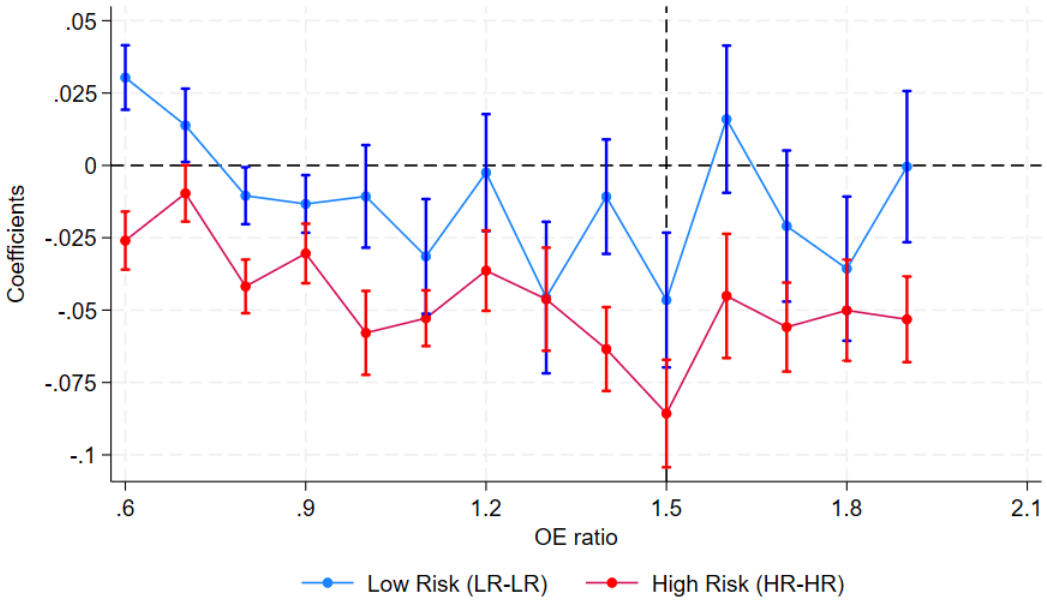
I estimate equation 1 with the new center size definitions. I have the same control group T_0 and break down patient-kidney offers in the treatment groups into low- and high-volume centers. I compare the β_m coefficient for the low- and high-volume centers in Figure 10.

Figure 10 shows that low-volume centers are less likely to accept a patient-kidney offer than a high-volume center. The effect is consistent throughout the different OE ratio. The results suggest low-volume centers are more sensitive to the OE ratio than high-volume centers. I interpret this as the low-volume centers having a higher standard error in their OE ratio. The higher standard error exacerbates the selection incentive when the OE ratio is close to the threshold.

Figure 9: Acceptance behavior across different $OE_{ct(k)}$ groups for different risk profiles



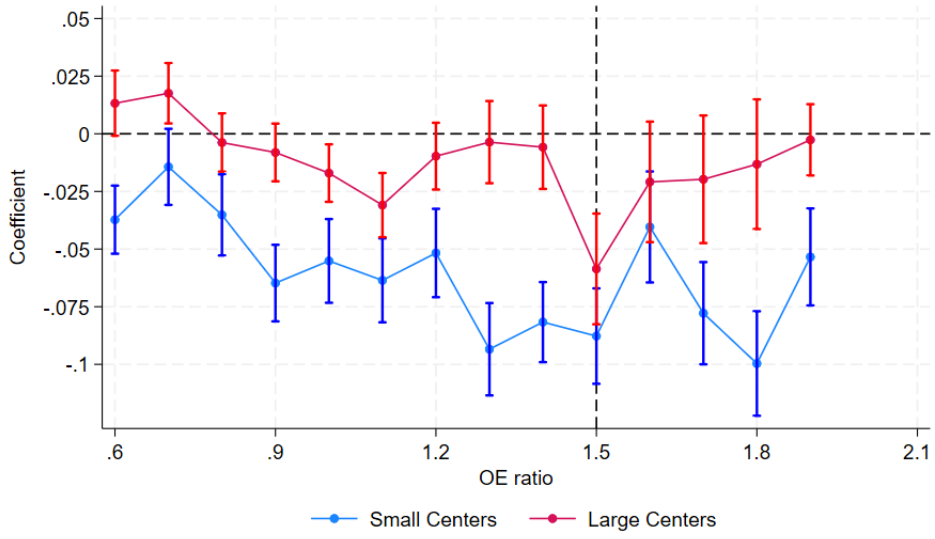
(a) Low-Risk v.s. Medium-Risk



(b) Low-Risk v.s. High-Risk

Note: The figure shows the acceptance behavior for each OE group m after CoP for different risk profile. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. I cluster standard errors at the center level. The point estimates are presented in Table B3 of the Appendix.

Figure 10: Acceptance behavior across different $OE_{ct(k)}$ groups for different center volume



Note: The figure shows the acceptance behavior for each OE group m after CoP for different center volume. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. I cluster standard errors at the center level. The point estimates are presented in Table B4 of the Appendix.

6 Results on Patient 1-year Mortality

Medicare implemented CoP to incentivize transplant centers to improve patient mortality Hamilton (2013). CoP affects patient mortality through two potential mechanisms: (i) better quality of care reduces post-transplant deaths, and(ii) selective transplant reduces treatment and increases deaths on the waitlist. The direction of the net effect of these channels is an empirical question.

