Causally Evaluating Selection: Biological, Behavioral and Institutional Effects on the Kidney Transplant Waitlist

Stephen Kastoryano*

September 15, 2024

Abstract

In causal studies, selection is usually considered a nuisance parameter to be held constant across treatment arms. This paper takes selection as the central object of interest in the context of the US kidney transplant waitlist. It advances an under-exploited dynamic causal setup, with corresponding dynamic treatment effect model and economic theory, to test different selection hypotheses. Using candidates' blood type as a conditionally exogenous shifter of transplant rates, the paper explains why groups of candidates with vastly different transplant rates have nearly identical overall survival rates. It also explains why groups of candidates with higher transplant rates display lower pre-transplant survival. We show that these patterns are due to institutional incentives which inadvertently lead transplant centers to excessively favour matching of high-quality kidneys to healthier candidates. This selection explains the large number of discards on the US kidney transplant waitlist. Somewhat unexpectedly, given the risk of transplant to less healthy candidates implied by the results, this skewed selection may be welfare improving if the aim is to only maximize survival, and not quality of life.

Keywords: Selection effects, Dynamic treatment effects, Kidney transplant, Survival models.

JEL Codes: I12, I18, C22, C41

*University of Reading, IZA, e-mail: s.p.kastoryano@reading.ac.uk.

Thanks to Jad Beyhum, Petyo Bonev, Jaap Abbring, Giovanni Mastrobuoni, Sang Yoon (Tim) Lee, and Christoph Breunig for valuable comments and help. The data reported here have been supplied by the Hennepin Healthcare Research Institute (HHRI) as the contractor for the Scientific Registry of Transplant Recipients (SRTR). The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy of or interpretation by the SRTR or the U.S. Government. The research protocol was approved by the IRB of the University of Reading.

1 Introduction

Medical treatments and social programs - education, welfare, labour market, housing - often operate through waitlists. The causal evaluations of these programs and treatments strive to ensure selection is held fixed among treated and control groups. However, after rolling out a treatment, or changing a policy, selection into treatment may not remain invariant. Faced with a new information setting imposed by regulators, awaiting individuals and institutional administrators may re-optimize their behaviour in response to changing incentives (Rosenzweig and Wolpin, 2000). These pre-treatment changes can alter the composition of individuals receiving treatment, if opting out or delaying treatment is possible, but also the ex-post effect of the treatment (Chassang et al., 2012). When this is the case, evaluating the effect of changes to the process of selection may be as important as evaluating the effect of the program or treatment itself.

The role of selection is apparent in the current context of kidney transplants in the US. As of 2022, around 140,000 candidates were on the waitlist for a kidney transplant, with around 40,000 added each year but only 25,000 removed after a successful transplantation (Lentine et al., 2023). Despite this under-supply, 25% of viable kidneys were discarded in 2021,[1] a share far exceeding that of other high-income countries.[2] It is as yet unclear which actors are driving this out-selection (candidates, transplant centers, or organ retrieval organisations), whether current market incentives are misaligned or misunderstood, and the welfare consequences of these losses.

This paper evaluates selection effects on the US kidney transplant waitlist using a dynamic causal framework. In particular, it exploits variation in blood types (AB,A,B,O) as a baseline shifter for the transplant rate to draw out selection patterns. The measured selection effects, combined with a discussion of incentives and tests based on economic theory, reveal the mechanisms leading to an excess discard of kidneys. Perhaps more importantly, these same mechanisms also explain how composition changes of the transplanted and heterogenous effects of the transplant lead groups of candidates with vastly different transplant rates (AB vs. O blood types) to have nearly identical overall survival rates. It also explains why groups of candidates with higher transplant rates display lower pre-transplant survival.

More specifically, the analysis shows that ethically motivated institutional considerations have produced misaligned incentives, in particular for transplant centers. Transplant centers, vying for higher ratings to attract more candidates and reduce costs, overtly favour matching high-quality kidneys to healthier candidates at the expense of other po-

[1] The US discard rates have been increasing since 2015.

[2] France had a kidney discard rate of 9.1% from 2004-2014, the United Kingdom has a rate ranging from 10% to 12% and Eurotransplant, a consortium of eight countries including Germany, reported a rate of about 8% (Aubert et al., 2019; Bernstein, 2023). The main reason for this difference is that the US system is unusually transparent compared to these other regions, leaving more room for discretion in which organs to accept in the hands of transplant centers and candidates on waitlists (Rasmussen et al., 2018).

tential life saving, but more risky, matches. Somewhat unexpectedly, the analysis also suggests this overt bias may be welfare improving if only survival, as opposed to quality of life, is being maximised. This is because the uncovered selection process suggests, contrary to common belief,[3] that less healthy individuals experience, on average, a negative effect of a transplant on survival relative to remaining on dialysis. The results and discussion of selection in relation to ethical considerations in this paper have wider implications for medical treatments and social programs. In particular, they underline welfare tradeoffs when the incentives determining the financial compensation of institutions providing the program (transplant centers) are responding to the ex-ante preferences of the beneficiaries of the program (transplant candidates) but these incentives are not aligned with the objectives of the regulating body or the ex-post preferences of the beneficiaries.

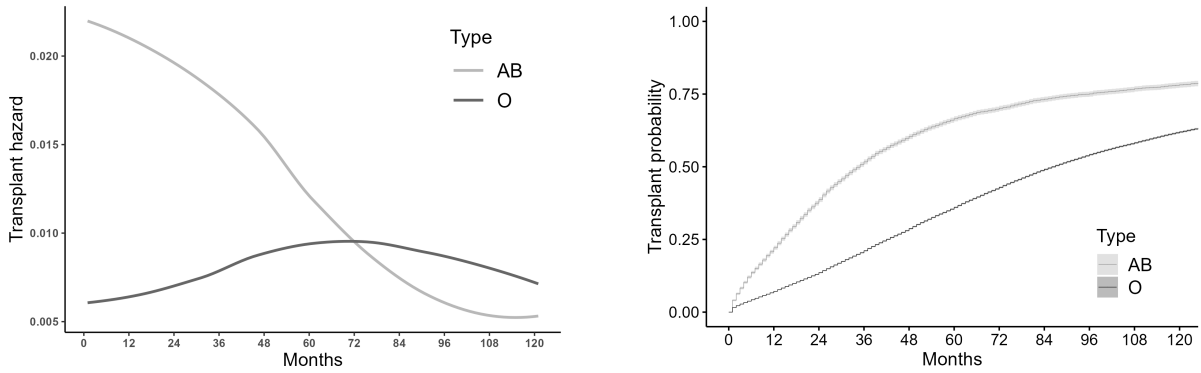
To assess hypotheses of selection, the paper promotes an under-exploited dynamic causal setup, set out in Abbring and Van den Berg (2005), with corresponding economic theory and dynamic treatment effect models. The setup of this framework is common to many waitlist allocation systems. Individuals enter a state at time 0 and are randomized to a *regime* at baseline. The regime is a special type of intention-to-treat variable which dictates a stochastic propensity for the timing of a future *treatment* among agents. Thereafter, at different moments in time and depending upon their regime, surviving agents are randomized to actually receive treatment. In the context of kidney transplants, individuals enter the waitlist when their kidney function falls below a certain threshold. For the regime variable, I propose a useful and previously unutilised biological randomization, which is that of individual ABO blood type. Blood type, as discussed and tested, will be assumed random with respect to the survival outcome conditional, among others, on race. However, due to compatibility constraints, and as required in the causal setup, blood type stochastically influences the time until an individual receives an offer of a kidney. The treatment in this paper's application is the duration until receiving a kidney transplant and the outcome is survival.

Figure 1 presents the central patterns of selection under study using detailed administrative data on all kidney transplants from the Organ Procurement and Transplantation Network (OPTN).[4] Figure 1a presents the hazard of receiving a kidney transplant from a deceased donor for AB-blood types, who are universal recipients, and O-blood type candidates, who can only receive a transplant from O-type donors.[5] We see that, due

[3]Due to selection issues and ethical considerations, unbiased causal estimates of the effects of transplants on survival have been elusive until recently (Agarwal et al., 2020). Surveys do, however, offer strong evidence that transplants improve quality of life relative to dialysis, the next best alternative (Axelrod et al., 2018).

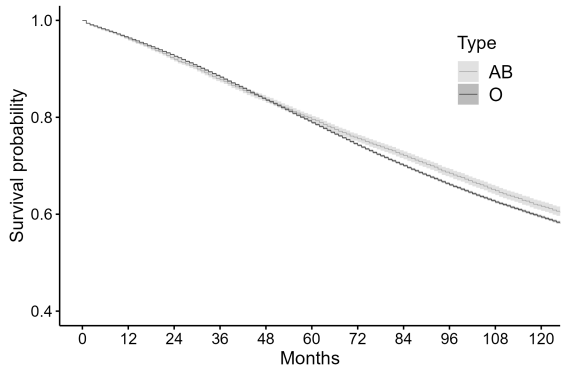
[4]These are obtained from the Organ Procurement and Transplantation Network/Scientific Registry of Transplant Recipients (OPTN/SRTR) and include socio-economic information and detailed medical histories for waitlist candidates and donated kidneys.

[5]Much of the early interest in allocation systems from economists focused on kidney transplant exchange programs among living donors (Roth et al., 2004, 2007; Agarwal et al., 2019). These exchange programs, along with the larger market for living donor kidneys, represent a minority of recorded transplantations. The majority, over 70%, of kidneys available for transplant are from deceased donors.

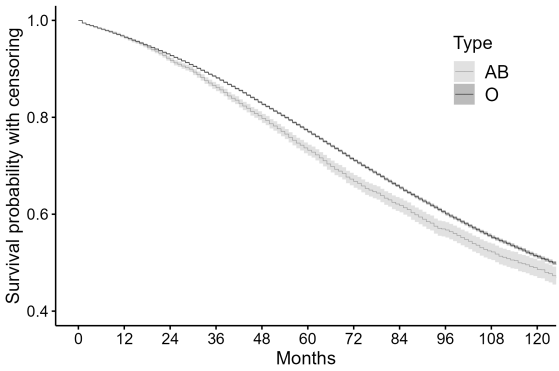


(a) Transplant Hazard

(b) Transplant Probability



(c) Survival



(d) Pre-Transplant Survival

Figure 1: Survival of candidates on the kidney transplant waitlist
Includes candidates entering waitlist after 01 June 2002. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.
Observations: $N_O = 186,838$, $N_{AB} = 14,683$.

to compatibility restrictions, AB-blood candidates have a higher hazard of receiving a transplant at any time in the first four years on the waitlist than O blood types. This results in a probability of transplant, presented in Figure 1b, which is approximately twice as large for the AB-blood group over the first 5 years on the wailtist.

Despite the large differences in transplant probabilities, and relatively steep declines in survival over time, Figure 1c reveals a surprising and previously unstudied pattern when it comes to survival. It shows that overall survival for AB and O blood types follow nearly identical trends up to 5 years since entering the waitlist, with only small divergence thereafter. It is unclear whether these curves show that transplants have only marginal effects on survival, or whether they reflect a particular selection of who receives a transplant and what quality kidney they receive.

The problem of selection becomes apparent when considering Figure 1d which plots the probability of survival for AB-types and O-types after censoring the duration to death at the moment a candidate receives a kidney transplant. We see that, prior to receiving a kidney transplant, AB-type candidates with a higher propensity to receive a transplant also have a significantly lower survival rate than O-type candidates.

To formally evaluate these patterns, the paper draws on a dynamic treatment effect

model, which combines the identifying assumptions of Robins (1986) with the decomposition of causal effects proposed in Abbring and Van den Berg (2005). Estimates from a semi-parametric proportional hazard estimator show, consistent with Figure 1d, a 25 percent higher pre-transplant probability of death for AB blood type candidates than O blood types even after conditioning on all variables determining kidney offers as well as additional important health variables. These pre-transplant effects are relatively constant, in percentage terms, over the first five years on the waitlist. They also appear across various model specifications and in various subsamples, including males and females, white and non-white people, a sample which includes candidates who receive kidney transplants from living donors, and candidates entering the waitlist before and after a reform in the allocation mechanism in December 2014.

Having established that these selection effects cannot readily be explained by observed variables, our analysis delves into institutional incentives, additional non-dynamic estimations, and dynamic discrete choice theoretical search models to explain the process of selection on unobservables. We first show that the pre-transplant patterns in Figure 1d can arise when candidates on the waitlist are not sophisticated, meaning they do not adjust their behavior in response to changes in the kidney offer rate. More specifically, these patterns can arise if kidney offer rates are higher in the AB-blood type group and if pre-transplant survival and kidney acceptance are both increasing in unobserved health.

In a second step, the paper revisits the question of why we only observe marginal differences in overall survival in Figure 1c. To address this question, the analysis presents two findings. First, it shows that AB types have a similar overall survival as O blood types after conditioning on relevant covariates. Second, it shows that AB blood types receiving a transplant in the first months on the waitlist have a higher survival than O types. For these two effects and previous results to coincide, I conjecture that less healthy AB type candidates receiving later transplants experience a negative effect of a transplant on survival. I further test this conjecture in a heterogenous effects model using type-II diabetes as a proxy for low health and, indeed, find patterns indicating opposite transplant effects on survival for high and low health candidates.

In a final part, the analysis tests whether the patterns in Figure 1c-1d could also be explained by a model with sophisticated forward looking candidates. Drawing insights from variations on a standard dynamic discrete choice theoretical search model (Mortensen, 1986),[6] the analysis shows, in theory and simulations, that the four panels in Figure 1 are not consistent with rational choices made by candidates. They are, in line with other results, consistent with a setting in which the decision-makers are the transplant center's physicians, a transplant has a negative effect for a majority of candidates (or a strong negative effect for a smaller share of candidates), and the transplant-center's physicians foresee these negative effects.

[6]These models can be seen as a limited version of the more complete agent choice model in Agarwal et al. (2021), but with context specific attention to agent responses to differences in kidney offer rates.

The discussion and results in this paper contribute to several fields. The methodological setup builds on a rich history of dynamic treatment effects work in economics (Ham and LaLonde, 1996; Abbring and Van den Berg, 2003; Heckman and Navarro, 2007; Heckman et al., 2016; Han, 2021). The paper further adds to a growing field addressing expectation effects in economics (see Haaland et al. 2023 for a survey). The results also contribute to the literature on transplants in general and kidney transplants in particular. Notable recent work focused on optimal allocation mechanisms for the deceased donor waitlist include theoretical work directly considering ABO blood compatibility issues (Sönmez et al., 2020; Kim and Li, 2022). This paper also draws on various important theoretical and empirical insights focused on candidate option value of a transplant and discretion to reject offers (Zhang, 2010; Agarwal et al., 2018, 2021) and the first legitimate causal evaluation of transplants on survival in Agarwal et al. (2020). In contrast to that body of work, the current paper focuses more deliberately on the implications of financial and reputational incentives from the side of transplant centers. Finally, the results also relate to research underlining the relevance of market dynamics in the supply of kidneys (Dickert-Conlin et al., 2019; Teltser, 2019).

The remainder of the paper proceeds as follows. It begins in the next section by briefly introducing the kidney transplant allocation system. Section 3 introduces the potential outcomes framework, describes causal effects of interest, discusses the identifying assumptions in relation to the paper's kidney transplant setting and describes the estimation approach. Section 4 describes the data and Section 5 assesses sequentially different hypotheses for selection on the kidney transplant waitlist.

2 Kidney Transplant Allocation System

In the United States, the allocation of deceased kidneys to candidates on the waitlist is designed, coordinated, and administered by the Organ Procurement and Transplantation Network (OPTN). A person is allowed to register on the kidney transplant waitlist and start accruing time when their kidney function falls below 20%.[7] Upon entering the waitlist, OPTN collects in-depth information about the candidate's health conditions, immunological profile, and any other characteristic needed to ascertain kidney compatibility.

When a potential kidney becomes available, OPTN gather extensive biological information about the kidney and the donor's medical history. Through its automated UNet system, OPTN then calculate a priority order for any candidate who is compatible with the potential kidney. The kidney offer process balances equity and efficiency, prioritising candidates on the waitlist following strict criteria. Unlike some other organs, such as livers and hearts, priority is not given to candidates who need a transplant most urgently.

Prior to December 2014, the main prioritisation criteria were the compatibility of

[7] More precisely when the Glomerular Filtration Rate (GFR) is 20 or below.

blood and tissue types, the age of the candidate, the time a candidate had been on the waitlist, and their geographic proximity to the center holding the deceased donated kidney (Smith et al., 2012).

This allocation system was deemed to result in excessive amounts of unutilised kidneys, lost graft years,[8] high re-transplant rates, and inequity in distribution for minorities. In an effort to address these issues, and update a system which had not undergone significant change in 20 years, a new allocation system was introduced in December 2014. Some of the changes included adding points for patients with high levels of antibodies, and giving priority to healthier candidates for the highest potential longevity kidneys. This prioritization is computed using an Estimated Post-Transplant Survival (EPTS) score which is a function of age, diabetes status, whether the candidate is on dialysis, the number of years they have been on dialysis, and whether they previously received a transplant.[9] Under the new allocation rules, less emphasis is placed on waiting time and geographical location.[10] In terms of ABO blood type prioritisation, B and O blood type candidates are prioritised for identical blood type kidneys unless perfect tissue compatibility (zero-antigen) exists with AB (or A for O donors) blood types (Israni et al., 2014). This prioritisation, which existed both before and after the reform, attempts to adjust for the lower supply of B and O blood type kidneys. Key to future assumptions, the UNet mechanism determining a kidney offer is a function of observed variables of which all health related ones are available in the data.

There also exists some discretion on the part of candidates and their physicians concerning which kidney they are willing to accept. Many kidneys will be rejected automatically based on age, health, and kidney function criteria set by candidates and their physicians upon entering the waitlist. These criteria choices are available in the data. Other kidneys will still be rejected after passing the initial screenings (Gordon et al., 2020). Looking at data on offers between January 1, 2010 and December 31, 2013, Agarwal et al. (2018) show that, among kidneys passing screening criteria, only 0.9% of kidney offers are accepted and only 70% of those are actually transplanted.[11] This low acceptance rate is largely due to highly undesirable kidneys being offered to several thousand patients.[12] Decisions to reject offers have no bearing on future offers. As a result of

[8]Candidates with shorter life expectancy may receive kidneys with the potential to function longer.

[9]More specifically, there was an emphasis on matching kidneys from the top 20% healthiest donors (KDPI$<$ 20%) to the candidates in the top 20% of EPTS.

[10]In 2021, the geographic criteria was further adjusted to attribute points, on a linearly decreasing scale with notches, based on the distance between a candidate's listed transplant center and the donor hospital (Potluri and Bloom, 2022).

[11]This last discrepancy is due to final crossmatch blood screenings which predict high chances of kidney graft failure. Throughout the text, I abstract from this discrepancy and often refer to transplanted kidneys as accepted kidneys, since the last crossmatch stage occurs after the main behavioral decisions focused on in the paper.