I examine the net effect using the same research design in Section 5, replacing the acceptance decision in equation 1 with patient mortality. Instead of using the top 2 patients of every deceased donor as described in Section 4.1, I use all patients and their first kidney offer. I estimate the following empirical specification:

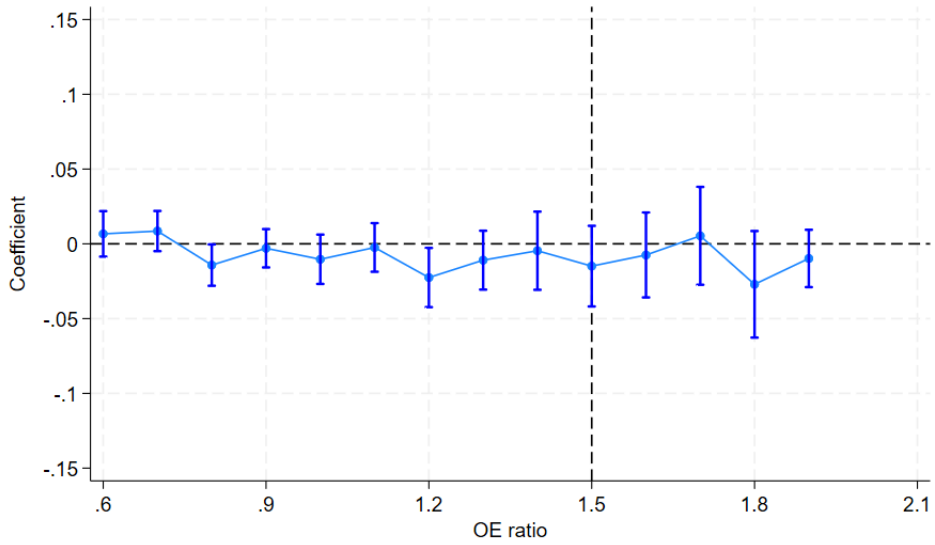
$$\begin{aligned} Death_{ickt} = \alpha_0 + \sum_{m=0.6}^{1.9} \alpha_m \times T_{m(ct)} + \sum_{m=0.6}^{1.9} \beta_m \times T_{m(ct)} \times CoP_k \\ + \delta_{w(t)} + \gamma_c + \gamma_1 X_i + \gamma_2 Z_k + \varepsilon_{ickt} \end{aligned} \tag{3}$$

$Death_{ickt}$ indicates if patient i at center c died within 1 year after receiving their first kidney offer k on day t. $CoP_{t(k)}$ is an indicator for kidney k arriving in the post-CoP period; $T_m(ct)$ is an indicator treatment group m, all defined at the start of Section 4. $\delta_{w(t)}$ is a six-month fixed effect that captures how the death 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. The parameter of interest is β_m. I expect β_m to be negative if CoP meets its original objective of improving patient outcomes.

I plot the full range of β_m in Figure 11. For $m \in [0.6, 1.9]$, β_m are small and statistically insignificant. The exception is $m = 1.2$. Overall, I conclude that the net effect of CoP on aggregate patient mortality is minimal. Next, similar to section 5.1, I divide patient-kidney pairs into low-, medium-, and high-risk and examine if there are differences in β_m across risk groups. I plot the results in Figure A6 and do not see any differences in mortality effect

across risk groups.

Figure 11: 1-year mortality after first kidney offer across different $OE_{ct(k)}$ groups



Note: The figure shows the 1-year mortality after the first kidney offer for each OE group m after CoP. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 3. I cluster standard errors at the center level. I present the point estimates in Column 1 of Table B6.

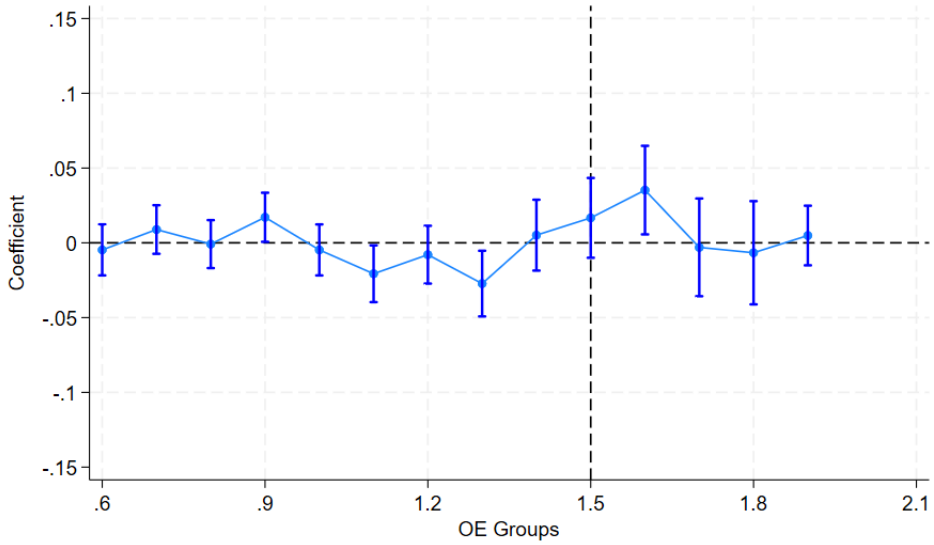
6.1 Decomposing the Net Effect

Next, I examine what drives the minimal net effect on patient mortality. Are the two countervailing mechanisms producing significant effects but canceling each other, or do they each have minimal effects? I estimate the effect of CoP on the post-transplant death rate by examining the mortality of transplanted patients. I use the exact empirical specification in equation 3. $Death_{ickt}$ now indicates if patient i at center c died within 1-year after undergoing a transplant with kidney k on day t.

The endogeneity concern is that centers select patient-kidney pairs into transplants. There are unobserved patient and kidney characteristics correlated with the OE ratio that influence post-transplant death (i.e., ε_{ickt} correlates with $T_m(ct)$). I address this concern by taking

advantage of the STAR dataset mentioned in Section 3 that provides me with detailed patient and kidney characteristics. I include these characteristics in X_i and Z_k to control for potential selection bias. I plot all β_m in Figure 12. For $m \in [0.6, 1.9]$, β_m are mostly small and statistically insignificant. I interpret this result as suggestive evidence that the effect of CoP on post-transplant mortality is minimal.

Figure 12: Post-transplant 1 year deaths across different $OE_{ct(k)}$ groups



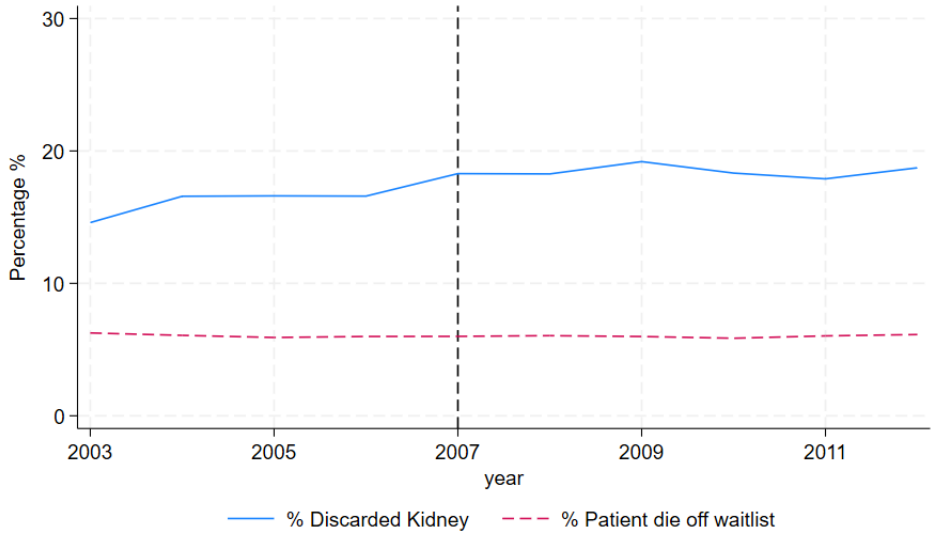
Note: The figure shows the post-transplant 1-year death rate for each OE group m after CoP. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 3. I cluster standard errors at the center level. I present the point estimates in Columns 1 of Table B5.

Next, I present descriptive evidence in Figure 13 that suggests patient mortality off the waitlist has been stable over the sample period. Results in Figure 12 and 13 suggest that the minimal net effect of CoP on patient mortality is because both mechanism has minimal mortality effect. So, this begs the question: why are there not more patients dying off the waitlist if CoP makes centers reject more kidney offers, as discussed in Section 5?

The solid lines in Figure 13 show that the percentage of discarded kidneys has been relatively stable over the sample period. I interpret this as evidence that although CoP makes transplant centers more selective, the rejected kidney trickles down the waitlist and is eventually

accepted by a transplant center whose OE ratio is not near the CoP threshold. This redistribution potentially benefits patients at the bottom of the ranking list who would otherwise not have received a kidney offer without CoP.

Figure 13: Post-transplant 1 year deaths across different $OE_{ct(k)}$ groups



Note: The figure shows that the percentage of discarded kidney and patients dying off the waitlist has been relatively stable over the sample period from 2003 to 2012.

7 Conclusion

This paper evaluates Medicare's CoP policy in the deceased donor kidney transplant program. My first result suggests that transplant centers are forward-looking and strategic in their behavior before submitting patient outcomes to Medicare for evaluation. Centers are more likely to decline patient-kidney offers or reduce treatments when the threat of punishment is high (i.e. OE ratio approaches CoP threshold). This result differs from previous literature that only examined the behavior of centers after submitting outcomes to Medicare and suggests strategic behavior is more prevalent than previously thought.

My second result examines the net effect of two potential countervailing mechanisms where CoP policy can affect patient mortality. I find little change in aggregate patient mortality. I provide suggestive evidence that the minimal mortality effect could result from a shift in transplant composition. As centers become more selective, kidney offers trickle down the waitlist and are more likely to reach patients who previously would not have received a transplant without CoP. In future work, I will build a structural model to estimate the effect of CoP on the distribution of health outcomes and study how two channels related to selection into transplant and improving quality of care affect patient mortaility.

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.

Agarwal, Nikhil, Itai Ashlagi, Michael A. Rees, Paulo Somaini, and Daniel Waldinger. 2021. "Equilibrium Allocations Under Alternative Waitlist Designs: Evidence From Deceased Donor Kidneys." *Econometrica*, 89(1): 37–76.

American Kidney Fund (AKF). 2008. "Deceased donor kidney transplants."

Bundorf, M. Kate, Natalie Chun, Gopi Shah Goda, and Daniel P. Kessler. 2009. "Do markets respond to quality information? The case of fertility clinics." *Journal of Health Economics*, 28(3): 718–727.

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.

Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. 2003. "Is More Information Better? The Effects of "Report Cards" on Health Care Providers." *Journal of Political Economy*, 111(3): 555–588.

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.

Kiberd, Bryce A., Karthik K. Tennankore, and Amanda J. Vinson. 2022. "Comparing the Net Benefits of Adult Deceased Donor Kidney Transplantation for a Patient on the Preemptive Waiting List vs a Patient Receiving Dialysis." *JAMA Network Open*, 5(7): e2223325–e2223325.

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.