[12]Several factors determine a kidney's quality in general. These include the donor's general health prior to death, their gender and age, as well as their kidney function and anatomy. Some donors die with infectious diseases which the candidate may contract if they proceed with a transplant. See Danovitch (2012) for a comprehensive review of kidney biology.

rejected offers, close to 25% of kidneys are never transplanted.[13]

3 Evaluation Framework

This section first introduces the potential outcomes notation and presents the evaluation framework, underlining which assumptions identify causal effects and what these assumptions imply in this paper's kidney transplant waitlist setting.

3.1 Potential Outcomes and Treatment Effects

The decomposition framework presented in this section combines the G-computation identifying assumptions of Robins (1986) with the decomposition of causal effects proposed in Abbring and Van den Berg (2005) and also draws on work from Lechner (2009) and Heckman et al. (2016). The potential outcomes notation combines the dynamic presentation of Abbring and Heckman (2007) with the more familiar notation in static treatment and mediation analyses (Angrist et al., 1996; Imai et al., 2010).

3.1.1 Regime and Potential Treatment Duration

We follow each agent from the moment of entering an initial state which is set as $t = 0$. In the empirical setting this will be the moment an individual enters the kidney transplant waitlist, but in other settings it could be the moment of contracting an illness, becoming unemployed, or advertising a new product on the market. At $t = 0$, agents are randomized to one of two regimes denoted by the random variable Z (the subscript i is suppressed for convenience). Z is observed by the econometrician and may be partially or fully known by the agents. The discussion focuses on the most simple setting in which agents can either be assigned to a baseline regime $Z = 0$ or to an alternative regime $Z = 1$.

The regime is a special type of two-in-one randomization which has received little attention in the literature. It is a variable which introduces administrative or other constraints which change the hazard rate to treatment, but does not determine a specific time for the future treatment. Although the framework allows a general interpretation of the regime randomization, it can find particularly interesting applications when being considered as an information randomization on expectations or beliefs. In the kidney transplant setting, $Z = 0$ for AB blood type individuals who face a high probability of rapidly finding a suitable kidney. $Z = 1$ for O blood type individuals who have a lower probability of rapidly finding a suitable kidney.

To define durations to treatment, we use a potential outcomes notation. Time is allowed to be discrete or continuous when defining potential outcomes and causal effects

[13] There is also a substantial amount of variation in kidney-acceptance rates among the 250 transplant centers (Wey et al., 2017).

of interest. Let the random variable S^z be the potential duration to treatment, or equivalently the treatment time, had an agent been subject to regime $Z = z$. The treatment time can take on any value $s > 0$. The paper considers the simplified setting in which there is only a single binary treatment which, once allocated, remains permanently thereafter. In the kidney transplant setting, s is the duration of time on the waitlist until an individual receives a kidney transplant.

3.1.2 Potential Exit Duration

We further define the potential outcome $T^{z,s}$ as the duration to exit had the agent been subject to regime $Z = z$, and been treated at $S^z = s$. In the previous example, $T^{1,6 \text{ months}}$ would be the potential time to death for a person on the kidney transplant waiting list with O blood type ($Z = 1$), and who (would have) received a kidney transplant after 6 months on the waitlist ($S^1 = 6$ months). Several comparisons can be made based on this notation. In this paper, the object of interest is the probability of survival past a period τ given treatment at s. Any comparison therefore requires formulating counterfactual potential outcomes of the form $\Pr(T^{z,s} > \tau)$.[14]

The usual problem with formulating causal comparisons with this potential outcome is that counterfactuals are never observed. On top of the usual static selection problems, there is dynamic selection in the sense that unobserved variables will influence the timing of exit, and some agents will exit before ever receiving treatment. Within this context, we use the shorthand notation $T^{z,\infty}$ to define the potential non-treated outcome. It represents the duration to exit had the agent been exposed to regime $Z = z$ and had not received treatment in finite time ($s \to \infty$), or, more concretely, by the time τ at which the outcome is evaluated.

3.1.3 Regime and Treatment Effect Decomposition

Using this notation, the relevant decomposition for a given treatment time s can already be defined:

$$
\begin{aligned}
\Pr(T^{z,s} > \tau) &= \beta_0 + \beta_z z + \beta_s \mathbb{1}(S = s) + \beta_{zs} z \mathbb{1}(S = s) \\
&\text{with } \beta_0 = \Pr(T^{0,\infty} > \tau) \\
&\quad \beta_z = [\Pr(T^{1,\infty} > \tau) - \Pr(T^{0,\infty} > \tau)] \\
&\quad \beta_s = [\Pr(T^{0,s} > \tau) - \Pr(T^{0,\infty} > \tau)] \\
&\quad \beta_{zs} = [(\Pr(T^{1,s} > \tau) - \Pr(T^{1,\infty} > \tau)) - (\Pr(T^{0,s} > \tau) - \Pr(T^{0,\infty} > \tau))]
\end{aligned} \tag{1}
$$

[14] In this paper the outcome of interest is the probability of survival instead of the expected exit duration $\mathbb{E}[T^{z,s}]$. Focusing on the expected duration to exit may be hampered by the fact that in many duration settings, as is the case in this study's application, a large fraction of exit outcomes are censored so the right tail of the exit distribution will be poorly approximated. Also, the expected potential exit duration outcomes can all be expressed as functions of the probability to exit, $\Pr(T^{z,s} \leq \tau) = 1 - \Pr(T^{z,s} > \tau)$. Alternatively, some studies focus on the relative effects of the potential hazard as causal effects. The drawback of focusing directly on the hazard to infer about causal effects is that its magnitude is difficult to interpret for cost-benefit analyses without transforming it into a survivor function.

This decomposition is a reformulation of one first proposed in Abbring and Van den Berg (2005). It parallels the well-known decomposition when evaluating heterogenous treatment effects.

As in a heterogeneity decomposition, β_0 represents the average probability of survival past τ had all candidates been of AB blood type and had none of them received a kidney transplant. β_z represents the pre-treatment difference in survival and is the central parameter of interest in this study. It represents the difference in the probability of survival past τ were a candidate to be of O blood type instead of AB blood type, withholding any effects of actually receiving a kidney transplant. β_s represents the difference in the probability of survival past τ were candidates to be of blood type AB and receive a kidney transplant at s compared to their probability of survival were they not to receive a kidney transplant by τ. β_{zs} represents the additional difference in the probability of survival past τ were candidates to be of O blood type and receive a kidney transplant at s relative to β_s.[15]

3.2 Assumptions and Identification of Causal Effects

In this section, I discuss the nonparametric identification of dynamic treatment effects originally formulated in Robins (1986). The evaluation framework will be presented in discrete time with exits and treatment times $t \in \mathbb{N}$ and $s \in \mathbb{N}$ to intuitively parallel assumptions from the static causal inference setting. To reduce notational burden, the setup assumes the observed period interval is the smallest discrete unit of time.

For each agent, when $T \geq S$, we observe the joint distribution (Z, S, T, X) where Z is the regime assignment, S the time to treatment, T the time to exit, and X is a set of baseline covariates at $t = 0$. Since identification is non-parametric, the model leaves the conditioning on observed baseline $(t = 0)$ covariates X implicit for notational convenience. However, the discussion does refer to covariates when considering the problems of intermediate or time-varying covariates.

The first identification requirement is that the regime is randomized and that treatments are randomized on survivors as first proposed by Robins (1986) and introduced into the economics literature by Lechner (2009),

[15] In some cases, the above formulation may not be the most policy relevant decomposition. A more relevant decomposition may require averaging effects over a treatment time interval, such as treatment before period s. When producing the decomposition relative to treatment in regime $Z = 1$, the average effects over the treatment interval $(0, s]$ are given by $\sum_{t=0}^{s} \cdot \Pr(S^1 = s) \Pr(T^{z,s} > \tau)$. The average policy parameters of interest would then be β_0, β_z, $\beta_{(0,s]} = \sum_{t=0}^{s} \Pr(S^1 = s) \cdot \beta_s$ and $\beta_{z(0,s]} = \sum_{t=0}^{s} \Pr(S^1 = s) \cdot \beta_{zs}$.

Assumption A.I: Dynamic Unconfoundedness: *For* $z = 0, 1$, $s \in \mathbb{N}$, $t \in \mathbb{N}$, $s \geq t$, *denote by* $\{T^{z,s}\}$, $\{S^z\}$ *the sets of all potential variables, then,*

$$\begin{aligned} (i) \quad & (\{T^{z,s}\}, \{S^z\}) \perp\!\!\!\perp Z \\ (ii) \quad & \{T^{z,s}\} \perp\!\!\!\perp \mathbb{1}(S = s) \mid S \geq t, T \geq t, Z = z \end{aligned}$$

In the context of this study, the first assumption states that the blood type is randomized at $t = 0$ with respect to the potential time to death and potential time until receiving a kidney transplant. The second assumption says that receiving a kidney transplant is randomized for the survivors on the waitlist who did not yet receive a kidney transplant. Assumption $A.Ii$ implies and is implied by the 'no-anticipation' assumption of Abbring and Van den Berg (2003), $\Pr(T^{z,s} > \tau) = \Pr(T^{z,s'} > \tau)$ for all $s, s' > \tau$.[16]

Concerning the plausibility of Assumption $A.I$, ABO blood types have been identified as risk factors in specific disease processes such as vascular disease and malignancy (Wu et al., 2008; Sun et al., 2015; Abegaz, 2021) as well as for a small number of infectious diseases (Cooling, 2015). However, on the whole, there is no convincing evidence that genetic features associated to ABO blood types consistently lead to higher overall mortality in combination with diseases. That said, distributions of ABO blood types are not similar across races and may be correlated to other variables. As such, the empirical examination of assumption $A.Ii$ proceeds in two parts. The analysis first produces additional specifications with added baseline covariates, including race and previous malignancies, among others. These results ascertain the robustness of the estimated pre-transplant effect β_z to some leading choices of confounders and known drivers of mortality. Thereafter, the evaluation exploits the four variations of blood type and present subsample results which, taken together, offer strong evidence that differences in pre-transplant survival operate through pre-existing differences in candidate blood type, rather than different processes of selection over the waitlist.

Assumption $A.Iii$ should be questioned in observational dynamic settings. OPTN, through the UNet system, allocate kidney *offers* to candidates on the waitlist purely based on candidate and donor characteristics, and registered restrictions. As discussed in section 2, although we observe all health related variables held by OPTN, candidates in consultation with clinicians, or clinicians and transplant centers independently, are likely selecting which kidney offers to *accept*. These decisions may be based on variables outside the set determining the allocation of offers, and possibly outside the additional ones available in the data. This selection problem is the main reason for not focusing on post-transplant effects β_s and β_{zs} in the paper.[17] However, an important contribution

[16] Assumption A.I implicitly contains structural assumptions concerning the timing of causes and effects within a period. In this model's formulation of the conditioning set, $S \geq t$, we allow treatment at period t to influence exit in the same period. So, receiving a kidney transplant can influence survival immediately. If there is censoring in the sample, one can add a similar dynamic unconfoundedness condition under which right censored observations are dynamically missing (completely) at random.

[17] See Agarwal et al. (2020) for a dedicated discussion of ex-post effects accounting for selection on unobserved variables in the context of kidney transplants

of the paper is explaining how a violation of assumption Iii, while biasing the ex-post transplant effects, does not prevent us from developing behavioral and policy relevant insights from the pre-transplant effect β_z. Under certain conditions, we can leverage the structure of assumption $A.Iii$ to separate out pre-transplant selection due to prominent assignment and health variables held and currently used by OPTN, as opposed to other unobserved variables.

One condition is that there is no predetermined time from the point of view of candidates or surgeons at which candidates will be offered a suitable kidney. This likely holds as Organ Procurement Organisations (OPOs) gather consent from people or relatives of people who experienced severe brain bleeding, have been declared brain dead, or for whom cardiac death is imminent. Retrieved kidneys are then transplanted within 24-48 hours. These time constraints imply that it is difficult to predict well in advance when a new and compatible donated kidney will be brought forward.

In addition, when it comes to kidney offers, we need to justify a source of exogenous randomization in the allocation system. Two such sources are assumed: regional differences in kidney availability and differences in donor-recipient tissue-type compatibility criteria (HLA-type). In terms of tissue compatibility, for two candidates similar on all measures, the one with a higher HLA tissue type compatibility is more likely to receive a kidney transplant.[18] Geographic location also introduces several layers of random variation in kidney offers. Prior to December 2014, large differences in allocation stemmed from regional differences in kidney availability (Ata et al., 2020; Massie et al., 2011).[19] As a result, among surviving agents who are similar on all measures, the one living within the same donation service area as that which procured the kidney would be prioritized. Because there were additional sources of geographical variation prior to December 2014, the initial analysis focuses on pre-2015 data.[20] As will be shown, the results and insights from the pre-2015 analysis carry over to the new kidney allocation mechanism, which continues to display large geographic variation in deceased donor kidney availability (King et al., 2020).

In addition, the model impose that regimes, treatments, and the decomposition can be evaluated,

[18]HLA tissue types are not known to be correlated to blood types (Erikouglu et al., 2011).

[19]Due to the limited time between a kidney being made available and the time of transplant, few candidates are listed in multiple Donation Service Areas (Dickert-Conlin et al., 2019).

[20]Prior to the December 2014 reform, one relevant condition under which some regional prioritization criteria could be ignored is if there was a perfect compatibility on blood and tissue type (zero-antigen mismatch) for a candidate in another donation area. When this occurred, and a kidney was transferred to another donation service area, the service area obtaining the kidney incurred a payback debt which required relinquishing the next available donated kidney with a similar ABO blood type. This debt payback system adds additional exogenous variation to the geographic and tissue compatibility dimension.

Assumption A.II: Overlapping Support: For $z = 0, 1$, $t \in (0, s]$, $s \in \mathbb{N}$, $t \in \mathbb{N}$,

$$\begin{aligned} (i) \quad & 0 < \Pr(Z = z) < 1 \\ (ii) \quad & 0 < \Pr(S = t | S \geq t, T \geq t, Z = z) < 1 \end{aligned}$$

This overlapping support assumption guarantees first, $A.II(i)$, that we observe agents under both regimes. $A.II(ii)$ provides that there is no time before period s at which the treatment is allocated to all, or none, of the untreated survivors.

We add to these two assumptions Rubin's Stable Unit Treatment Value Assumption (Rubin, 1980) which we rejoin with Robins' consistency assumption (Robins, 1997; Murphy, 2003) in the following,

Assumption A.III: SUTVA: *For* $z = 0, 1$, $s \in \mathbb{N}$,

$$\begin{aligned} S^z &= S \qquad if\ Z = z \\ T^{z,s} &= T \qquad if\ S = s, Z = z \end{aligned}$$

In the empirical setting, this assumption states that the potential duration until an individual receives a kidney transplant and the potential time to death equal their corresponding observed duration to transplant and duration to death in the sample. The usual interpretation of the above is that potential outcomes for an agent do not depend on the observed or counterfactual outcomes of any other agent. However, since the data includes the universe of candidates on the waitlist in the US, and this study's main goal is to diagnose which forms of pre-transplant selection may be at play for that population, the questions of spillovers and extrapolation when rolling-out a treatment are of second order importance since the data were collected from the hypothetical causal setting of interest.

3.2.1 Identification of Regime and Treatment Effects

Under $A.I - A.III$, we can consider the identification result for the separate components of the causal decomposition in equation 1.

$$\begin{aligned} \beta_0 &= \prod_{t=0}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 0) \\ \beta_z &= \prod_{t=0}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 1) - \prod_{t=0}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 0) \\ \beta_s &= \prod_{t=0}^{\tau} \Pr(T > t | S = s, T \geq t, Z = 0) - \prod_{t=0}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 0) \\ \beta_{zs} &= \prod_{t=0}^{\tau} \Pr(T > t | S = s, T \geq t, Z = 1) - \prod_{t=0}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 1) - \beta_s \end{aligned} \tag{2}$$

where we use a shorthand notation,

$$\prod_{t=0}^{\tau} \Pr(T > t|S = s, T \geq t, Z = z) \equiv$$
$$\prod_{t=s}^{\tau} \Pr(T > t|S = s, T \geq t, Z = z) \cdot \prod_{t=0}^{s-1} \Pr(T > t|S > t, T \geq t, Z = z)$$

We can also identify the probability of treatment at period s under regime z,

$$\alpha_z = \Pr(S^z = s) = \Pr(S = s|S \geq s, T \geq s, Z = z) \cdot \prod_{t=0}^{s-1} \Pr(S > t|S \geq t, T \geq t, Z = z)$$

The above equations are adapted versions of Robins (1986) and Gill and Robins (2001) "g-computation formula" which also admits non-binary and non-permanent treatments as well as non-duration outcomes. As in the general formula, this version allows for causal dependence of treatment histories on outcomes but also of outcome histories on the treatment. The main parameter of interest throughout the results is β_z representing pre-transplant effects.

3.3 Estimation Approach

The estimator builds on a semi-parametric proportional hazard model.[21] The joint exit and treatment hazards in the proportional hazard specification are written as,

$$\theta_t^T(x, z, s) = \lambda_t^T(z, \mathbb{1}(s \leq t)) \exp(x'\beta^T)$$
$$\theta_t^S(x, z) = \lambda_t^S(z) \exp(x'\beta^S)$$

For the estimation, I first parameterize the duration dependence functions $\lambda_t^T(z, \mathbb{1}(s \leq t))$ and $\lambda_t^S(z)$ as separate piecewise constant baseline hazards depending on Z and whether treatment occurred $\mathbb{1}(s \leq t)$. I then estimate the parameters of the joint density of the above by Maximum Likelihood.[22]

Each treated causal effect taking the form $\prod_{t=\tau_1}^{\tau_2} \Pr(T > t|S = s, T \geq t, Z = z)$ is then computed as,

$$\sum_{\{i\}} \hat{w}_i(s) \exp(-\int_{\tau_1}^{\tau_2} \hat{\theta}_t^T(x_i, z, s)dt) = \sum_{\{i\}} \hat{w}_i(s) \exp(-\int_{\tau_1}^{\tau_2} \hat{\lambda}_t^T(z_i, \mathbb{1}(s_i \leq t)) \exp(x_i'\hat{\beta}^T)dt)$$

[21] I opt for a continuous time estimator rather than a discrete time one mainly for computational reasons. Also, for settings in which identifying assumptions are likely to hold and time intervals are small, continuous and discrete time estimators produce similar results in treatment effect duration models (Kastoryano and van der Klaauw, 2022). The latter also provide non-parametric discrete time duration estimators adaptable to this paper's setting and discuss the trade-offs between discrete and continuous time estimation approaches.

[22] R programming code is available at www.skastoryano.com.