Lunsford, Keri E., Gagan S. Prakash, and James V. Guarrera. 2022. "Does Liver Transplant Center Size Matter? Center Size Disparities and the Unintended Consequences of Transplant Regulation - A Commentary." *Transplantation*, 106(5): 911–912.

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."

Ramanarayanan, Subramaniam. 2011. "Reputations and Firm Performance: Evidence from the Dialysis Industry."

Roth, Alvin, Tayfun Sönmez, Utku Unver, 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.

Vatter, Benjamin. 2023. "Quality Disclosure and Regulation: Scoring Design in Medicare Advantage." *SSRN Electronic Journal*.

Volk, Michael L., Amber R. Wilk, Cameron Wolfe, and Daniel R. Kaul. 2017. "The "PHS Increased Risk" Label Is Associated With Nonutilization of Hundreds of Organs per Year." *Transplantation*, 101(7): 1666–1669.

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.

Zhang, Juanjuan. 2010. "The Sound of Silence: Observational Learning in the U.S. Kidney Market." *Marketing Science*, 29(2): 315–335.

Appendices

A Supplementary Figures

Line		Center 1 Year	National 1 Year
	Adult (Age 18+)		
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.

		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)			
Characteristic Covariates	Reference Group	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)		0.007999	1.01	0.0026	0.0017	0.009014	1.01	0.0026	0.0005	0.010833	1.01	0.0026	<0.0001
Cold ischemia time: missing		0.301176	1.35	0.1068	0.0048	0.370824	1.45	0.0976	0.0001	0.424076	1.53	0.0928	<0.0001
Deceased donor kidney was pumped: missing	no	0.372378	1.45	0.2695	0.1671	0.366577	1.44	0.2521	0.1459	0.470206	1.6	0.2907	0.1057
Deceased donor kidney was pumped: yes	no	-0.177109	0.84	0.0567	0.0025	-0.121466	0.89	0.0559	0.0299	-0.117816	0.89	0.0549	0.0320
Deceased donor with history of diabetes: missing	no	-0.560700	0.57	0.4098	0.1712	-0.758828	0.47	0.4485	0.0907	-0.385158	0.68	0.3553	0.2783
Deceased donor with history of diabetes: yes	no	0.208408	1.23	0.0776	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	0.0589	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		0.051958	1.05	0.0057	<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	0.0845	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	0.1390	0.2452	0.152541	1.16	0.1365	0.2638
Donor race: Black	White	0.137920	1.15	0.0670	0.0396	0.162123	1.18	0.0638	0.0111	0.123100	1.13	0.0635	0.0526
Donor race: Hispanic/Latino	White	-0.036753	0.96	0.0741	0.6199	-0.017480	0.98	0.0704	0.8040	-0.006509	0.99	0.0681	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	0.0466	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	-0.216594	0.81	0.1914	0.2577	-0.196728	0.82	0.1892	0.2983
Donor serum creatinine: missing		-0.001372	1	0.3596	0.9970	-0.085849	0.92	0.3583	0.8106	-0.007086	0.99	0.3418	0.9835
Donor to recipient weight ratio: continuous		-0.018267	0.98	0.0829	0.8256	-0.070278	0.93	0.0795	0.3769	-0.158743	0.85	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

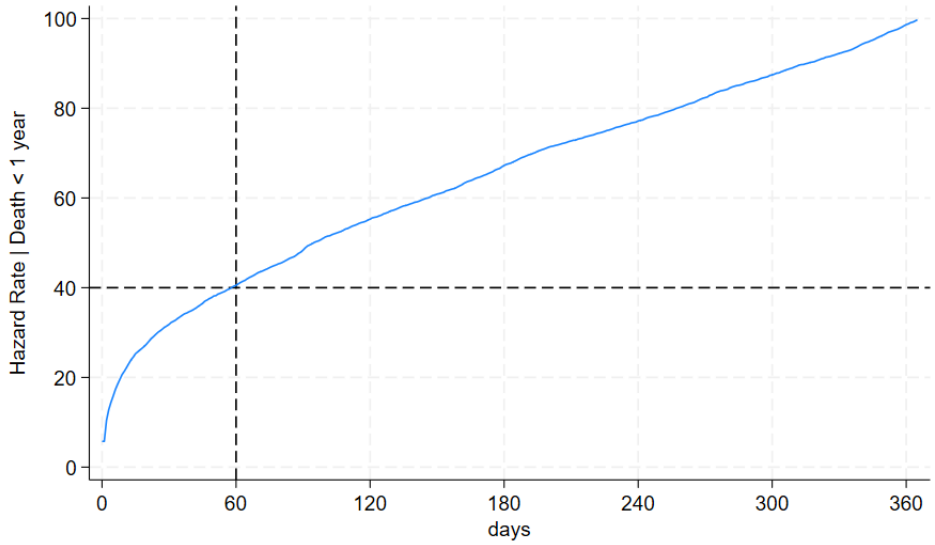
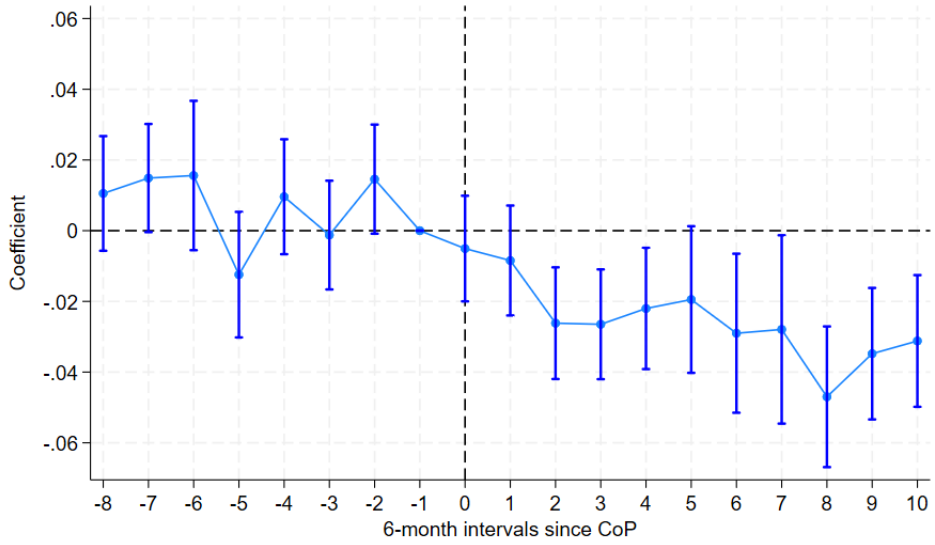


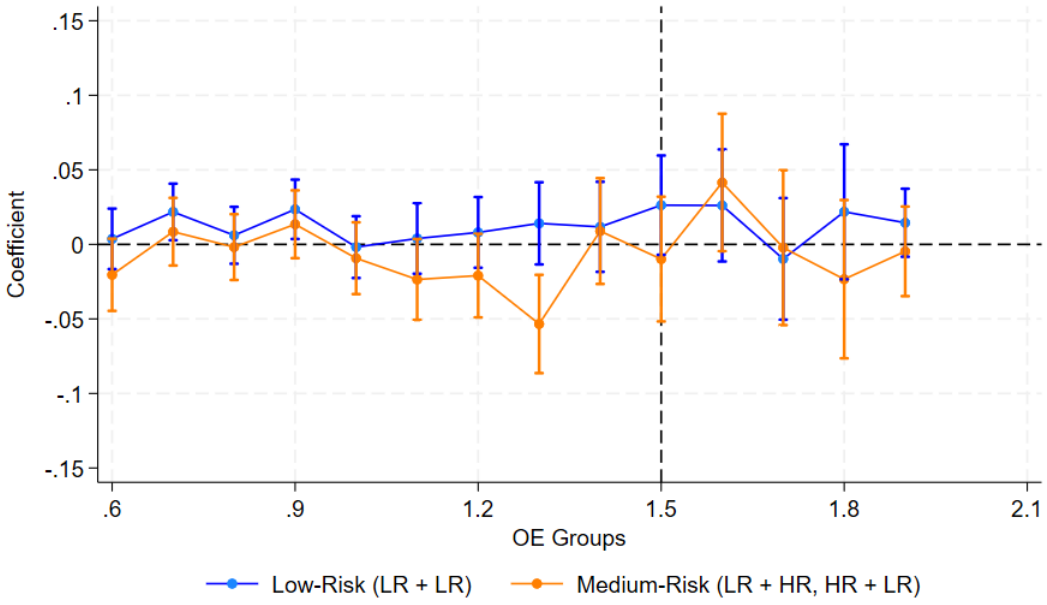
Figure A3: Hazard rate conditional on transplant death within 1-year. The intersection of the dash lines indicates 40% of transplants that die within 1-year die within 60 days/2 months after the transplant.

Figure A4: Impact of CoP on dynamic acceptance behavior for $T_{-0} = \mathbb{1}\{OE_{ct(k)} \geq 1.0\}$

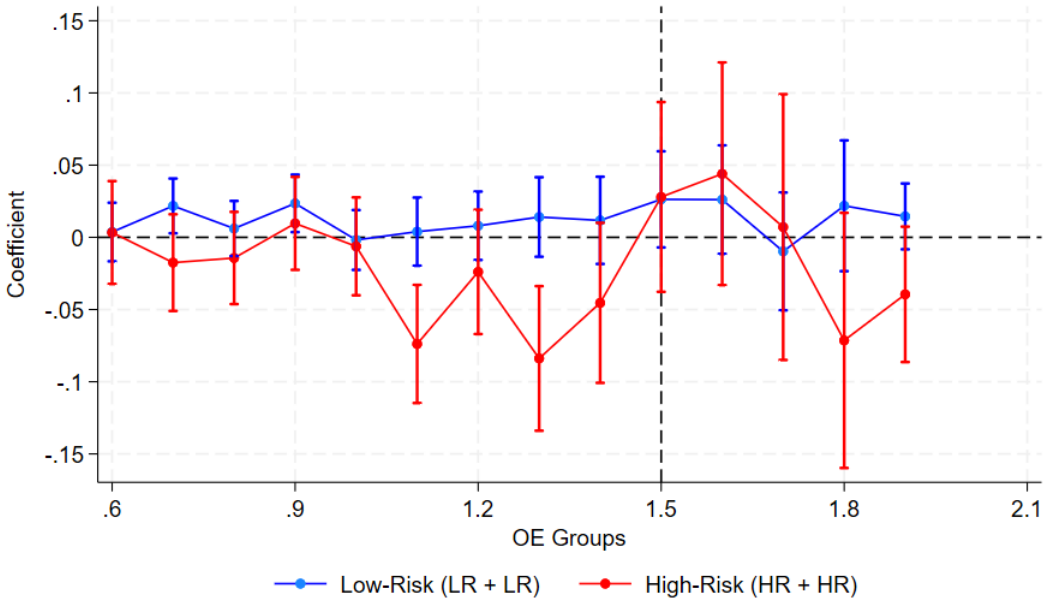


Note: The figure shows OLS estimates and 95% confidence intervals of the coefficients μ_s form equation 2. I plot all coefficients relative to when CMS introduced CoP (s=0). I cluster standard errors at the center level.