While non-treated causal effects taking the form $\prod_{t=\tau_1}^{\tau_2} \Pr(T > t | S > t, T \geq t, Z = z)$ are computed as,

$$\sum_{\{i\}} \hat{w}_i(s) \exp(-\int_{\tau_1}^{\tau_2} \hat{\theta}_t^T(x_i, z, S > \tau_2) dt) = \sum_{\{i\}} \hat{w}_i(s) \exp(-\int_{\tau_1}^{\tau_2} \hat{\lambda}_t^T(z_i, 0) \exp(x_i' \hat{\beta}^T) dt)$$

with weights given by,

$$\hat{w}_i(s) = \frac{\hat{\theta}_s^S(x_i, 1) \exp(-\int_0^s \hat{\theta}_\tau^S(x_i, 1) d\tau)}{\sum_{\{i\}} \hat{\theta}_s^S(x_i, 1) \exp(-\int_0^s \hat{\theta}_\tau^S(x_i, 1) d\tau)}$$

Finally, I use the delta method to compute standard errors around the causal effects of interest.[23]

Appendix G provides some simulation results assessing the robustness of the estimation method on simulated data generated from a dynamic discrete choice search model. It also briefly discusses how the parameters of the above estimation relate to the identifying assumptions and to the search model. The estimator performs well when the underlying data generating process is not excessively non-linear.

4 Data

This study uses data from the Scientific Registry of Transplant Recipients (SRTR). The SRTR data system includes data on all donor, wait-listed candidates, and transplant recipients in the US, submitted by the members of the Organ Procurement and Transplantation Network (OPTN). The Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services provides oversight to the activities of the OPTN and SRTR contractors.

The dataset contains detailed information on candidate characteristics, as well as information on the date a candidate enters the waitlist, when they receive a kidney transplant, and the date of death. The initial analysis includes all candidates who entered the waitlist between June 1st 2002 and December 1st 2014, the month in which the new kidney allocation system was implemented. The analysis excludes any candidate who was scheduled to receive multiple transplants and candidates who were under 18 years of age, as both of these groups follow special allocation rules. The baseline analysis also excludes candidates who receive a transplant from a living donor since those candidates effectively jump the line based on unobservables.[24] Finally, the analysis only include first spells on the waitlist thereby excluding repeated observations for candidates who experience graft

[23] To obtain the expected value of the effects over an interval of treatment times $[1, s]$, one should replace the denominators in the weights with a sum over $t = 1, \ldots, s$, $\sum_{t=1}^{\bar{s}} \sum_{\{i\}} \hat{\theta}_t^S(x_i, 1) \exp(-\int_0^t \hat{\theta}_\tau^S(x_i, 1) d\tau)$.

[24] Note that this data selection may be endogenous to the blood type if candidates with a different propensity to finding a suitable kidney have different propensities to seeking out living donors. I return to this point in the results section.

failure or renewed kidney failure and the limited group of candidates with rare subgroup blood types. It also excludes the small fraction of immune sensitised candidates with a first cPRA score above 0 given that these candidates face special allocation rules. Appendix B provides a full description of the data selection and the fraction of dropped observations at each step.

The main decomposition estimation takes the unit of time to be two months. The outcome duration T of interest is the time from the moment a candidate with kidney failure enters the waitlist until the time of death, if observed. In the SRTR data, information on the date of death is obtained from social security records (Massie et al., 2014)

The duration to treatment S is the time until a candidate receives a kidney transplant. The regime randomization Z in the baseline analysis is the biological randomization to a blood type AB or O. I purposively choose these two groups to draw inference from pre-transplant differences in survival. As illustrated in Figure A3a of Appendix D, AB blood type candidates, who are universal recipients, have the highest hazard to transplant over the first 3 years on the waitlist among all ABO blood types. O blood type candidates and B blood types, as will be discussed later, have the lowest propensity to receive a transplant, while A blood type candidates have a propensity lying between that of AB types and the others.

The main specification controls for every health related variable relevant to the offer allocation mechanism and other prominent variables related to a candidate's health which do not contain excessive amounts of missing values. Given the sample selection, only the candidate age variable is relevant to offer allocations prior to the 2014 reform. In the post-2015 analysis, the only offer relevant variable in the selected sample is the Estimated Post Transplant Survival (EPTS) score, which is the candidate health score used to match donors to candidates.[25] In addition, the model includes time $t = 0$ variables for: age, diabetes status, whether the candidate is on dialysis, whether they previously received a transplant, BMI, biological sex, race, whether they previously had a malignant tumour, whether they would accept a kidney from a hepatitis B positive donor.[26] At no point does the model control for intermediate variables which require stringent assumptions to be considered exogenous.

The main set of variables for each blood type are described with summary statistics in appendix B Figure A1 columns 1-4. The average age of candidates on the waitlist is just under 52 years of age, 60% of which are men. Most candidates are on dialysis upon entering the waitlist ($> 90\%$) and slightly over a third of them have diabetes upon entering the waitlist. In general, different blood types are comparable along all available

[25]EPTS scores are calculated as follows: $EPTS = 0.047 \cdot \max(Age - 25, 0) - 0.015 \cdot Diabetes \cdot \max(Age - 25, 0) + 0.398 \cdot PriorOrganTransplant - 0.237 \cdot Diabetes \cdot PriorOrganTransplant + 0.315 \cdot \log(YearsonDialysis + 1) - 0.099 \cdot Diabetes \cdot \log(YearsonDialysis + 1) + 0.130 \cdot (YearsonDialysis = 0) - 0.348 \cdot Diabetes \cdot (YearsonDialysis = 0) + 1.262 \cdot Diabetes$.

[26]The model does not include other relevant variables which present a large fraction of missing values: hypertension (26% missing), creatinine measure of candidate (33% missing). The main analysis also drops the limited (6.7%) of candidates who state they would accept a kidney from Hepatitis C donor.

health measures, but are not balanced on race. The O and A blood type groups have lower shares of minorities, with a relatively larger share of Black people in the B blood type group, and higher share of Asian people in the AB and B blood type groups. ABO blood types are also slightly unbalanced on education level, a point we return to frequently in the analysis.

5 Empirical assessment of selection hypotheses

The following section discusses and tests the various theories consistent with the patterns observed in Figures 1a and 1d of the introduction. It further relates these hypotheses and their necessary assumptions to the current discussions about reforms to the organ transplantation system in the US.

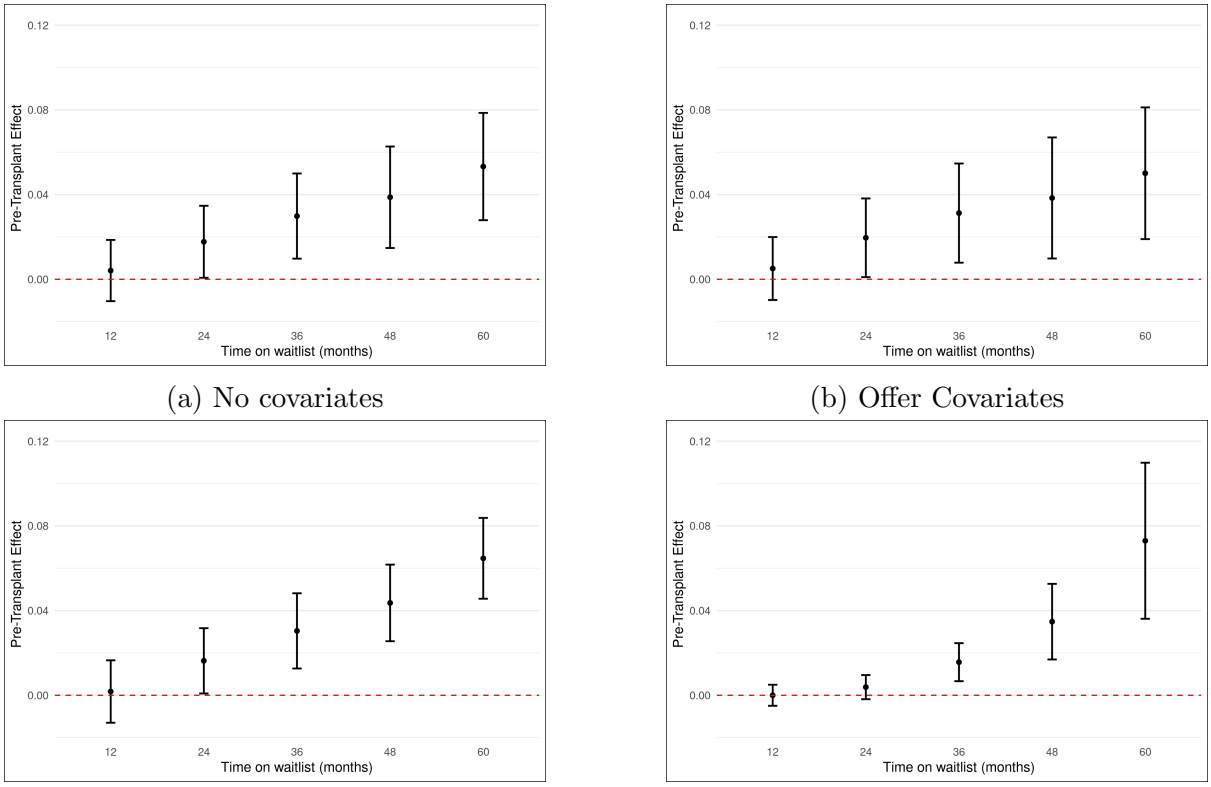
5.1 Selection on kidney observed variables

A first point to inspect is how much of the difference in pre-transplant survival displayed in Figure 1d can be explained by selection on observed variables used in the allocation of kidney offers or other prominent health variables. Figure 2a presents the pre-transplant effect, β_z in the decomposition, on survival past one to five years when excluding any covariates. Figure 2b presents these same effects when adding covariates which are relevant in the allocation of kidney offers.[27] Prior to 2015, and given the data selection, this includes only age which is introduced into the model as a series of age bracket indicators. The results indicate that O-blood types have a higher survival in both specifications.[28]

The point estimates in Figures 2b indicate that the difference in pre-transplant survival of candidates with O blood type relative to AB blood types increases to 5.0 percentage points over the first 5 years on the waitlist. Given the baseline survival of the O blood group in Figure 1d, selection on unobserved variables leads to a 25 percent higher pre-transplant probability of death for AB blood type candidates. Figure 2c adds all additional $t = 1$ baseline covariates related to candidate's health in order to inspect how much remaining variation can be captured by major health risk factors. The results do not indicate that the additional set of covariates substantially affect the pre-transplant estimates. The results also indicate the effects remain in Figure 2d when considering the sample of candidates who entered the waitlist after the 2014 changes to the kidney allocation system. This result is consistent with the agent based structural model in Agarwal et al. (2021) showing the new system only marginally improves patient welfare and discard rates relative to the pre-2014 mechanism. Hereafter, all modelled pre-transplant

[27] I present in Table A2 of Appendix E the full set of decomposition effects.

[28] In the estimation, I specify the segments of the duration dependence terms to be one year each. This choice is guided by the desire to calibrate the model on the observed patterns of Figure 1d in the estimation without covariates. The estimate of β_z in 2a approximates well the observed patterns in Figure 1d.



(a) No covariates

(b) Offer Covariates

(c) All Covariates

(d) Post-2015

Figure 2: Pre-transplant difference in survival effect β_z

Includes candidates entering waitlist after 01 June 2002. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.

(a) includes no covariates, (b) includes offer covariates: age bracket dummies, (c-d) includes all variables presented in Appendix Table A1 omitting eduction variables. Observations (a-c): $N_{O:no-Tr} = 93,922$, $N_{O:Tr} = 35,989$, $N_{AB:no-Tr} = 5125$, $N_{AB:Tr} = 5032$. Observations (d): $N_{O:no-Tr} = 43,338$, $N_{O:Tr} = 11,698$, $N_{AB:no-Tr} = 2241$, $N_{AB:Tr} = 2176$.

effects include all covariates.[29]

Taken together, the results from Figures 2a-2d show a pre-transplant effect which cannot be readily explained by selection into treatment on the basis of variables used when allocating kidney offers nor on the basis of other major risk factor variables influencing survival. Figure A6 of Appendix D shows that these effects appear for White people and minorities, and on both sexes, albeit noticeably lower for females. The results also remain robust to including the potentially non-random candidates who receive a living donor kidney transplant. Given that selection patterns cannot be readily explained by observed variables, the analysis proceeds to test several hypothesis for selection on unobserved variables.

[29]The estimated model does not include education variables in the covariate set. Appendix C discusses this choice in length while presenting an additional set of results when including education variables. The main problem is that the education variables may be bad controls if they are proxies for unobserved selection variables which are the object of interest in this study.

5.2 Selection in response to institutional incentives hypothesis

A first hypothesis which can explain pre-transplant effects is that institutional incentives are leading to selection on unobserved variables. Let us first assume agents are not sophisticated in the sense that they do not adjust their behavior in response to changes in the kidney offer rate. Thus, the kidney quality acceptance threshold is invariant to the offer rate across blood-type regimes. We further assume that pre-transplant survival is increasing in unobserved variables. These will be unobserved health factors known by the candidate and their clinicians, but not relevant in the kidney offer allocation mechanism.

Under this assumption, the pre-transplant empirical patterns documented in Figure 1d of the introduction and the modelled effects of Figure 2 will only arise if the probability of accepting a transplant is also increasing in unobserved health variables. If the probability of accepting a transplant is increasing in unobserved health variables, and the AB-type candidates have a higher hazard to transplant, then the distribution of the non-transplanted AB type group at any time $t > 0$ will be skewed towards less healthy individuals. As a result, the AB type group will have lower pre-transplant survival. Table A3 of Appendix F presents a simple stylised one period example to gain intuition for these survival dynamics.

5.2.1 Do institutional incentives support the assumptions?

Several features of the kidney waitlist suggest transplants are, indeed, increasing in unobserved health due to institutional incentives influencing candidate and transplant center decisions. In particular, candidates must be informed by law if they are being offered certain sub-optimal kidneys (OPTN Final rule 121.11(b)(iV)). Because the new kidney allocation system after 2014 prioritizes longer potential life candidates, unhealthier candidates are more likely to be offered, and therefore informed, of these sub-optimal kidneys. If we assume, as seems plausible, that being informed about an offer of a sub-optimal kidney leads to higher rejection independent of underlying health and offered kidney quality, then we would expect a higher rejection rate from low-health individuals.

On the side of transplant centers, prior to July 2023, transplant center assessments were solely based on a 1-year post-transplant risk-adjusted candidate survival score. This score indicated the relative survival success of transplanted candidates in a center over one year post-transplant adjusting for the donor kidney quality, the candidate health, and the interaction between the two. If selection were purely on observed variables, then the predictive statistical score model should not, in principle, have incentivized transplant centers to favour healthier kidneys or healthier candidates. However, selection is likely not only on observed variables. For example, financial costs of a transplant for transplant centers are not included in the score model and these have been shown to be higher in absolute and marginal terms after medicare compensation for lower health candidates (Axelrod et al., 2017, 2018). As a result, a statistical model trained on ac-

cepted kidneys will predict unusually high expected survival since candidates receiving a transplant, especially at the lower end of the health distribution, likely display positive skewed unobserved characteristics.[30] This leads to a feedback loop. Transplant centers have an objective incentive to only allocate kidneys to high unobserved health candidates which further skews the updated OPTN score model, and further forces transplant centers to focus on healthy candidates. These dynamics may explain the increasing discard of healthy donated kidneys from 18% in 2011 to 25% in 2021 despite the growing number of people on the waitlist.

Beyond these tangible incentives, there is also suggestive evidence of bias in decision-making on behalf of candidates and clinicians in transplant centers. A first piece of evidence is from Zhang (2010) who shows that, for two kidneys from a same donor, the probability of rejecting a kidney increases with the amount of previous rejections when one kidney was randomly rejected more often in initial offers. This suggests the number of rejections, rather than the actual quality of a kidney, is sometimes erroneously used as a heuristic for rejection decisions by either the patient, the transplant center, or both.

Recent changes in the measures used to evaluate transplant center performance and initiate official enquiries also pre-suppose transplant center bias in decision making. Since July 2023, publicly displayed performance reviews and official enquiries for transplant centers include a risk-adjusted offer-to-acceptance rate. In addition, in July 2024, an additional score was implemented based on the pre-transplant mortality of candidates on a transplant center's waitlist. These developments, explicitly focused on reducing the excessive waste of viable kidneys, suggest OPTN is attempting to nudge transplant centers to focus more on the mortality and outcomes of less healthy candidates.[31]

While the adjustments may partially be attempting to counterbalance financial incentives, there is also reason to believe they are attempting to counteract behavioral biases in the understanding of the 1-year risk adjustment scores. When it comes to candidate and clinician decisions based on observed metrics, evidence suggests, unsurprisingly, that both are partial to healthy kidneys (Mohan et al., 2018; King et al., 2023).[32] More problematically, anecdotal reports suggest clinicians misinterpret the risk-adjusted 1-year survival score and replace it with the simpler notion of a 1-year survival score (Mohan et al., 2018; Kimberly et al., 2018), which would erroneously incentivize transplant centers to overtly favour healthier kidneys and healthier candidates. As yet, there exists no literature on candidate and clinician behavioral biases in the context of transplants.

All of the previously described mechanisms also create a chain reaction. If transplant centers and candidates are less likely to accept sub-optimal kidneys, then the 56 organ procurement organizations - government-chartered nonprofits that collect, test and dis-

[30] The risk adjusted score model and some of its potential pitfalls are further discussed in appendix I.