Figure A5: Post-transplant 1 year deaths across different $OE_{ct(k)}$ groups for different risk profiles



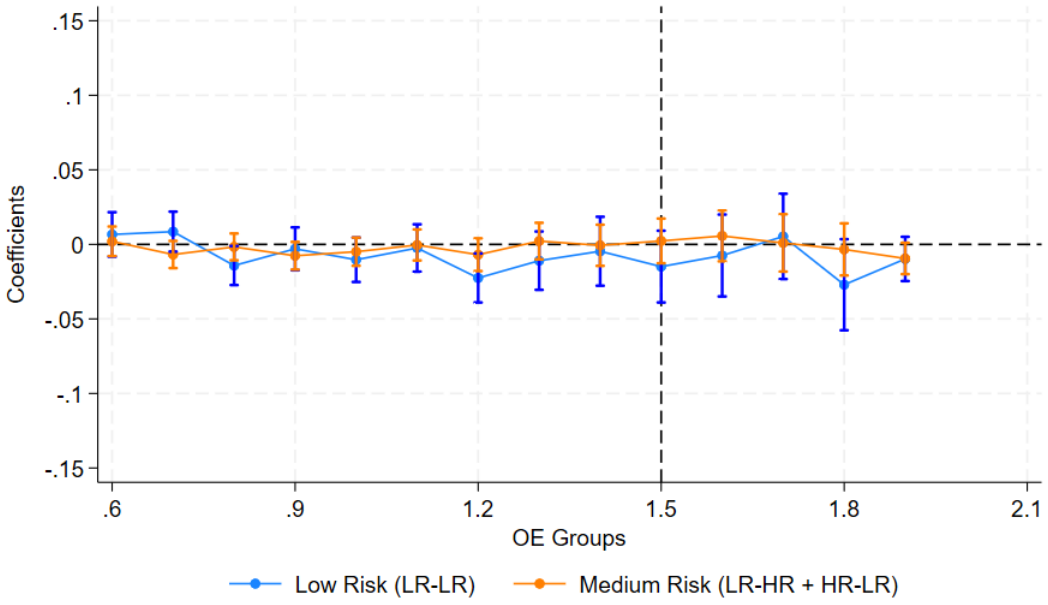
(a) Low-Risk v.s. Medium-Risk



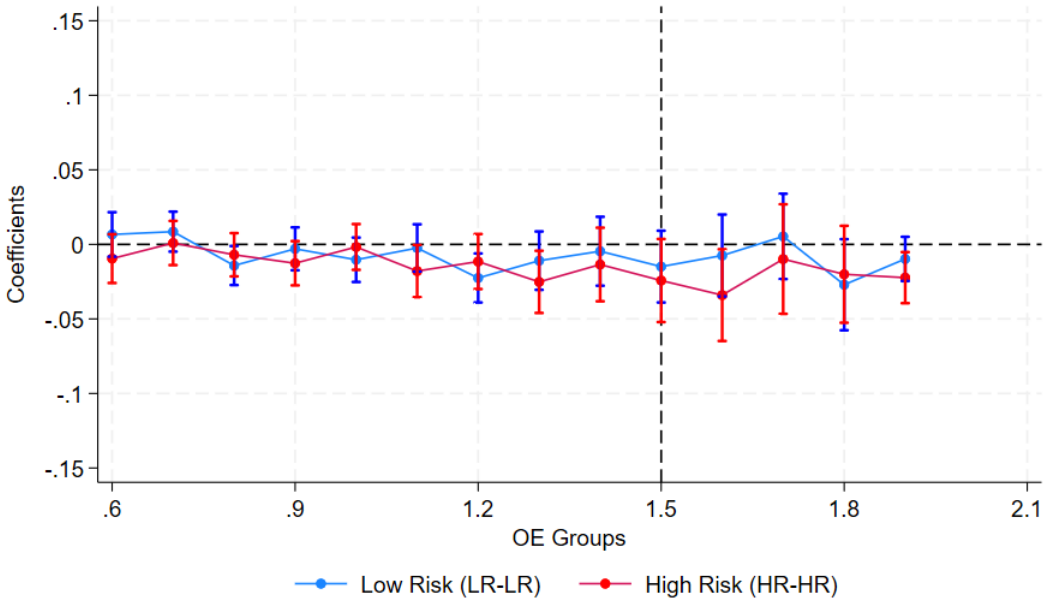
(b) Low-Risk v.s. High-Risk

Note: The figure shows the post-transplant 1-year death rate for each OE group m after CoP for different risk profiles. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. I cluster standard errors at the center level. The point estimates are presented in Columns 2 - 4 of Table B5 of the Appendix.

Figure A6: 1-year mortality after first kidney offer across different $OE_{ct(k)}$ groups for different risk profiles



(a) Low-Risk v.s. Medium-Risk



(b) Low-Risk v.s. High-Risk

Note: The figure shows the 1-year mortality after the first kidney offer for each OE group m after CoP for different risk profiles. I plot the OLS estimates and 95% confidence intervals of the coefficients β_m from equation 1. I cluster standard errors at the center level. The point estimates are presented in Columns 2 - 4 of Table B6 of the Appendix.