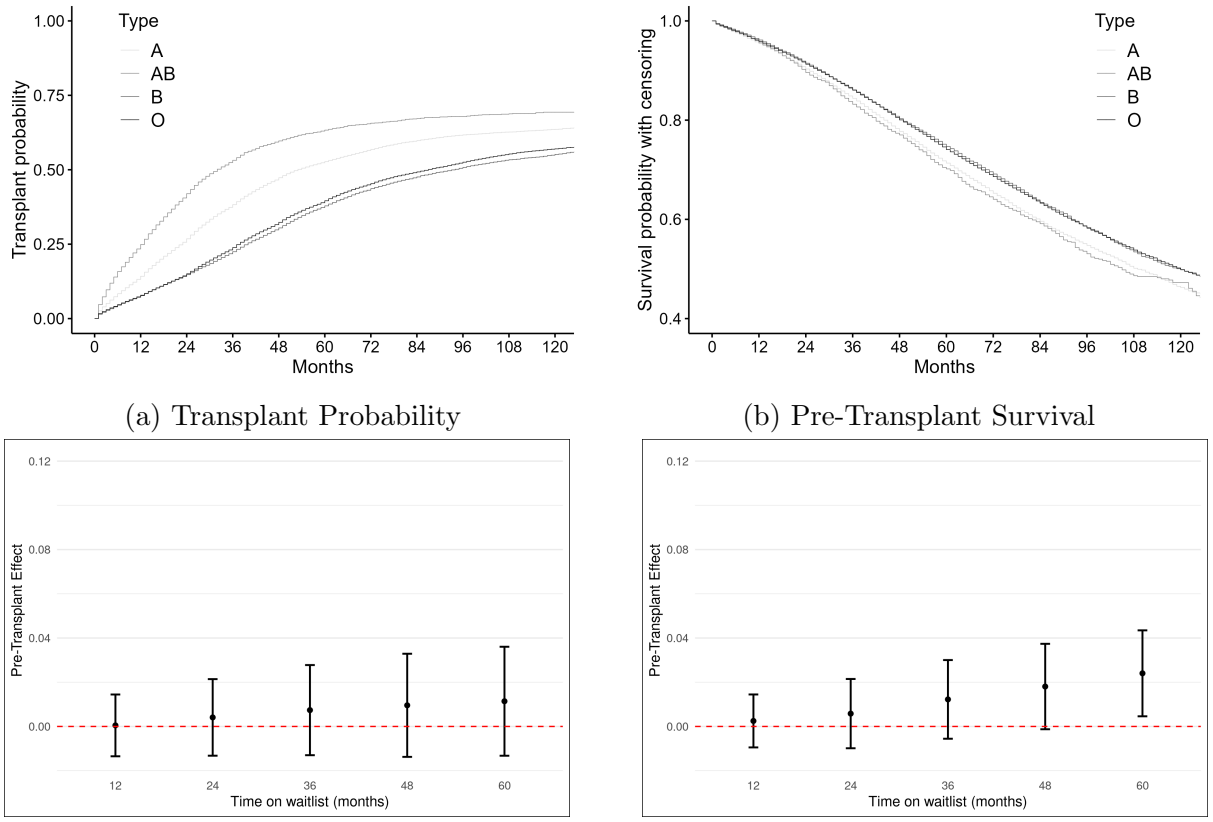
[31] See discussion of the Membership and Professional Standards Committee here and formal information guidelines about change here.

[32] There is also existing evidence of transplant centers attempting to game UNOS/SRTR evaluations (Stith and Hirth, 2016).

seminate organs in their regions - also have lower financial incentives to retrieve these kidneys (Kimberly et al., 2018). This reinforcing process would lead to fewer kidney offers for lower health individuals.[33]

5.3 Unobserved blood-type correlated variables hypothesis

While the previously described incentives form a compelling explanation, a methodical evaluation of selection requires exploring other hypotheses consistent with the observed patterns of Figure 1. One alternative hypothesis is that there are intrinsic unobserved baseline features correlated to people's blood types which affect survival. As discussed previously, evidence for overall differences in mortality across blood types is lacking, but may be different for the specific populations on the transplant waitlist.



(a) Transplant Probability

(b) Pre-Transplant Survival

(c) Pre-Transplant: O vs. B

(d) Pre-Transplant: O vs. A

Figure 3: Treatment probability and Pre-transplant survival effects β_z for A and B blood type candidates

Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B omitting education variables.
(c-d) include all variables presented in Appendix Table A1. Observations (c): $N_{O:no-Tr} = 93,922$, $N_{O:Tr} = 35,989$, $N_{B:no-Tr} = 28,432$, $N_{B:Tr} = 10,303$. Observations (d): $N_{O:no-Tr} = 93,922$, $N_{O:Tr} = 35,989$, $N_{A:no-Tr} = 50,005$, $N_{A:Tr} = 31,637$.

The argument of correlated baseline confounding appears unconvincing on the basis of pre-transplant survival patterns in relation to transplant rates across blood-types.

[33]Given the increase in discarded kidneys over time, this effect is likely of second order importance.

To illustrate this argument, Figures 3a-3b present differences in transplant probabilities alongside differences in pre-transplant survival for all major ABO blood types. Figure 3c further shows the modelled pre-transplant effect of O vs B blood types and Figure 3d shows that same effect for O vs A blood types. Taken together, we notice there is a one-to-one ordinal relation between the transplant probabilities, pre-transplant survival rates, and modelled effects. The higher the transplant probability, the lower the pre-transplant survival probability. These differences cannot be attributed to differences in racial compositions between blood types since, as presented in Figure A4 of Appendix D, the results on the subsample of White people follow the same patterns.[34]

It is difficult to conjure compelling arguments for why, even after conditioning on socio-economic variables, intrinsic unobserved variables simultaneously affect variation in transplant rates and covariation in pre-transplant survival rates, and, coincidentally, that these so happen to form a one-to-one ordinal relation with respect to blood type compatibility in the world. A more likely explanation, as discussed in the previous section, is that differences in transplant rates across blood types are influencing the type and share of candidates receiving a transplant. This selection in turn determines which candidates remain on the waitlist which influences the measured pre-transplant survival.

5.4 Selection and the effect of a transplant on survival

So far, the analysis and discussion of selection have remained disconnected to patterns of overall survival in Figure 1c. These overall survival patterns, however, carry important information with regards to selection. To further inspect these patterns the analysis considers results from the following equation:

$$\mathrm{Y}_i = \alpha + \delta\, \mathrm{Z}_i + \beta\, \mathrm{X} + u_i \tag{3}$$

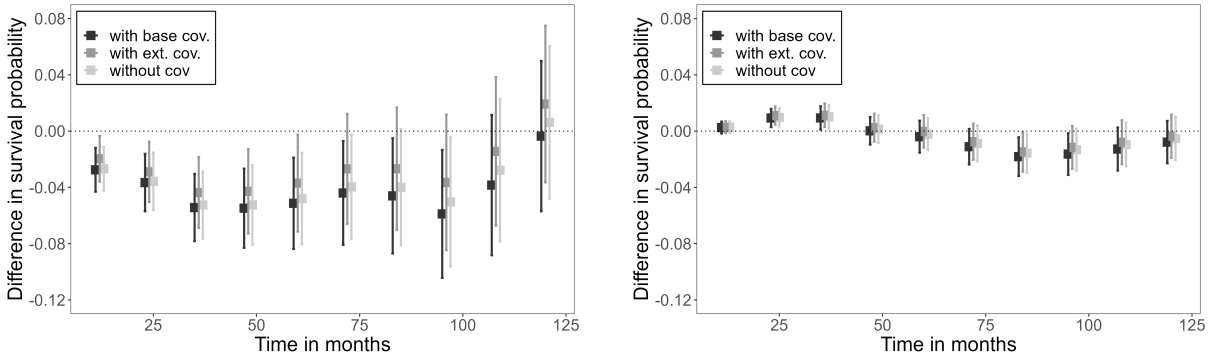
In this model, $\mathrm{Y}_i = \mathbb{1}(T_i > \tau)$ is an indicator for survival past a certain duration τ. Z_i is the blood type indicator with $\mathrm{Z}_i = 1$ for a O-type candidates and $\mathrm{Z}_i = 0$ for AB-type candidates. X, depending on the specification, includes all variables relevant to the transplant allocation mechanism, as well as other prominent mortality related variables. The parameter of interest, δ, represents the difference in the probability of survival past τ for O blood types compared to AB types.

The first set of results, presented in Figure 4a, conditions the analysis on candidates receiving a transplant within the first three months on the waitlist.[35] We see that transplanted O blood types have a sustained lower survival than transplanted AB blood types. The second set of results in Figure 4b present differences in overall survival between O

[34]The appendix also presents these same patterns when conditioning on a balanced subgroup in terms of education in Figure A5 of Appendix D.

[35]Recall that we cannot compare transplanted candidates across blood types at later time periods since, due to the previously described dynamic selection, there will be relatively fewer surviving positively selected AB blood type candidates than O blood type candidates remaining on the waitlist.

and AB blood types. We notice a slightly oscillating, but not consequentially different, pattern of overall survival between both blood groups.



(a) Survival for early transplanted

(b) Overall Survival

Figure 4: Difference in survival between AB and O blood types over time

Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.
(*without cov*) includes no covariates, (*with base cov*) includes offer covariates: age bracket dummies, (*with ext. cov*) includes all variables presented in Appendix Table A1. Observations (a): $N_O = 3757$, $N_{AB} = 955$. Observations (b): $N_O = 125324$, $N_{AB} = 9697$.

We can ask whether the results from Figures 4a-4b are consistent with the institutional incentive selection hypothesis. Consider first the results of Figure 4a. Assume that candidates and/or clinicians make a simultaneous decision over several kidney offers in a non-null amount of cases.[36] Then, because there are more kidney offers in the AB-blood type regime, the expected value of the chosen maximum quality kidney will be higher for AB types than for O types with similar characteristics.[37]

Thus, assuming the distribution of offered kidneys in terms of quality is the same in the AB and O group, expected survival should be higher for AB blood candidates who receive a transplant than for a comparable group of O blood individuals receiving a transplant. However, if we assume a top down kidney allocation strategy first prioritizing transplants for healthy candidates before moving on to unhealthy candidates, then a larger share of marginally less healthy candidates will receive a transplant in the AB group. This phenomenon can be seen in terms of observed variables in Appendix B Table A1 when comparing the full sample of O and AB blood groups (columns 1-2) to the transplanted in the first three months for each group (columns 5-6). We see that the transplanted are positively selected on nearly all health related proxy measures (education, age, bmi, diabetes). What is more, the differences in means shows that O blood types receiving a transplant are relatively more positively selected than AB blood types.

The results from Figure 4a can therefore only hold if, when it comes to survival,

[36] Based on Agarwal et al. (2021), a crude estimate for the average number of kidney offers per candidate a month is around 17-18. If the trend in the transplant hazards in Figure 1a are a close proxy to the trend in kidney offer rates on the waitlist, the average numbers are slightly higher in initial months. Because candidates and/or transplant centers receive offers some days in advance, it is indeed likely some decisions are made jointly over multiple kidneys.

[37] A simple proof for this result is presented in Appendix H.

the higher quality kidneys offered over-compensate for the larger share of unhealthy candidates receiving a transplant in the AB blood group. Again, selection on observed variables in Appendix B Table A1 seems to corroborate the theory that AB candidates receive better kidneys. This is reflected in the lower share of transplanted AB type candidates willing to accept a kidney from a Hepatitis C (HCV+) donor, despite the older and less healthy population compared to the O blood group based on age and diabetes status.

Moving to the results of Figure 4b, we can ask what the institutional incentive selection hypothesis in conjunction with results on transplanted candidates in Figure 4a imply concerning the effect of a transplant on survival. In particular, note, in addition to the early transplanted effects in Figure 4a, that AB-blood candidates as a whole have a higher transplant rate over the first four years on the waitlist (as shown in Figure 1a), but do not have a higher overall survival rate than O candidates. For these patterns to be consistent under a strategy which first prioritizes transplants to healthy candidates, it must hold that, on average, AB type candidates receiving a transplant past the initial months experience a negative effect of a transplant on survival. Re-stated, it implies that at some point in the unobserved (and observed) health distribution, kidney transplants have a negative effect on survival.

This conjecture is plausible given the risks of a transplant which are likely emphasized for low health candidates. These include the risk of dying during surgery, having a graft rejected, or becoming terminally ill as a result of the required immunosuppressant drugs intended to prevent graft rejection. All of these are likely increasing with the lower quality kidneys offered to low health candidates. Finally, it may also be that low health candidates are more likely to transplant kidneys with higher antigen mismatches.

We can augment equation 3 to informally test for heterogenous effects of transplants on survival depending on candidate health. To do so, I include in the model a binary proxy for low health which is whether a candidate had type-II diabetes upon entering the waitlist.[38] As described in Appendix K, type-II diabetes candidates have a lower pre-transplant and overall survival.[39]

The results, fully presented in Appendix K.1, are consistent with stated conjectures about the heterogenous effects of kidneys on several dimensions. First, when it comes to candidates receiving a transplant in the first three months, the results indicate that high health AB candidates have a higher survival than high health O candidates. For transplanted low-health candidates, however, the results show that the difference in survival between blood types is insignificant and tightly bound near 0. These two results

[38]Diabetes status is only part of the set of variables determining kidney allocations after December 2014, but even then, no distinction is made between type-I and type-II.

[39]Appendix K.2 further applies this heterogenous effect model to assess transplant decisions by replacing the survival outcome with an indicator for treatment in the first three months on the waitlist. Results from this model are primarily discussed to show they are consistent with the institutional incentive hypothesis.

suggest that the higher availability of quality kidneys previously suggested only favours high health candidates in the AB blood group. When it comes to overall survival, the results indicate that high health AB candidates have a zero to positive survival compared to high health O candidates. In contrast, the results indicate that low health AB candidates have a zero to negative survival effect compared to low health O candidates. Given the higher rate of transplants for the AB blood group, these differences fail to reject the argument that transplants have heterogenous and opposite sign effects depending on a candidate's underlying health.

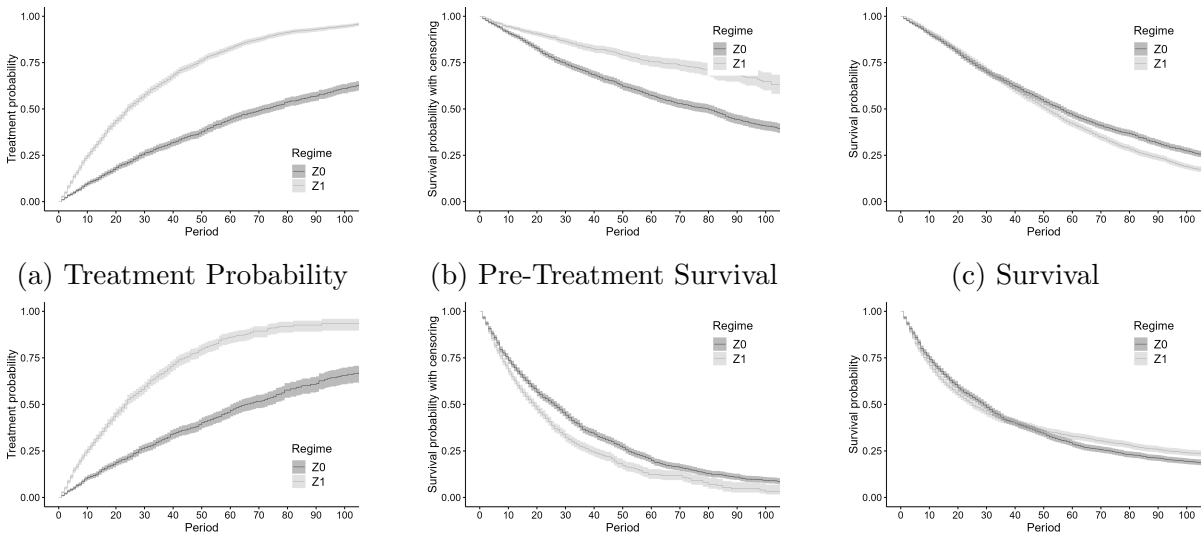
5.5 Lifestyle change behavioral hypothesis

In this last section we turn to alternative incentive-based explanations for the observed pre-transplant patterns. We consider the general setting of optimising forward looking agents within a standard dynamic discrete choice search model (Mortensen, 1986) which is extensively discussed in appendix G.

In short, candidates wish to survive each period but face biological pressures. These pressures can be influenced to some degree by adopting a more or less healthy lifestyle. More precisely, each rational candidate forms a reservation survival value before receiving a transplant. This value is a function of their observed and unobserved (health) types, the mental and physical costs of remaining alive, the arrival rate of health shocks, the probability of receiving a transplant and the rational expected benefit of receiving a transplant. Also, while candidates know the distribution of health shocks they receive each period, they do not perfectly foresee or control these health shocks. A candidate's life ends upon receiving a negative health shock above their reservation survival threshold. From the setup of the model, it follows that poor health candidates are more likely to succumb to a strong negative health shock. In this first model, we assume candidates will accept a kidney if its quality is above an individual specific threshold, and also assume this threshold is invariant to the kidney offer rate.

Appendix G shows that if candidates on the waitlist foresee a kidney transplant to improve their utility of surviving, and they have a high probability of rapidly receiving a kidney transplant (AB blood type), then they will take on a more healthy lifestyle in order to stay alive before receiving the transplant. This result is also illustrated in Figures 5a-5c, where simulations predict a higher pre-transplant survival rate for AB blood-types (Z1), which is contrary to the observed patterns in the data. The pre-transplant survival and overall survival patterns in the data can, in fact, only be reproduced in the rational expectation model if we assume a negative shock from the transplant which is foreseen by candidates, as shown in Figures 5a-5c.[40]

[40] Appendix J discusses a final selection hypothesis. We consider again a forward looking agent model as described in the previous section, but assume reservation survival thresholds are invariant to differences in the kidney offer rate. Instead, the setup assumes only that candidates' kidney quality acceptance thresholds change, and allow this change to be heterogeneous depending on underlying unobserved health. The empirical patterns in Figure 1 of the introduction can be rationalized in this model if, when



(a) Treatment Probability (b) Pre-Treatment Survival (c) Survival

(d) Treatment Probability (e) Pre-Treatment Survival (f) Survival

Figure 5: Simulations with (a-c) positive and (d-f) negative treatment effects
Data generating process and simulation details presented in Appendix G. Observations (a-c): $N_{Z=1:no-Tr} = 1657$, $N_{Z=1:Tr} = 822$, $N_{Z=0:no-Tr} = 738$, $N_{Z=0:Tr} = 1783$. Observations (d-f): $N_{Z=1:no-Tr} = 1919$, $N_{Z=1:Tr} = 560$, $N_{Z=0:no-Tr} = 1488$, $N_{Z=0:Tr} = 1033$.

While the predictions from this model make little sense from the point of view of a candidate, they are consistent with previously stated incentives when considering the transplant center as the main decision maker. The survival patterns will arise in a setting where: a transplant has either a moderate size negative effect for a majority of candidates or a strong negative effect for a smaller share of candidates, and the transplant-center foresees these negative effects but is still constrained to allocate kidneys to such candidates.

6 Conclusion

This paper's analysis shows that ethically motivated institutional considerations of OPTN have led transplant centers to overtly favour matching high-quality kidneys to healthier candidates. Minimizing the number of kidneys allocated to unhealthy candidates may seem sub-optimal at first glance. However, this paper conjectures, based on survival patterns, that kidney transplants lead to shorter life-spans for lower health candidates. As a result, this paper raises important welfare questions in relation to the trade-offs of quality vs. quantity of life, especially when comparing the benefits of transplants to dialysis for low health individuals. Addressing such questions requires a better understanding of the heterogenous effects of transplants depending on candidates' underlying health. This need for further evaluations is urgent since the stated and implemented policies of

facing a higher offer rate, lower health candidates become relatively more selective about which kidneys to accept compared to higher health candidates. However, I argue in Appendix J that both tangible data driven arguments and intangible arguments based on information acquisition do not lend credence to this selection theory when it comes to the effects on overall survival presented in section 5.4.

OPTN (Senate hearing: UNOS) in recent times aim to nudge transplant centers towards focusing more on the mortality and outcomes of less healthy candidates.

In terms of research, the paper advances a particular dynamic causal setup, with corresponding economic theory and dynamic treatment effect models, which allows researchers to methodically assess hypotheses of selection. In cases where selection on unobserved variables is negligible, it also permits a full decomposition of pre-transplant and post-transplant effects. The potential applications of this setup are broad and may support investigations of selection in health, education, labour and finance settings among others.

References

Abbring, J. H. and Heckman, J. J. (2007). Econometric evaluation of social programs, part iii: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. *Handbook of Econometrics*, 6:5145–5303.

Abbring, J. H. and Van den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517.

Abbring, J. H. and Van den Berg, G. J. (2005). Social experiments and instrumental variables with duration outcomes. *Tinbergen Institute Discussion paper no. TI 05-047/3*.

Abegaz, S. B. (2021). Human abo blood groups and their associations with different diseases. *BioMed research international*, 2021:1–9.

Agarwal, N., Ashlagi, I., Azevedo, E., Featherstone, C. R., and Karaduman, Ö. (2019). Market failure in kidney exchange. *American Economic Review*, 109(11):4026–4070.

Agarwal, N., Ashlagi, I., Rees, M. A., Somaini, P., and Waldinger, D. (2021). Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys. *Econometrica*, 89(1):37–76.

Agarwal, N., Ashlagi, I., Somaini, P., and Waldinger, D. (2018). Dynamic incentives in wait list mechanisms. In *AEA: Papers and Proceedings*, volume 108, pages 341–347.

Agarwal, N., Hodgson, C., and Somaini, P. (2020). Choices and outcomes in assignment mechanisms: The allocation of deceased donor kidneys. Technical report, National Bureau of Economic Research.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Ata, B., Friedewald, J., and Randa, A. C. (2020). Structural estimation of kidney transplant candidates' quality of life scores: Improving national kidney allocation policy under endogenous patient choice and geographical sharing. *Structural Estimation of Kidney Transplant Candidates' Quality of Life Scores: Improving National Kidney Allocation Policy Under Endogenous Patient Choice and Geographical Sharing (January 25, 2020)*.

Aubert, O., Reese, P. P., Audry, B., Bouatou, Y., Raynaud, M., Viglietti, D., Legendre, C., Glotz, D., Empana, J.-P., Jouven, X., et al. (2019). Disparities in acceptance of deceased donor kidneys between the united states and france and estimated effects of increased us acceptance. *JAMA Internal Medicine*, 179(10):1365–1374.

Axelrod, D., Schnitzler, M. A., Xiao, H., Naik, A. S., Segev, D. L., Dharnidharka, V. R., Brennan, D. C., and Lentine, K. L. (2017). The changing financial landscape of renal transplant practice: a national cohort analysis. *American Journal of Transplantation*, 17(2):377–389.

Axelrod, D. A., Schnitzler, M. A., Xiao, H., Irish, W., Tuttle-Newhall, E., Chang, S.-H., Kasiske, B. L., Alhamad, T., and Lentine, K. L. (2018). An economic assessment of contemporary kidney transplant practice. *American Journal of Transplantation*, 18(5):1168–1176.

Bernstein, L. (2023). Troubled u.s. organ transplant system targeted for overhaul. *Washington Post*.

Chassang, S., Padró i Miquel, G., and Snowberg, E. (2012). Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review*, 102(4):1279–1309.

Cooling, L. (2015). Blood groups in infection and host susceptibility. *Clinical Microbiology Reviews*, 28(3):801–870.

Danovitch, G. M. (2012). *Handbook of kidney transplantation*. Lippincott Williams & Wilkins.

Dickert-Conlin, S., Elder, T., and Teltser, K. (2019). Allocating scarce organs: How a change in supply affects transplant waiting lists and transplant recipients. *American Economic Journal: Applied Economics*, 11(4):210–239.

Erikouglu, M., Buyukdogan, M., and Tülin, C. (2011). The relationship between hla antigens and blood groups. *European Journal of General Medicine*, 8(1):65–68.

Gill, R. D. and Robins, J. M. (2001). Causal inference for complex longitudinal data: the continuous case. *Annals of Statistics*, pages 1785–1811.

Gordon, E. J., Knopf, E., Phillips, C., Mussell, A., Lee, J., Veatch, R. M., Abt, P., Dunn, S., and Reese, P. P. (2020). Transplant candidates' perceptions of informed consent for accepting deceased donor organs subjected to intervention research and for participating in posttransplant research. *American Journal of Transplantation*, 20(2):474–492.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.

Ham, J. C. and LaLonde, R. J. (1996). The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica*, pages 175–205.

Han, S. (2021). Identification in nonparametric models for dynamic treatment effects. *Journal of Econometrics*, 225(2):132–147.

Heckman, J. J., Humphries, J. E., and Veramendi, G. (2016). Dynamic treatment effects. *Journal of Econometrics*, 191(2):276–292.

Heckman, J. J. and Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2):341–396.

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25(1):51–71.

Israni, A. K., Salkowski, N., Gustafson, S., Snyder, J. J., Friedewald, J. J., Formica, R. N., Wang, X., Shteyn, E., Cherikh, W., Stewart, D., et al. (2014). New national allocation policy for deceased donor kidneys in the united states and possible effect on patient outcomes. *Journal of the American Society of Nephrology: JASN*, 25(8):1842.

Kastoryano, S. and van der Klaauw, B. (2022). Dynamic evaluation of job search assistance. *Journal of Applied Econometrics*, 37(2):227–241.

Kim, J. and Li, M. (2022). Optimal organ allocation policy under blood-type barriers with the donor-priority rule. *Theoretical Economics*, 17(1):331–369.

Kimberly, K., Bernstein, L., and Keating, L. (2018). Lives lost, organs wasted. *Washington Post*.

King, K. L., Husain, S. A., Schold, J. D., Patzer, R. E., Reese, P. P., Jin, Z., Ratner, L. E., Cohen, D. J., Pastan, S. O., and Mohan, S. (2020). Major variation across local transplant centers in probability of kidney transplant for wait-listed patients. *Journal of the American Society of Nephrology: JASN*, 31(12):2900.

King, K. L., Husain, S. A., Yu, M., Adler, J. T., Schold, J., and Mohan, S. (2023). Characterization of transplant center decisions to allocate kidneys to candidates with lower waiting list priority. *JAMA Network Open*, 6(6):e2316936–e2316936.

Lechner, M. (2009). Sequential causal models for the evaluation of labor market programs. *Journal of Business & Economic Statistics*, 27(1):71–83.

Lentine, K. L., Smith, J. M., Miller, J. M., Bradbrook, K., Larkin, L., Weiss, S., Handarova, D. K., Temple, K., Israni, A. K., and Snyder, J. J. (2023). Optn/srtr 2021 annual data report: kidney. *American Journal of Transplantation*, 23(2):S21–S120.

Massie, A. B., Caffo, B., Gentry, S. E., Hall, E. C., Axelrod, D., Lentine, K. L., Schnitzler, M. A., Gheorghian, A., Salvalaggio, P. R., and Segev, D. L. (2011). Meld exceptions and rates of waiting list outcomes. *American Journal of Transplantation*, 11(11):2362–2371.

Massie, A. B., Kuricka, L., and Segev, D. L. (2014). Big data in organ transplantation: registries and administrative claims. *American Journal of Transplantation*, 14(8):1723–1730.

Mohan, S., Chiles, M. C., Patzer, R. E., Pastan, S. O., Husain, S. A., Carpenter, D. J., Dube, G. K., Crew, R. J., Ratner, L. E., and Cohen, D. J. (2018). Factors leading to the discard of deceased donor kidneys in the united states. *Kidney international*, 94(1):187–198.

Mortensen, D. T. (1986). Job search and labor market analysis. *Handbook of Labor Economics*, 2:849–919.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(2):331–355.

Potluri, V. S. and Bloom, R. D. (2022). Effect of policy on geographic inequities in kidney transplantation. *American Journal of Kidney Diseases*, 79(6):897–900.

Rasmussen, S., Thomas, A., Garonzik-Wang, J., Henderson, M., Stith, S., Segev, D., and Nicholas, L. H. (2018). Reported effects of the srtr 5-tier rating system on us transplant centers. In *American Journal of Transplantation*, volume 18, pages 749–750.

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7(9-12):1393–1512.

Robins, J. M. (1997). Causal inference from complex longitudinal data. In *Latent variable modeling and applications to causality*, pages 69–117.

Rosenzweig, M. R. and Wolpin, K. I. (2000). Natural "natural experiments" in economics. *Journal of Economic Literature*, 38(4):827–874.

Roth, A. E., Sönmez, T., and Ünver, M. U. (2004). Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488.

Roth, A. E., Sönmez, T., and Ünver, M. U. (2007). Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences. *American Economic Review*, 97(3):828–851.

Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, 75(371):591–593.

Smith, J., Biggins, S., Haselby, D., Kim, W., Wedd, J., Lamb, K., Thompson, B., Segev, D., Gustafson, S., Kandaswamy, R., et al. (2012). Kidney, pancreas and liver allocation and distribution in the united states. *American Journal of Transplantation*, 12(12):3191–3212.

Sönmez, T., Ünver, M. U., and Yenmez, M. B. (2020). Incentivized kidney exchange. *American Economic Review*, 110(7):2198–2224.

Stith, S. S. and Hirth, R. A. (2016). The effect of performance standards on health care provider behavior: evidence from kidney transplantation. *Journal of Economics & Management Strategy*, 25(4):789–825.

Sun, W., Wen, C.-P., Lin, J., Wen, C., Pu, X., Huang, M., Tsai, M. K., Tsao, C. K., Wu, X., and Chow, W.-H. (2015). Abo blood types and cancer risk—a cohort study of 339,432 subjects in taiwan. *Cancer Epidemiology*, 39(2):150–156.

Teltser, K. F. (2019). Do kidney exchanges improve patient outcomes? *American Economic Journal: Economic Policy*, 11(3):427–453.

Wey, A., Salkowski, N., Kasiske, B., Israni, A., and Snyder, J. (2017). Program-specific offer acceptance behavior for kidney programs across the spectrum of kdpi. In *American Journal of Transplantation*, volume 17, pages 440–440.

Wu, O., Bayoumi, N., Vickers, M., and Clark, P. (2008). Abo (h) blood groups and vascular disease: a systematic review and meta-analysis. *Journal of Thrombosis and Haemostasis*, 6(1):62–69.

Zhang, J. (2010). The sound of silence: Observational learning in the us kidney market. *Marketing Science*, 29(2):315–335.

APPENDIX

A Identification of Causal Effects

In the derivation of identification results we maintain the overlap assumptions in A.II throughout and apply SUTVA/consistency A.III whenever converting potential variables to observed variables. All identification results remain the same when there is censoring if we assume censoring always occurs before treatment and exit at time t, and censored observations are dynamically missing at random. We derive effects in discrete time $t \in \mathbb{N}$ and $s \in \mathbb{N}$.

Proof of Proposition 1 and Corollary 1:

Using the dynamic unconfoundedness assumption A.I, we can straightforwardly derive all causal effects from the general form:

$$\begin{aligned}
\Pr(T^{z,\infty} > \tau) &= \Pr(T^{z,\infty} > \tau | S > \tau, Z = z) \\
&= \prod_{t=0}^{\tau} \Pr(T > t | S > t, T \geq t, Z = z) \\
\Pr(T^{1,s} > \tau) &= \Pr(T^{z,s} > \tau | S = s, Z = z) \\
&= \prod_{t=s}^{\tau} \Pr(T > t | S = s, T \geq t, Z = z) \cdot \prod_{t=0}^{s-1} \Pr(T > t | S > t, T \geq t, Z = z) \\
\Pr(S^z = s) &= \Pr(S^z = s | Z = z) \\
&= \Pr(S = s | S \geq s, T \geq s, Z = z) \cdot \prod_{t=0}^{s-1} \Pr(S > t | S \geq t, T \geq t, Z = z)
\end{aligned}$$

Proof of Proposition 2:

Under assumptions A.I-A.IV, with $T^{1,\infty} \geq T^{0,\infty}$, $\Pr(T^{0,\infty} > \tau | cs_s) = \Pr(T^{0,s} > \tau | cs_s) = 0$ and all probabilities for never-survivors are equal to 0. We then have that,

$$\begin{aligned}
\beta_0 &= \Pr(T^{0,\infty} > \tau) = \Pr(T^{0,\infty} > \tau | as_s) \cdot \Pr(as_s) \\
&= \Big(\prod_{t=s}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 0) \Big) \cdot \Pr(as_s) \\
\beta_s &= [\Pr(T^{0,s} > \tau) - \Pr(T^{0,\infty} > \tau)] = [\Pr(T^{0,s} > \tau | as_s) - \Pr(T^{0,\infty} > \tau | as_s)] \cdot \Pr(as_s) \\
&= \Big(\prod_{t=s}^{\tau} \Pr(T > t | S = s, T \geq t, Z = 0) - \prod_{t=s}^{\tau} \Pr(T > t | S > t, T \geq t, Z = 0) \Big) \cdot \Pr(as_s) \qquad (4) \\
\Pr(as_s) &= \Pr(T^{1,\infty} \geq s, T^{0,\infty} \geq s) = \prod_{t=0}^{s-1} \Pr(T > t | S > t, T \geq t, Z = 0) \\
\Pr(cs_s) &= \Pr(T^{1,\infty} \geq s, T^{0,\infty} < s) = \prod_{t=0}^{s-1} \Pr(T > t | S > t, T \geq t, Z = 1) - \Pr(as_s) \\
\Pr(ns_s) &= 1 - \Pr(as_s) - \Pr(cs_s)
\end{aligned}$$

In these derivations it is useful to note that, for example, $\prod_{t=0}^{s-1} \Pr(T > t|S > t, T \geq t, Z = 0) = 1$ for always-survivors.

Under the additional exclusion restriction A.V we have that $\Pr(T^{1,s} > \tau|as_s) = \Pr(T^{0,s} > \tau|as_s)$ from which we can identify $\Pr(T^{1,s} > \tau|cs_s) = [\Pr(T^{1,s} > \tau) - \Pr(T^{1,s} > \tau|as_s) \cdot \Pr(as_s)]/\Pr(cs_s)$. In addition, assuming that for complier-survivors defined at s, $T^{1,\infty} < \tau$ if $\tau >> s$, it follows that $\Pr(T^{1,\infty} > \tau|cs_s) = 0$. Under these restrictions, we can identify all remaining effects,

$$
\begin{aligned}
\beta_z &= [\Pr(T^{1,\infty} > \tau) - \Pr(T^{0,\infty} > \tau)] = [\Pr(T^{1,\infty} > \tau|as_s) - \Pr(T^{0,\infty} > \tau|as_s)] \cdot \Pr(as_s) \\
&= \Big(\prod_{t=s'}^{\tau} \Pr(T > t|S > t, T \geq t, Z = 1) - \prod_{t=s}^{\tau} \Pr(T > t|S > t, T \geq t, Z = 0) \Big) \cdot \Pr(as_s) \\
\beta_{zs} &= [(\Pr(T^{1,s} > \tau) - \Pr(T^{1,\infty} > \tau)) - (\Pr(T^{0,s} > \tau) - \Pr(T^{0,\infty} > \tau))] \qquad (5) \\
&= \Pr(T^{1,s} > \tau|cs_s)] \cdot \Pr(cs_s) - [\Pr(T^{1,\infty} > \tau|as_s) - \Pr(T^{0,\infty} > \tau|as_s)] \cdot \Pr(as_s) \\
&= \prod_{t=s}^{\tau} \Pr(T > t|S > t, T \geq t, Z = 1) - \prod_{t=s}^{\tau} \Pr(T > t|S > t, T \geq t, Z = 0) \cdot \Pr(as_s) - \beta_z
\end{aligned}
$$

B Descriptive Statistics and Data selection

Table A1: Descriptive Statistics

	Pre-treatment Analysis				Treated at $t = 1$	
	O	AB	B	A	O	AB
$Educ_{HSless}$	0.54	**0.49**	0.51	0.52	0.48	0.48
	(0.5)	**(0.5)**	(0.5)	(0.5)	(0.5)	(0.5)
$Educ_{COLsome}$	0.24	0.26	0.25	0.25	0.26	0.25
	(0.43)	(0.44)	(0.44)	(0.43)	(0.44)	(0.43)
$Educ_{COLdeg}$	0.22	0.25	0.23	0.23	0.27	0.27
	(0.41)	(0.43)	(0.42)	(0.42)	(0.44)	(0.44)
Age	51.32	51.91	51.56	51.90	49.07	**51.31**
	(13.06)	(13.02)	(12.83)	(13.08)	(13.94)	**(13.57)**
$Transp$	0.10	0.12	0.09	0.12	0.11	0.09
	(0.30)	(0.32)	(0.29)	(0.32)	(0.31)	(0.29)
$Diabetes$	0.35	0.37	0.37	0.36	0.25	**0.32**
	(0.48)	(0.48)	(0.48)	(0.48)	(0.43)	**(0.47)**
$Dialysis$	0.92	0.94	0.92	0.93	1.00	1.00
	(0.27)	(0.24)	(0.27)	(0.26)	(0.04)	(0.06)
BMI	28.41	28.40	28.37	28.47	27.44	27.81
	(5.73)	(5.85)	(5.80)	(5.75)	(5.49)	(5.54)
$Female$	0.39	0.38	0.38	0.38	0.39	0.36
	(0.49)	(0.49)	(0.49)	(0.48)	(0.49)	(0.48)
$PrevMalign$	0.06	0.06	0.05	0.07	0.06	0.06
	(0.24)	(0.24)	(0.23)	(0.25)	(0.24)	(0.24)
$Acpt_{HepB}$	0.56	0.54	0.57	0.54	0.48	0.47
	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)
$Acpt_{HCV+}$	0.06	0.05	0.07	0.05	0.09	**0.07**
	(0.24)	(0.21)	(0.25)	(0.22)	(0.29)	**(0.25)**
$White$	0.62	**0.55**	**0.45**	**0.70**	0.78	**0.68**
	(0.48)	**(0.50)**	**(0.50)**	**(0.46)**	(0.42)	**(0.47)**
$Asian$	0.05	**0.12**	**0.12**	0.05	0.03	**0.07**
	(0.22)	**(0.32)**	**(0.33)**	(0.22)	(0.17)	**(0.26)**
$Black$	0.30	0.32	**0.41**	**0.23**	0.18	**0.24**
	(0.46)	(0.47)	**(0.49)**	**(0.42)**	(0.39)	**(0.43)**
$Other$	0.02	0.01	0.02	0.02	0.01	0.01
	(0.15)	(0.12)	(0.12)	(0.13)	(0.11)	(0.09)

Means for each group presented with standard deviations in parenthesis. Normalized differences relative to the O blood-type group greater or equal to 0.1 are marked in bold. Full sample includes 159,150 observations. *Age*: Candidate age at listing, *Transp*: Received previous transplant (other than kidney transplant), *Diabetes*: Had Diabetes upon entering the waitlist, *Dialysis*: On Dialysis upon entering the waitlist, BMI, $Female$, , $PrevMalign$: Any previous Malignancy, $Acpt_{HepB}$: Will accept an Hepatitis B Core Antibody Positive Donor?, $Acpt_{HCV+}$: Will accept an HCV Positive donor?, $White$: Race-White, *Asian*: Race-Asian, *Black*: Race-Black, *Other*: Race-Other Non-White

Main analysis data selection:

The initial data contains 1,010,051 observations. I remove candidates with no activation date (954,406), and select only people set to receive a kidney transplant (894,372). I then select for the initial analysis only candidates entering the waitlist between June 1st 2002 and December 1st 2014 (440,926). Among these, I only keep the first observed kidney transplant (360,171), for candidates who are over 18 (349,831). I further remove candidates with unusual A1, A2, A1B, A2B blood types (344,882). In the AB vs. O blood types analysis, I also remove the A and B blood types (182,394). I drop candidates with a positive cPRA (170,966) and drop candidates who received a transplant from a living donor (145,274). All remaining reductions in sample for the analysis with covariates

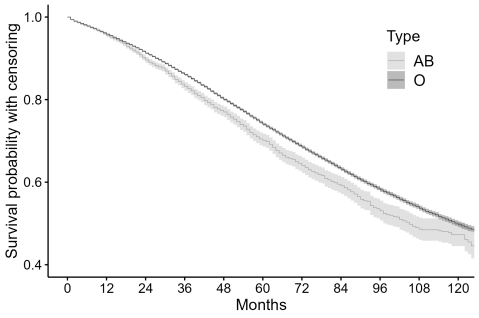
result from missing values in the covariate matrix. These in one or the other specification include: $< 1\%$ of missing values for education, $< 0.01\%$ of missing values for BMI, $< 2.8\%$ of missing values for previous malignancy. I take December 1st 2014 as the 2014 reform cutoff date.