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: Acceptance behavior across different $OE_{ct(k)}$ groups

	(1)	(2)	(3)
$\beta_{0.6}$	0.00766	0.00361	0.000164
	(0.00662)	(0.00640)	(0.00655)
$\beta_{0.7}$	0.0172**	0.00665	-0.00680
	(0.00624)	(0.00605)	(0.00625)
$\beta_{0.8}$	0.00681	-0.0101	-0.0291***
	(0.00618)	(0.00598)	(0.00618)
$\beta_{0.9}$	-0.00774	-0.0164**	-0.0291***
	(0.00625)	(0.00604)	(0.00624)
$\beta_{1.0}$	-0.0192**	-0.0264***	-0.0371***
	(0.00632)	(0.00611)	(0.00635)
$\beta_{1.1}$	-0.0161*	-0.0325***	-0.0475***
	(0.00678)	(0.00655)	(0.00681)
$\beta_{1.2}$	-0.0126	-0.0205**	-0.0296***
	(0.00697)	(0.00674)	(0.00701)
$\beta_{1.3}$	-0.0329***	-0.0419***	-0.0496***
	(0.00780)	(0.00754)	(0.00784)
$\beta_{1.4}$	-0.0302***	-0.0278***	-0.0330***
	(0.00833)	(0.00805)	(0.00832)
$\beta_{1.5}$	-0.0580***	-0.0581***	-0.0608***
	(0.00955)	(0.00923)	(0.00951)
$\beta_{1.6}$	-0.0269*	-0.0111	-0.0191
	(0.0106)	(0.0102)	(0.0105)
$\beta_{1.7}$	-0.0443***	-0.0313**	-0.0424***
	(0.0115)	(0.0111)	(0.0113)
$\beta_{1.8}$	-0.0367**	-0.0412***	-0.0506***
	(0.0116)	(0.0112)	(0.0115)
$\beta_{1.9}$	-0.00539	-0.0149*	-0.0268***
	(0.00727)	(0.00706)	(0.00779)
Center FE			✓
6-months period FE	✓	✓	✓
Pat. and Kid. Controls		✓	✓
Observations	282393	282393	282392

Standard errors in parentheses

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

Table B3: Acceptance behavior across different $OE_{ct(k)}$ groups for different risk profiles

	(1) Low Risk	(2) Medium Risk	(3) High Risk
$\beta_{0.6}$	0.0304***	-0.0157***	-0.0260***
	(0.00519)	(0.00296)	(0.00467)
$\beta_{0.7}$	0.0138*	-0.0199***	-0.00966
	(0.00590)	(0.00298)	(0.00455)
$\beta_{0.8}$	-0.0105*	-0.0381***	-0.0418***
	(0.00458)	(0.00352)	(0.00431)
$\beta_{0.9}$	-0.0133*	-0.0404***	-0.0304***
	(0.00464)	(0.00365)	(0.00479)
$\beta_{1.0}$	-0.0107	-0.0481***	-0.0579***
	(0.00826)	(0.00481)	(0.00676)
$\beta_{1.1}$	-0.0314**	-0.0626***	-0.0528***
	(0.00924)	(0.00626)	(0.00449)
$\beta_{1.2}$	-0.00248	-0.0477***	-0.0364***
	(0.00942)	(0.00651)	(0.00647)
$\beta_{1.3}$	-0.0457**	-0.0585***	-0.0462***
	(0.0122)	(0.00685)	(0.00830)
$\beta_{1.4}$	-0.0108	-0.0364***	-0.0634***
	(0.00922)	(0.00670)	(0.00675)
$\beta_{1.5}$	-0.0465***	-0.0641***	-0.0857***
	(0.0108)	(0.00828)	(0.00866)
$\beta_{1.6}$	0.0160	-0.0320***	-0.0451***
	(0.0119)	(0.00740)	(0.0100)
$\beta_{1.7}$	-0.0209	-0.0659***	-0.0559***
	(0.0122)	(0.00828)	(0.00717)
$\beta_{1.8}$	-0.0357**	-0.0565***	-0.0501***
	(0.0116)	(0.00873)	(0.00814)
$\beta_{1.9}$	-0.000432	-0.0385***	-0.0532***
	(0.0122)	(0.00753)	(0.00689)
Center FE	✓	✓	✓
6-months period FE	✓	✓	✓
Pat. and Kid. Controls	✓	✓	✓
Observations	116046	139896	67636

Standard errors in parentheses

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

Table B4: Acceptance behavior across different $OE_{ct(k)}$ groups for different center volumes

	(1) Small Centers	(2) Large Centers
$\beta_{0.6}$	-0.0372***	0.0133
	(0.00688)	(0.00722)
$\beta_{0.7}$	-0.0143	0.0176**
	(0.00770)	(0.00667)
$\beta_{0.8}$	-0.0352***	-0.00378
	(0.00820)	(0.00644)
$\beta_{0.9}$	-0.0647***	-0.00809
	(0.00774)	(0.00639)
$\beta_{1.0}$	-0.0551***	-0.0170**
	(0.00846)	(0.00637)
$\beta_{1.1}$	-0.0636***	-0.0309***
	(0.00849)	(0.00710)
$\beta_{1.2}$	-0.0517***	-0.00970
	(0.00894)	(0.00739)
$\beta_{1.3}$	-0.0934***	-0.00362
	(0.00934)	(0.00909)
$\beta_{1.4}$	-0.0817***	-0.00580
	(0.00809)	(0.00923)
$\beta_{1.5}$	-0.0877***	-0.0586***
	(0.00965)	(0.0122)
$\beta_{1.6}$	-0.0404**	-0.0209
	(0.0112)	(0.0133)
$\beta_{1.7}$	-0.0778***	-0.0197
	(0.0103)	(0.0141)
$\beta_{1.8}$	-0.0996***	-0.0132
	(0.0106)	(0.0143)
$\beta_{1.9}$	-0.0534***	-0.00262
	(0.00981)	(0.00789)
Center FE	✓	✓
6-months period FE	✓	✓
Pat. and Kid. Controls	✓	✓
Observations	71633	190098