C Selection on education

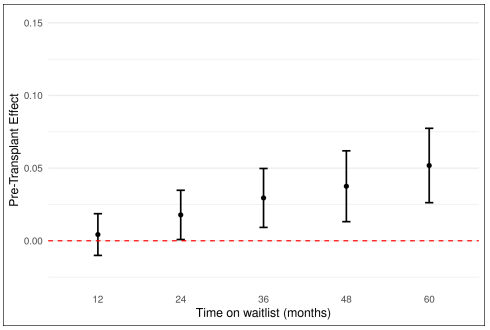
As noted in the main text, including the education variable in the set of covariates may be questionable. Adding them will account for small imbalances between ABO blood types noted in appendix B Figure A1 columns 1-4. However, including them may also proxy for unobserved selection variables which are the object of this study, thereby dimming effects of interest.[41] To examine this issue the analysis produces several results. First, Figures A1a-A1b produce Kaplan-Meier curves and pre-transplant effect estimates after re-balancing the O and AB samples on education. More precisely, the estimation randomly drops O blood candidates with 'high school education or less' until the sample is balanced on all education variables. The measured effects remain the same as the ones in the main text.

We can also examine pre-transplant effects conditioning the sample of the three different education categories. These are presented in Figures A1c-A1h and show that pre-transplant effects are stable across education brackets. The analysis also produces yet another set of results conditioning the analysis on white people who received some college education, the education category highly balanced between AB and O blood groups. Again, the results remain the same as the ones in the main text. If, on the other hand, we add the education covariates in the estimation, then results are dimmed as shown in Figure A2c, and even in Figure A2d for the sample which forces balance of education across AB and O blood groups. All these results suggest including the education variable leads to proxy unobserved selection problems, rather than accounting for slight unbalance across blood groups. A last important point backing this is that the higher share of low educated individuals in the O blood group should, ceteris paribus, lead to lower pre-transplant survival in the O blood group relative to the AB blood group. The potential bias is therefore opposite to measured pre-transplant effects and can therefore not explain the measured selection effect

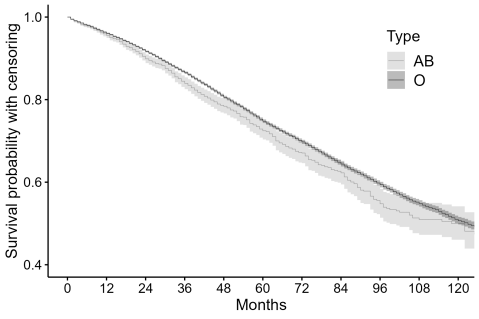
[41] A somewhat smaller issue is that 9% of observations do not include any education information.



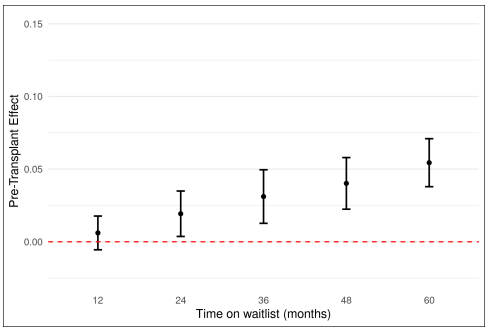
(a) Survival after re-balancing



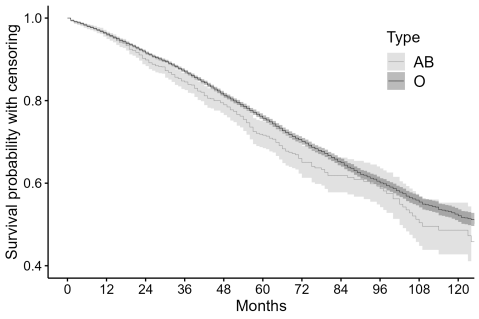
(b) Effects after re-balancing



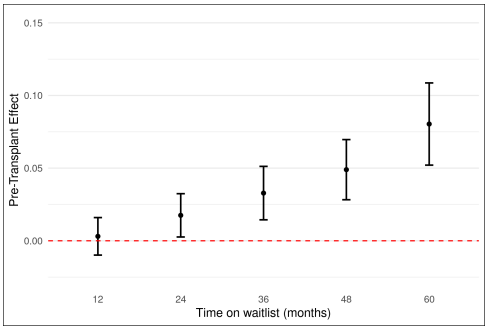
(c) Survival for low educated



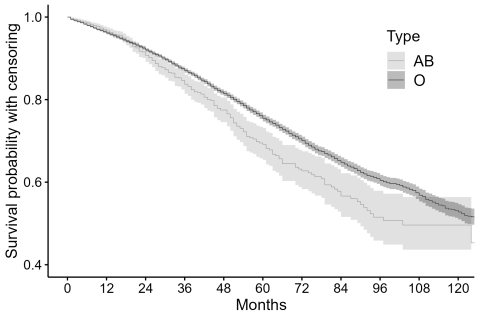
(d) Effects for low educated



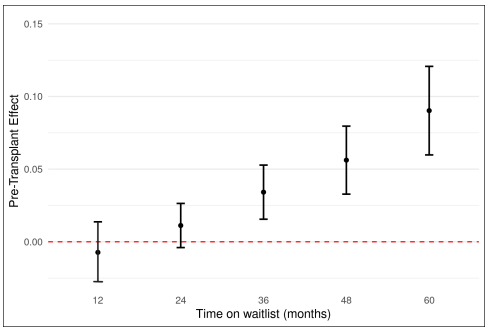
(e) Survival for some college



(f) Effect for some college

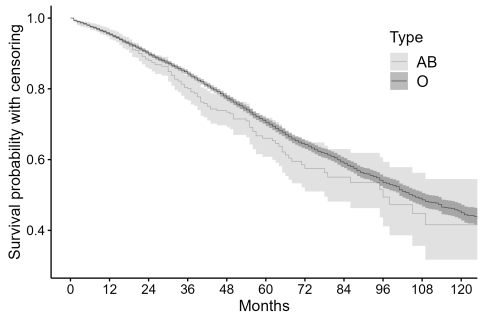


(g) Survival for college degree

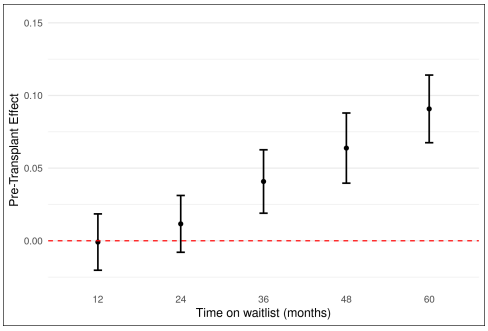


(h) Effects for college degree

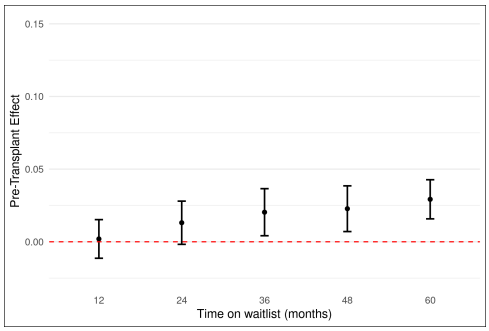
Figure A1: Pre-transplant survival and effect β_z for different subsamples. Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.



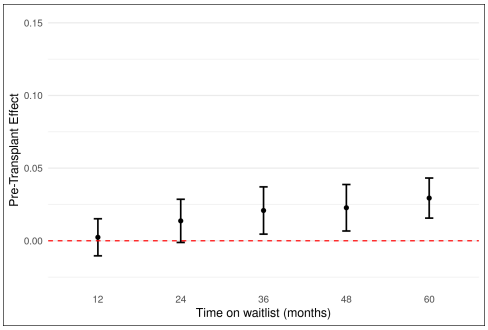
(a) Survival for White, some college



(b) Effects for White, some college



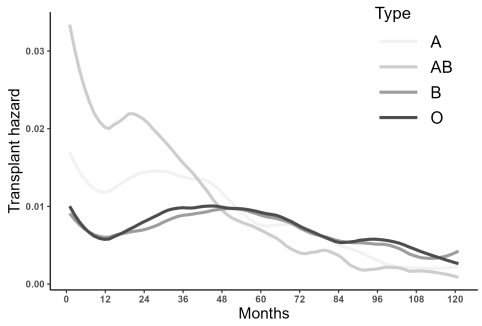
(c) Effects with education covariates



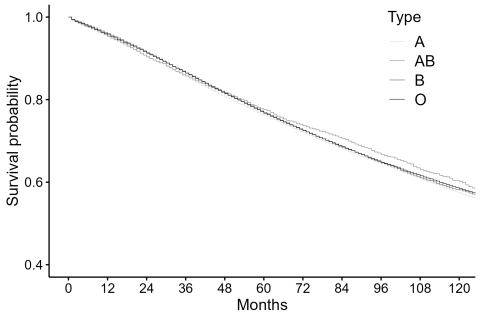
(d) Effects after re-balancing with education covariates

Figure A2: Pre-transplant survival and effect β_z for different subsamples. Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.

D Additional figures and estimates on subgroups



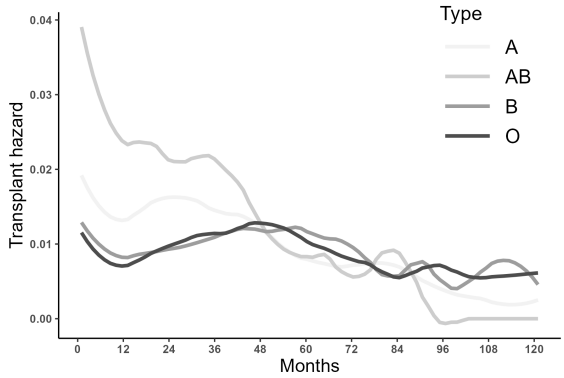
(a) Hazard rate for all blood types



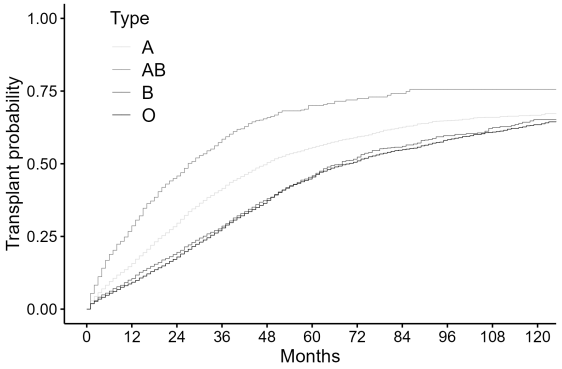
(b) Survival for all blood types

Figure A3: Hazard and overall survival for different blood types

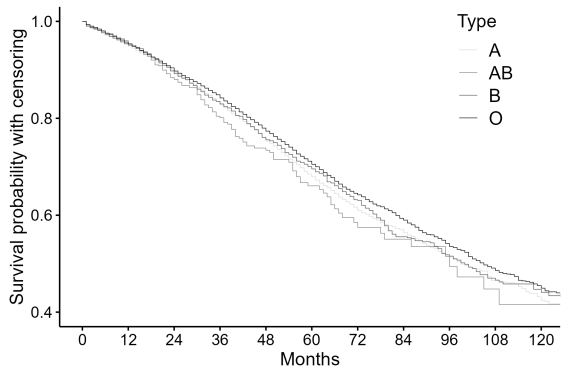
Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B. Observations: $N_O = 134,730$, $N_{AB} = 10,544$, $N_B = 40,217$, $N_A = 84,929$.



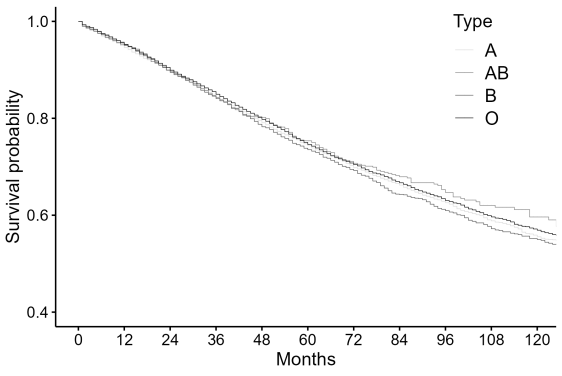
(a) Transplant Hazard



(b) Transplant Probability



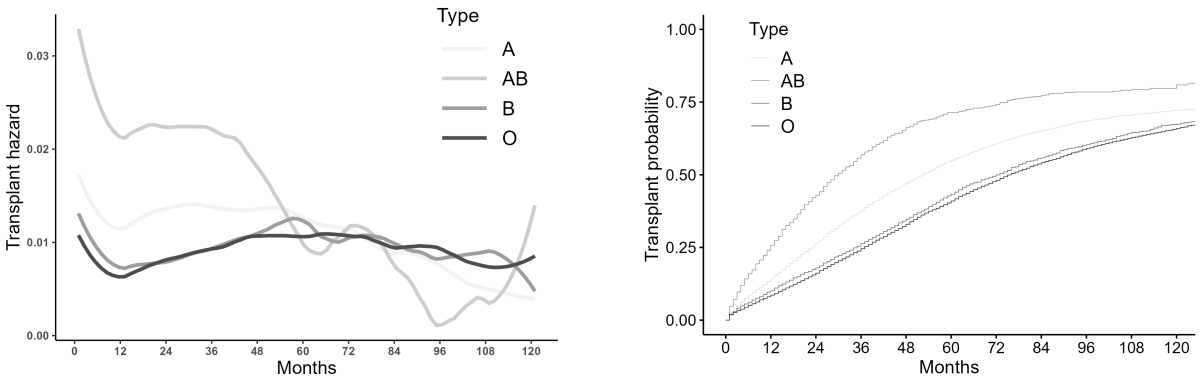
(c) Pre-Transplant Survival



(d) Survival

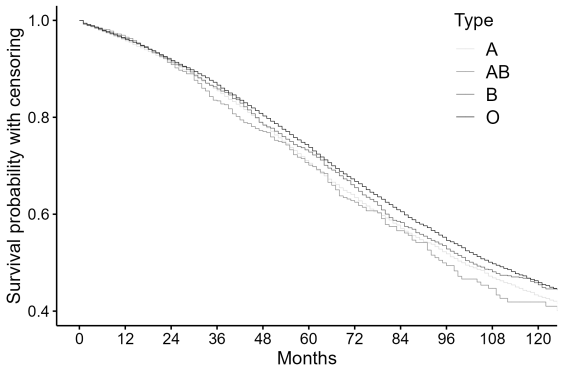
Figure A4: Hazard, transplant and survival probabilities for different blood types conditional on being White

Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B. Observations: $N_O = 23,476$, $N_{AB} = 1680$, $N_B = 5249$, $N_A = 17,885$.

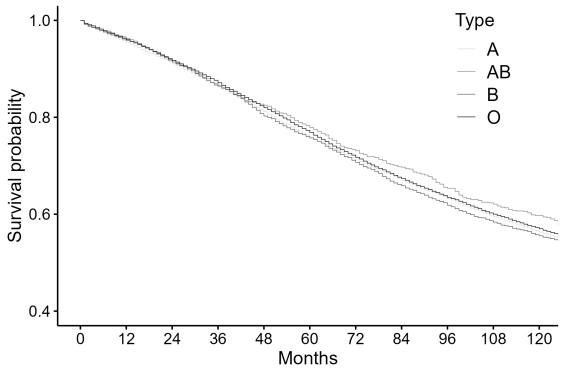


(a) Transplant Hazard

(b) Transplant Probability



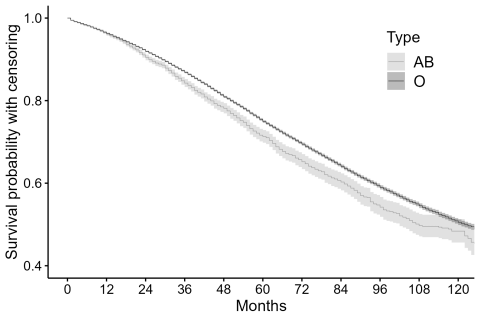
(c) Pre-Transplant Survival



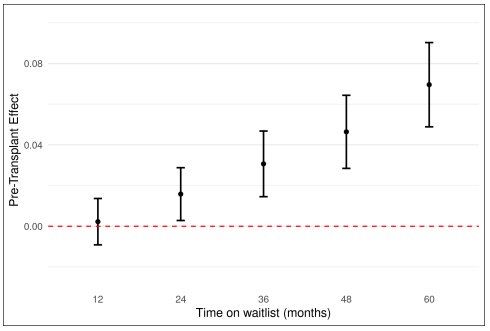
(d) Survival

Figure A5: Hazard, transplant and survival probabilities for different blood types conditional on being White and attending some college.

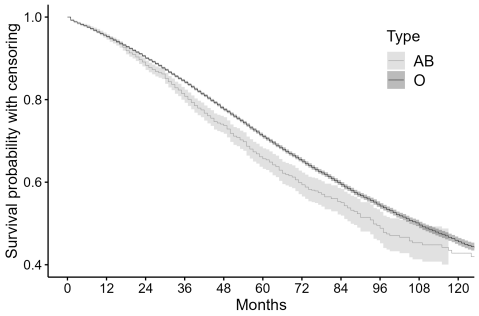
Includes all candidates entering waitlist after 01 June 2002. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B. To increase comparison sample, the Figure includes all candidates entering waitlist pre- and post-reform Observations: $N_O = 80,058$, $N_{AB} = 5385$, $N_B = 16,830$, $N_A = 56,877$.



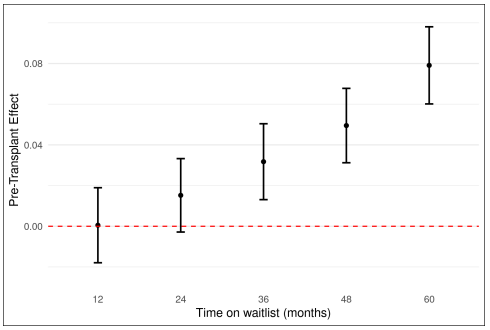
(a) Survival including living donors



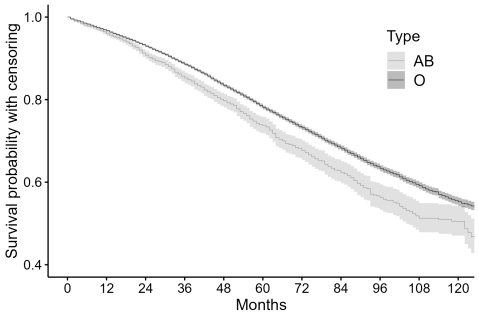
(b) Effects including living donors



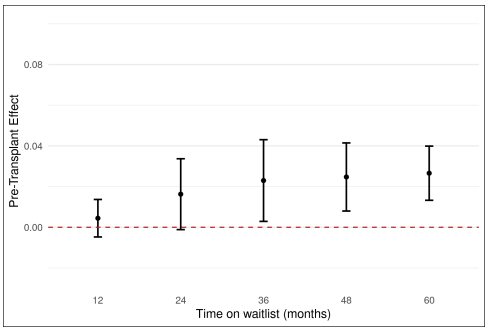
(c) Survival for white people



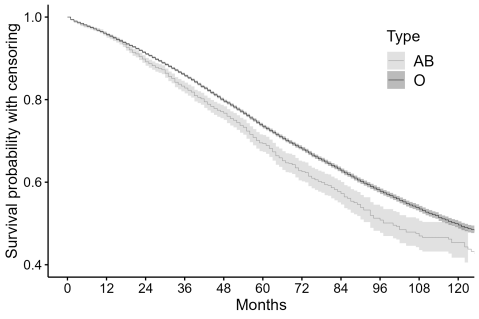
(d) Effects for white people



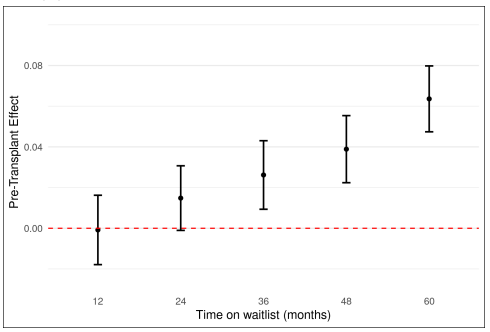
(e) Effects for non-white people



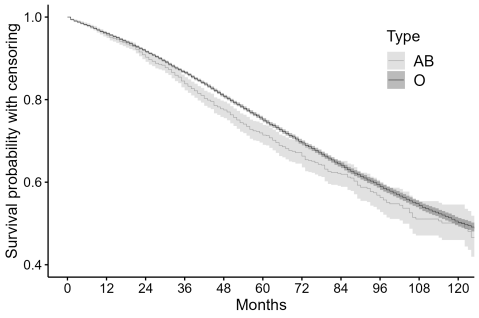
(f) Survival for non-white people



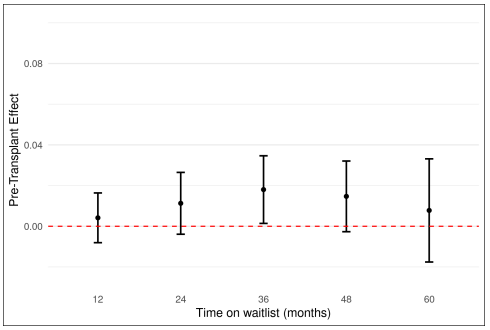
(g) Survival for males



(h) Effects for males



(i) Survival for females



(j) Effects for females

Figure A6: Pre-transplant survival and effect β_z for different subsamples. Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.

E Robustness Checks

Table A2: Causal Effect Decomposition

	(1)	(2)	(3)	(4)	(5)
β_0	0.690	0.679	0.760	0.733	0.713
	(0.043)	(0.150)	(0.104)	(0.133)	(0.138)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β_z	0.053	0.065	0.073	0.011	0.024
	(0.013)	(0.010)	(0.019)	(0.013)	(0.010)
	[0.000]	[0.000]	[0.000]	[0.365]	[0.015]
β_s	0.171	0.187	0.130	0.101	0.126
	(0.020)	(0.075)	(0.050)	(0.042)	(0.050)
	[0.000]	[0.013]	[0.009]	[0.016]	[0.012]
β_{zs}	-0.094	-0.108	-0.099	-0.021	-0.046
	(0.014)	(0.043)	(0.038)	(0.011)	(0.019)
	[0.000]	[0.012]	[0.009]	[0.062]	[0.017]
$N_{Z=0:no-Tr}$	5125	5125	2241	28,432	50,005
$N_{Z=0:Tr}$	5032	5032	2176	10,303	31,637
$N_{Z=1:no-Tr}$	93,922	93,922	43,338	93,922	93,922
$N_{Z=1:Tr}$	35,989	35,989	11,698	35,989	35,989

Standard errors in parenthesis. P-values in brackets. Effect is estimated over 5 years (τ). Each column presents ex-post transplant effects values $\beta_{s=1}$ and $\beta_{z,s=1}$ for transplant in the first two months. Column (1) presents pre Dec. 2014 decomposition with $Z = 0$: AB-blood types, $Z = 1$: O-blood types, without covariates. Column (2) presents pre Dec. 2014 decomposition with $Z = 0$: AB-blood types, $Z = 1$: O-blood types, with all covariates. Column (3) presents post Dec. 2014 decomposition with $Z = 0$: AB-blood types, $Z = 1$: O-blood types, with all covariates. Column (4) presents pre Dec. 2014 decomposition with $Z = 0$: B-blood types, $Z = 1$: O-blood types, with all covariates. Column (5) presents pre Dec. 2014 decomposition with $Z = 0$: A-blood types, $Z = 1$: O-blood types, with all covariates.

F One period example of unsophisticated selection

Here we consider in a simple example how the patterns of Figure 1d can arise even in the absence of sophisticated selection, in the sense that candidates are not adjusting their behavior in response to difference in the kidney offer rate. For each blood type, there are 5 high unobserved health (h) and 5 low unobserved health (l) candidates. In all scenarios we assume survival is increasing in health by imposing that h-types who receive a transplant always survive while l-types only survive if they receive a transplant. In the upper panel, we consider the scenario in which the probability of accepting a kidney is higher for h-types than for l-types by assuming h-types always accept an offered kidney while l-types only accept an offered kidney with a 1/3 probability. Further assume that 6 kidneys are randomly offered in the AB-blood type regime but only 2 are offered in the O-blood type regime. In this setting, for the AB-blood type regime there will be 2 h-types and, in expectation, 4 l-types who do not receive a transplant. Of these only the h-types survive. As such, the pre-transplant survival will be 2/6 in the AB-blood regime. A similar calculation shows that the pre-transplant survival rate in the O-blood regime is 4/9. As such, the difference in the pre-transplant survival $\beta_z = 1/9$ is higher in the O-blood regime as in the empirical results.

A similar exercise is presented in the second panel in which we assume h-types and l-types have the same 1/3 probability to accept an offered kidney. In this scenario, and holding all other previous assumptions, the pre-transplant survival remains the same in the O-type and AB-type regime. In the last panel we consider the case in which h-types are assumed to have a lower probability of accepting an offered kidney than l-types. This scenario results in higher pre-transplant survival in the AB-blood regime than in the O-blood regime. We see that only the first case are compatible with the observed empirical patterns in the kidney setting and the main results.

Table A3: Survival Dynamics in Different Acceptance/Offer Scenarios

	cand.	offered kidneys	accepted kidneys	pre-transpl. survival	pre-transpl. survival ratio	β_z
$\partial \Pr(\text{accept kidney}\|U=u)/\partial u > 0$						
AB	5h	3	3	2h/2h	2/6	1/9
	5l	3	1	0l/4l		
O	5h	1	1	4h/4h	4/9	
	5l	1	0	0l/5l		
$\partial \Pr(\text{accept kidney}\|U=u)/\partial u = 0$						
AB	5h	3	1	4h/4h	4/8	0
	5l	3	1	0l/4l		
O	5h	1	0	5h/5h	5/10	
	5l	1	0	0l/5l		
$\partial \Pr(\text{accept kidney}\|U=u)/\partial u < 0$						
AB	5h	3	1	4h/4h	4/6	-1/9
	5l	3	3	0l/2l		
O	5h	1	0	5h/5h	5/9	
	5l	1	1	0l/4l		

G Dynamic discrete choice model of candidate behavior

To relate the dynamic treatment effect model to the literature on dynamic discrete choice models we consider in this appendix a standard job search model. The model is then extended to include a treatment. I loosely adapt the discussion of the search model to this paper's kidney transplant application. For expositional purposes, the model is solved under simplistic assumptions. The discussion then explains how the dynamic discrete choice model relates to the parameters in the proportional hazard model in Section 3.3. Thereafter I describe the data generating process and present simulation results for the estimator.

G.1 Standard dynamic discrete choice model

Consider an agent who enters an initial state at time $t = 0$ and assume that time is discrete. In each subsequent period the agent faces the choice of staying within this state or leaving. Whenever he is in the initial state, the agent derives utility w_0. In the application, $t = 0$ is the moment a candidate enters the kidney transplant waitlist and each period the candidate, which we consider as a unit combined of psychological and biological factors, puts in a certain amount of effort to remain alive. w_0 represents the combined psychological and biological utility of staying alive.

Next, with probability λ the agent receives an offer to leave the state. An offer can be interpreted as a negative health shock on the body due to kidney failure. The agent also faces a cost c corresponding to the necessary biological effort to prevent a health shock, with lower costs corresponding to higher effort. Once the agent receives an offer, he has to decide immediately whether or not to accept it. An offer is characterized by its instantaneous utility w drawn from the distribution $G(w)$. Upon accepting an offer, the agent derives the same instantaneous utility w for each subsequent period. So once a candidate's biological constitution receives a health shock beyond what it can fight, they die and receive the (perceived) utility of death thereafter.

We assume the agent optimization behavior follows a dynamic discrete choice model which nests search models and optimal stopping models. The agent forms expectations over future instantaneous utilities. Denoting by ρ the discount rate, the present discounted combined psychological and biological value of remaining alive at the start of period t, $V_{0,t}$, can be described by the Bellman equation,

$$V_{0,t} = w_0 - c + \rho\lambda\mathbb{E}[\max\{V_1(w), V_{0,t+1}\}] + \rho(1-\lambda)V_{0,t+1}$$

$V_1(w)$ is the discounted value that the agent would acquire by failing to fight off a health shock, dying as a consequence, and reaping instantaneous biological utility w. The agent follows a stationary reservation utility strategy where w^* is the minimum offer of w required to induce the agent to exit in the following period. In terms of this study, w^* represents the utility associated to the limit at which someone's biological constitution prefers to stop functioning, rather than exert the continued effort to keep a person alive, despite the psychological desire to remain alive. For a healthy person with a well functioning immune system, we would expect w^* to be very high, since only an extreme negative health shock (large w) would induce someone's biological functions to give in. A lower w^* implies, all else equal, that a person's health is worse, making lower utilities from death w more appealing. Under a reservation utility strategy, we can reformulate the Bellman equation as

$$V_{0,t} = w_0 - c + \rho\lambda \int_{w^*}^{\infty} (V_1(w) - V_{0,t+1})\, dG(w) + \rho V_{0,t+1}$$

We can augment this model to include a treatment prescribed by a regime. Let us assume that the agent knows that he has been assigned to a certain regime z which allocates future treatment. For simplicity we consider in this section a stochastic assignment regime where the agent faces the same probability π to receive treatment at each period. So each regime z is fully characterized by its value of π. This type of randomization can correspond as in the empirical setting to a randomization set by nature but may also be a rule imposed by a policymaker. In the empirical setting, it can best be interpreted as a situation in which agents receive different signals of how likely they are to receive

treatment, and interpret this signal as a constant hazard π to treatment each period.

In principle, we could model π from an additional search process of by specifying a search model of offered kidneys for the accepted kidney. I do not add this layer of complexity in order to convey insights for the setting of section 5.5 in which the acceptance/offer rates are invariant to the number of offers.

In the application, candidates on the waitlist were randomized at birth to have an AB blood type or a B/O blood type. This blood type, while inconsequential during most peoples' lives, is an important determinant to the time a candidate must wait until receiving a kidney transplant, which is the treatment in this setting. From the point of view of a candidate, their blood type may be the most salient feature determining the duration until they receive a transplant, but is not the only factor influencing the timing of a kidney transplant. Given the many factors determining the waiting time until a match, the blood type randomization can be seen as a stochastic treatment assignment mechanism since it influences the chances of finding a kidney match but does not determine the exact date a candidate will receive the transplant. Receiving a kidney transplant would likely reduce the arrival of health shocks (λ), change the effort a candidate needs to put into their general health upkeep (search costs c), or change the distribution of the (perceived) utility from death ($G(w)$).

To simplify the exposition, assume treatment only affects the distribution of the (perceived) utility from death by prescribing a higher mean to $G^{tr}(w)$ relative to $G(w)$ for an agent upon receiving the treatment. Allowing agents to form expectations over treatment outcomes, the alive value functions before receiving a kidney transplant, $V_{0,t}$, and after receiving a kidney transplant, $V^{tr}_{0,t}$, are given by,

$$V_{0,t} = w_0 - c + \rho\lambda\mathbb{E}[\max\{V_1(w), (1-\pi)V_{0,t+1} + \pi V^{tr}_{0,t+1}\}] + \rho(1-\lambda)[(1-\pi)V_{0,t+1} + \pi V^{tr}_{0,t+1}]$$
$$V^{tr}_{0,t} = w_0 - c + \rho\lambda\mathbb{E}[\max\{V^{tr}_1(w), V^{tr}_{0,t+1}\} + \rho(1-\lambda)V^{tr}_{0,t+1}$$

We can solve this model for the reservation utilities after and before treatment,[42]

$$w^*(\pi) = \frac{(1-\rho)(1-\pi)}{1-\rho+\rho\pi}(w_0 - c) + \frac{\rho\lambda(1-\pi)}{(1-\rho)(1-\rho+\rho\pi)}\int_{w^*}^{\infty}(1-G(w))\,dw + \frac{\pi}{1-\rho+\rho\pi}w^{tr*}$$
$$w^{tr*} = w_0 - c + \frac{\rho\lambda}{1-\rho}\int_{w^{tr*}}^{+\infty}\left(1-G^{tr}(w)\right)dw$$

Under the assumption that $V^{tr}_{0,t} > V_{0,t}$ one can demonstrate that the pre-treatment reservation utility $w^*(\pi)$ is increasing in π. This means that if agents foresee positive effects of receiving treatment on the w offer distribution, and are in a regime with higher chances to receive treatment, then they will remain for longer in the initial state if they are not yet treated. In terms of the empirical application, this implies that if candidates on the waitlist foresee a kidney transplant to improve their biological utility of life, and they

[42] Full derivation of solutions are presented in subsection G.4

have a high probability of rapidly receiving a kidney transplant (AB blood type), then they will exert more effort each period to stay alive before receiving the transplant. The empirical findings are not consistent with this prediction. They show that candidates with a higher probability of receiving a kidney transplant display a higher pre-transplant mortality. One way to consolidate the empirical results with the biological predictions of the model is to allow a candidate's behavior to asymmetrically influence their biological functions in response to their probability of receiving a future kidney transplant. In particular, we must assume candidates with a high probability of treatment pay less attention to their general health, leading to lower survival.

It is worth considering the interpretation of parameters in the model of subsection 3.3 under different treatment assignment mechanisms prescribed by the regime. If the regime enforces a constant treatment hazard each period and there is full compliance, then $\theta_t^S(z) = \pi^S(z)$. If in addition agents know the regime, do not vary their search strategy over time, and the expected value of future variables over intermediate shocks is constant over time, then $\lambda_t^T(z)$ will be constant. This constant is then interpreted as the effect on the exit hazard of a constant treatment hazard regime. This is the setting considered in the dynamic discrete choice search model above.

G.2 Data Generating Process

The data generating process for the simulation data follows the dynamic discrete choice model presented in the above section with the following specifications,

$$\begin{aligned}
U_{it}^{no-exit} &= w_{0it} - c_{it} \\
U_{it}^{exit} &= w_{it} \\
c_{it} &= \beta_a^c \cdot a_i + \beta_e^c \cdot e_i \\
w_{it} &= \beta_a^w \cdot a_i + \beta_s^w \cdot I(s < t) + \xi_{it} \quad \xi_{it} \sim \mathcal{N}(0, \sigma_\xi^2) \\
w_{0it} &= 0.75 \cdot \beta_a^w \cdot a_i
\end{aligned}$$

where $U_{it}^{no-exit}$ and U_{it}^{exit} are the instantaneous utilities when the agent chooses to remain in the initial state or exit. λ_{it} follows a Poisson distribution with mean $\beta_a^\lambda \cdot a_i + \beta_e^\lambda \cdot e_i$. The treatment outcome for the group under regime $Z = 1$ is drawn each period from a binary distribution with probability $\pi_{Z=1} = Pr(S = s|S \geq s, Z = 1) = 0.03$. The regime assignment for the $Z = 0$ group is $\pi_{Z=0} = Pr(S = s|S \geq s, Z = 0) = 0.01$. I impose that the treatment takes place before the exit decision in period t.

Using this model I generate the treatment durations, exit durations, and accepted offers for a population of 5000 agents over 5000 periods. Within this population, a_i and e_i assume discrete values in the intervals $[1, 6]$ and $[1, 3]$ respectively. The initial randomization assigns half of the population to each regime $Z = 0$ or $Z = 1$. As in the discussion above, I choose to focus on a situation where the treatment affects only the

offer distribution $G(w;a)$. The treatment effect is negative and is calibrated to equal one standard deviation of the offer distribution, σ_w, with $\beta_s^w = 0$. The full choice of parameters for each policy setting is presented below. To estimate the stationary solution for the accepted offer I simulate expectations of w_{it} over 1000 draws and iterate over the value function until convergence.