Standard errors in parentheses

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

Table B5: Post-transplant 1-year deaths across different $OE_{ct(k)}$ groups

	(1) All	(2) Low Risk	(3) Medium Risk	(4) High Risk
$\beta_{0.6}$	-0.00473	0.00369	-0.0204	0.00333
	(0.00870)	(0.0103)	(0.0124)	(0.0182)
$\beta_{0.7}$	0.00891	0.0218*	0.00850	-0.0175
	(0.00828)	(0.00968)	(0.0116)	(0.0171)
$\beta_{0.8}$	-0.000842	0.00610	-0.00183	-0.0143
	(0.00818)	(0.00973)	(0.0113)	(0.0163)
$\beta_{0.9}$	0.0171*	0.0235*	0.0135	0.00967
	(0.00838)	(0.0102)	(0.0116)	(0.0164)
$\beta_{1.0}$	-0.00472	-0.00182	-0.00929	-0.00624
	(0.00870)	(0.0106)	(0.0123)	(0.0173)
$\beta_{1.1}$	-0.0207*	0.00394	-0.0235	-0.0739***
	(0.00969)	(0.0120)	(0.0138)	(0.0209)
$\beta_{1.2}$	-0.00793	0.00802	-0.0209	-0.0239
	(0.00986)	(0.0121)	(0.0143)	(0.0220)
$\beta_{1.3}$	-0.0273*	0.0141	-0.0534**	-0.0839**
	(0.0112)	(0.0141)	(0.0168)	(0.0255)
$\beta_{1.4}$	0.00513	0.0118	0.00893	-0.0454
	(0.0121)	(0.0154)	(0.0181)	(0.0283)
$\beta_{1.5}$	0.0167	0.0263	-0.00986	0.0280
	(0.0136)	(0.0170)	(0.0213)	(0.0335)
$\beta_{1.6}$	0.0353*	0.0261	0.0415	0.0440
	(0.0151)	(0.0192)	(0.0235)	(0.0393)
$\beta_{1.7}$	-0.00302	-0.00974	-0.00222	0.00708
	(0.0167)	(0.0208)	(0.0265)	(0.0469)
$\beta_{1.8}$	-0.00664	0.0219	-0.0234	-0.0714
	(0.0176)	(0.0231)	(0.0271)	(0.0451)
$\beta_{1.9}$	0.00490	0.0145	-0.00469	-0.0395
	(0.0102)	(0.0116)	(0.0153)	(0.0239)
Center FE	✓	✓	✓	✓
6-months period FE	✓	✓	✓	✓
Pat. and Kid. Controls	✓	✓	✓	✓
Observations	102499	51432	48264	30348

Standard errors in parentheses

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

Table B6: 1-year mortality after first kidney offer across different $OE_{ct(k)}$ groups

	(1) All	(2) Low Risk	(3) Medium Risk	(4) High Risk
$\beta_{0.6}$	-0.00111	0.00666	0.00204	-0.00954
	(0.00426)	(0.00773)	(0.00479)	(0.00852)
$\beta_{0.7}$	-0.00389	0.00854	-0.00681	0.000921
	(0.00380)	(0.00682)	(0.00436)	(0.00690)
$\beta_{0.8}$	-0.00614	-0.0142*	-0.00170	-0.00696
	(0.00395)	(0.00702)	(0.00466)	(0.00775)
$\beta_{0.9}$	-0.00878	-0.00297	-0.00764	-0.0127
	(0.00447)	(0.00651)	(0.00495)	(0.00793)
$\beta_{1.0}$	-0.00588	-0.0103	-0.00491	-0.00173
	(0.00429)	(0.00838)	(0.00436)	(0.00865)
$\beta_{1.1}$	-0.00398	-0.00243	-0.000411	-0.0179
	(0.00459)	(0.00825)	(0.00569)	(0.00997)
$\beta_{1.2}$	-0.0109	-0.0225*	-0.00693	-0.0115
	(0.00564)	(0.0101)	(0.00586)	(0.0111)
$\beta_{1.3}$	-0.00531	-0.0109	0.00220	-0.0252
	(0.00567)	(0.01000)	(0.00572)	(0.0145)
$\beta_{1.4}$	-0.00144	-0.00460	-0.000629	-0.0135
	(0.00602)	(0.0133)	(0.00746)	(0.0132)
$\beta_{1.5}$	-0.00458	-0.0149	0.00231	-0.0243
	(0.00776)	(0.0137)	(0.00857)	(0.0158)
$\beta_{1.6}$	-0.00289	-0.00746	0.00566	-0.0340
	(0.00676)	(0.0144)	(0.00840)	(0.0181)
$\beta_{1.7}$	0.00151	0.00534	0.000979	-0.00985
	(0.00847)	(0.0166)	(0.00931)	(0.0232)
$\beta_{1.8}$	-0.00850	-0.0271	-0.00339	-0.0201
	(0.00825)	(0.0181)	(0.00929)	(0.0179)
$\beta_{1.9}$	-0.00993	-0.00977	-0.00946	-0.0223
	(0.00559)	(0.00974)	(0.00647)	(0.0121)
Center FE	✓	✓	✓	✓
6-months period FE	✓	✓	✓	✓
Pat. and Kid. Controls	✓	✓	✓	✓
Observations	307374	93359	197692	85965

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 $KDPI \leq 50$ [18].

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 $EPTS \leq 50$ [19].

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.

[18] Source:https://optn.transplant.hrsa.gov/data/allocation-calculators/kdpi-calculator/learn-about-kdpi/
[19] Source:https://optn.transplant.hrsa.gov/data/allocation-calculators/epts-calculator/learn-about-epts/