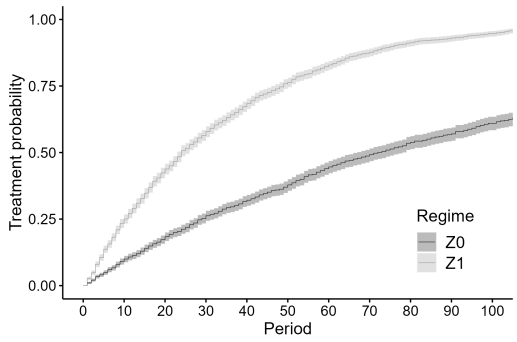
Table A4: Parameter choices in simulations

μ_w	13.762,	μ_c	0.893,	μ_λ	0.092,	ρ	0.995
σ_w	5.497,	σ_c	0.257,	σ_λ	0.031,	σ_ξ	3
β_a^w	4,	β_a^c	0.2,	β_a^λ	0.5/21,	$\pi_{Z=0}$	0.01
β_s^w	5.497,	β_e^c	0.1,	β_e^λ	0.1/21,	$\pi_{Z=1}$	0.03

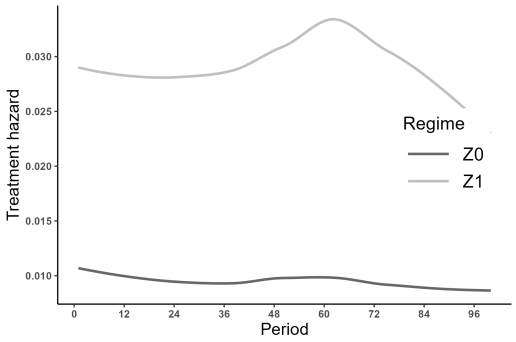
G.3 Simulation Results and Discussion

To apply the continuous time methods under study in this paper, it is preferable to have a large dataset where the unit of time represents a relatively short period. In practice, if the unit of time is too large it may be challenging to account for dynamic selection and for the simultaneity of treatment and exit outcomes within a period. In the estimation, I treat effort e_i as an unobserved characteristic for the researcher, and fully stratify a_i. The researcher observes individual treatment and exit duration outcomes, ability, w_{0it}, an indicator Z for the regime assignment, and an indicator if the observation is right censored. After generating the dynamic discrete choice data I censor all observations greater than $t = 60$ ($\sim$ 43.7%) and apply random right censoring to $\sim$ 6.3% of the remaining observations. Descriptive survival curves and hazards are presented in Figure A7 in the case of a positive treatment effect and in Figure A8 in the case of a negative treatment effect.

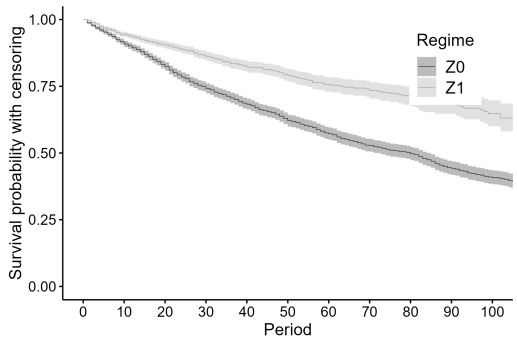
Table A5 shows the results when applying the estimation method. It shows the average effects for β_0, β_z, $\beta_{(0,s]}$, $\beta_{z(0,s]}$ over the treatment times $s = 1,\ldots,30$ when specifying the duration dependence to six 10 period intervals. The estimator performs relatively well with all sample sizes observed. Figure A9 provide further descriptions on the performance of the estimator. They present the estimates of β_0, β_z, β_s, β_{zs} with τ fixed at 60 over the first 30 periods using a sample of 3000 observations from the full population of 5000. The estimator seems to fit the data very well for all causal estimates, and in the case of β_0, β_s, β_{zs}, also match closely the DGP values.



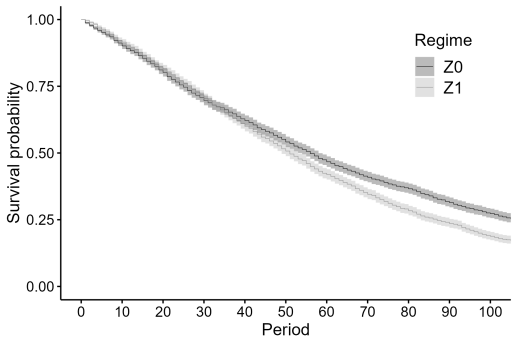
(a) Transplant Probability



(b) Transplant Hazard



(c) Pre-Transplant Survival



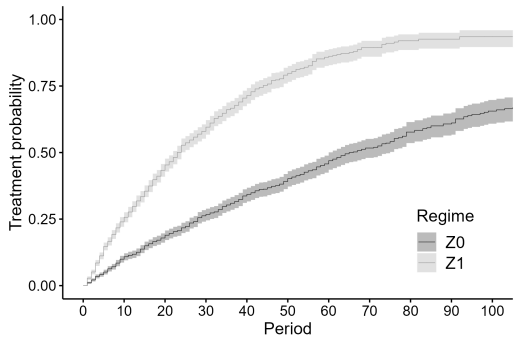
(d) Survival

Figure A7: Results from simulation with positive treatment effects

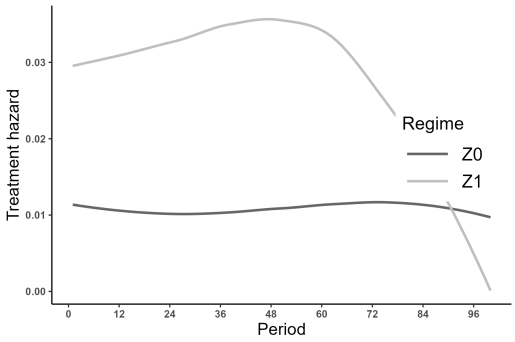
Based on selected sample of Scientific Registry of Transplant Recipients data. Selected sample is described in section 4. $N_{Z=0:no-Tr} = 1013$, $N_{Z=0:Tr} = 453$, $N_{Z=1:no-Tr} = 441$, $N_{Z=1:Tr} = 1093$.

Table A5: Simulation results of dynamic discrete choice model, s from 1,. . . ,30 and $s+\triangle$ fixed at 90

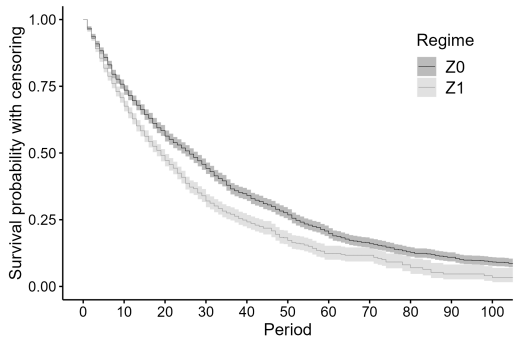
	Estimates	Bias	Variance	MSE
N=5000				
β_0	0.590	0.001	0.006	0.006
β_z	0.170	-0.043	0.000	0.002
$\beta_{(0,s]}$	-0.297	0.040	0.002	0.003
$\beta_{z(0,s]}$	-0.180	-0.021	0.004	0.004
N=3000				
β_0	0.593	0.002	0.006	0.006
β_z	0.175	-0.039	0.000	0.002
$\beta_{(0,s]}$	-0.337	0.000	0.002	0.002
$\beta_{z(0,s]}$	-0.144	0.057	0.005	0.008
N=1000				
β_0	0.618	0.027	0.005	0.006
β_z	0.177	-0.036	0.000	0.002
$\beta_{(0,s]}$	-0.322	0.014	0.003	0.003
$\beta_{z(0,s]}$	-0.211	-0.010	0.004	0.004
N=500				
β_0	0.602	0.011	0.006	0.006
β_z	0.203	-0.011	0.001	0.001
$\beta_{(0,s]}$	-0.324	0.013	0.003	0.003
$\beta_{z(0,s]}$	-0.251	-0.050	0.007	0.010



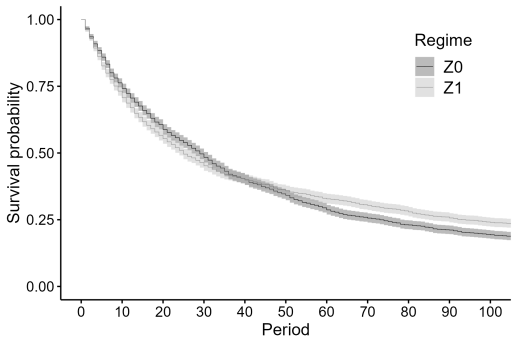
(a) Transplant Probability



(b) Transplant Hazard

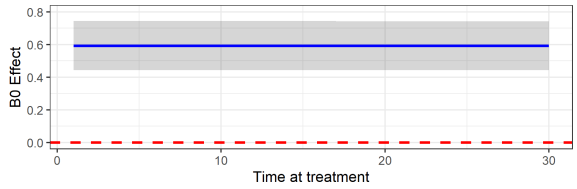


(c) Pre-Transplant Survival

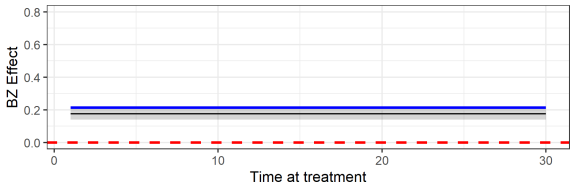


(d) Survival

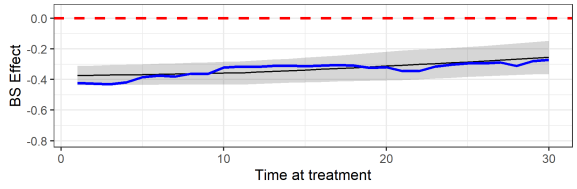
Figure A8: Results from simulation with negative treatment effects
Based on selected sample of Scientific Registry of Transplant Recipients data. Selected sample is described in section 4. $N_{Z=0:no-Tr} = 1013$, $N_{Z=0:Tr} = 453$, $N_{Z=1:no-Tr} = 441$, $N_{Z=1:Tr} = 1093$.



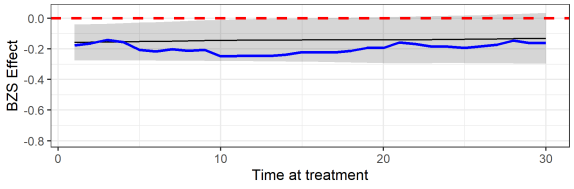
(a) β_0 effect



(b) β_z effect



(c) β_s effect



(d) β_{zs} effect

Figure A9: Assessing performance of empirical evaluation model
Based on selected sample of Scientific Registry of Transplant Recipients data. Selected sample is described in section 4. $N_{Z=0:no-Tr} = 1013$, $N_{Z=0:Tr} = 453$, $N_{Z=1:no-Tr} = 441$, $N_{Z=1:Tr} = 1093$.

G.4 Reservation utilities in dynamic discrete choice model

Consider the post-treatment Bellman equations from the setting described in section G.

$$V_{0,t} = w_0 - c + \rho\lambda\mathbb{E}_w[\max\{V_1(w), (1-\pi)V_{0,t+1} + \pi V^{tr}_{0,t+1}\}] + \rho(1-\lambda)[(1-\pi)V_{0,t+1} + \pi V^{tr}_{0,t+1}]$$
$$V^{tr}_{0,t} = w_0 - c + \rho\lambda\mathbb{E}_{w^{tr}}[\max\{V^{tr}_1(w), V^{tr}_{0,t+1}\} + \rho(1-\lambda)V^{tr}_{0,t+1}$$

This second equation can be written as,

$$V_{0,t}^{tr} = w_0 - c + \rho\lambda \int_{w^{tr*}}^{\infty} \left(V_1(w^{tr}) - V_{0,t+1}^{tr}\right) dG^{tr}(w) + \rho V_{0,t+1}^{tr}$$

Since all parameters and distributions in the model are time independent, we have a stationary reservation utility strategy. In a stationary strategy $V_{0,t}^{tr} = V_{0,t+1}^{tr} = V_0^{tr}$ for all $t > 0$ Furthermore, the reservation utility w^{tr*} is such that the agent would refuse any offer below it and accept any offer above it so $V_1(w) = \frac{w}{1-\rho}$ if $w \geq w^{tr*}$, V_0^{tr} if $w^{tr} < w^{tr*}$, and $V_1(w^{tr*}) = \frac{w^{tr*}}{1-\rho} = V_0^{tr}$ if $w = w^{tr*}$. It follows that,

$$\begin{aligned} V_0^{tr} &= w_0 - c + \rho\lambda \int_{w^{tr*}}^{\infty} \left(V_1(w^{tr}) - V_0^{tr}\right) dG^{tr}(w) + \rho V_0^{tr} \\ &= w_0 - c + \frac{\rho\lambda}{1-\rho} \int_{w^{tr*}}^{\infty} \left(w - w^{tr*}\right) dG^{tr}(w) + \frac{\rho w^{tr*}}{1-\rho} \end{aligned}$$

Replacing again $V_0^{tr} = \frac{w^{tr*}}{1-\rho}$, rearranging this equation and using integration by parts we obtain the post-treatment reservation utility,

$$w^{tr*} = w_0 - c + \frac{\rho\lambda}{1-\rho} \int_{w^{tr*}}^{\infty} \left(1 - G^{tr}(w)\right) dw$$

Note that this reservation utility does not depend on the regime π.

Now consider the reservation utility before treatment with a treatment assignment policy π. Since the problem is still stationary, the agent will again accept any value of w higher than his reservation w^*. We can therefore rewrite
$\mathbb{E}_w[\max\{V_1(w) - (1-\pi)V_{0,t+1} - \pi V_{0,t+1}^{tr}, 0\}] = \int_{w^*}^{\infty} w - w^* dG(w) + (1-\pi)V_{0,t+1} - \pi V_{0,t+1}^{tr}$
which results in the same V_0 value function,

$$V_0 = w_0 - c + \frac{\rho\lambda}{1-\rho} \int_{w^*}^{\infty} w - w^* dG(w) + \rho[(1-\pi)V_0 + \pi V_0^{tr}]$$

Since the agent will accept any value of w higher than his reservation w^* we know that $V_1(w^*) = V_0 = \frac{w^*}{1-\rho}$ which we replace in the above equation and rearrange to get,

$$w^* = \frac{1-\rho}{1-\rho+\rho\pi}(w_0 - c) + \frac{\rho\lambda}{1-\rho+\rho\pi}\left(\int_{w^*(\pi)}^{+\infty} w - w^* dG(w)\right) + \frac{\rho\pi}{1-\rho+\rho\pi} w^{tr*}$$

We can further show prove that w^* is increasing in π if $V_0^{tr} > V_0$:
First we rewrite the previous equation to isolate $\int_{w^*}^{+\infty} w - w^* dG(w)$,

$$\int_{w^*}^{+\infty} w - w^* dG(w) = \frac{1-\rho}{\rho\lambda}\left[c - w_0 + \frac{1-\rho(1-\pi)}{1-\rho} w^* - \frac{\rho\lambda\pi + \rho(1-\lambda)\pi}{1-\rho} w^{tr*}\right]$$

Let us hold w^* constant in the previous equation. w^{tr*} is also constant because it does

not depend on π. If we increase π, the derivative of the right-hand side with respect to π is

$$\frac{1-\rho}{\rho\lambda}[\frac{\rho}{1-\rho}w^* - \frac{\rho}{1-\rho}w^{tr*}] \qquad \leftrightarrow \qquad \frac{1-\rho}{\lambda}[(V_0 - V_0^{tr})]$$

The last equation is negative when $V_0 - V_0^{tr} < 0$, so when the value of treatment is higher than that of no-treatment. Therefore, $\int_{w^*}^{+\infty} w - w^* dG(w)$ is decreasing in π. Furthermore, written as a function of π,with $w^* = w^*(\pi)$, it is also decreasing in w^* which implies that w^* increases in π. The agent is more willing to wait for treatment.

H Proof that accepted kidney quality is increasing with non-sequential kidney offers

The goal is to prove that the expected accepted kidney quality will be increasing in the number of kidneys simultaneously offered to candidates (in consultation with clinicians). We can translate this problem into the following. Say we have an urn with balls labelled $i \in 1, ..., 100$. There are an infinite amount of balls for each label i. We want to prove that the expected value of the maximum labelled ball is increasing in the number of draws.

Let X_n be a random variable representing the maximum labelled ball after n draws. We can express $\Pr(X_n \leq k)$ as the product of the probabilities of each individual draw being less than or equal to k:

$$\Pr(X_n \leq k) = \prod_{i=1}^{n} \Pr(\text{Draw}_i \leq k)$$

Since there are an infinite number of balls for each label, the probability of drawing a ball with label i or less is $i/100$ for each individual draw so $\Pr(\text{Draw}_i \leq k) = \frac{k}{100}$. Substituting this into the expression for $\Pr(X_n \leq k)$, we get

$$\Pr(X_n \leq k) = \left(\frac{k}{100}\right)^n$$

It follows that we can express the probability mass function of X_n as:

$$\Pr(X_n = k) = \Pr(X_n \leq k) - \Pr(X_n \leq k-1) = \left(\frac{k}{100}\right)^n - \left(\frac{k-1}{100}\right)^n$$

The expected value of X_n is then given by:

$$\mathbb{E}[X_n] = \sum_{k=1}^{100} k \cdot P(X_n = k) = \mathbb{E}[X_n] = \sum_{k=1}^{100} k \left[\left(\frac{k}{100}\right)^n - \left(\frac{k-1}{100}\right)^n\right]$$

From which it follows straightforwardly that $\mathbb{E}[X_n]$ is increasing in n since the nth power terms for larger k dominate those for smaller k. As a result, the expected value of the maximum labelled ball is increasing in the number of draws. The result can even more simply be shown in the non-infinite case.

I SRTR risk adjusted score model

The SRTR risk adjustment models are based on Bayesian methods (Bayesian Methods for Assessing Transplant Program Performance). In the case of the 1-year post-transplant survival, the model produces survival predictions using the previous 2.5 years of data and adjusting for a set of donor and candidate characteristics, as well as their interactions (Developing Statistical Models to Assess Transplant Outcomes Using National Registries: The Process in the United States). Given the amount of included variables and the amount of missing variables, some adjustments are necessary. First, some lower threshold of national events are imposed (over 25 events), then a chained equation mean matching algorithm is used to impute values when missing, and last LASSO is used for variable selection. For the offer-acceptance risk adjustment score, any kidney which was never accepted does nto enter the calculation.

Some systematic issues raise concern about the precision of these predictive scores. As remarked in the following official Q&A:How can my program have 0 graft failures, but still be ranked as a tier 4 program?, some transplant centers with 0 graft failures receive scores of 4 out of 5 in their performance evaluation. SRTR's explanation for this seeming anomaly is that it results from transplant centers offering very healthy kidneys to very healthy candidates. Since the expected graft failure is very low for this combination, transplant centers receive reduced scores. While the final outcome may be desirable, reducing incentives for transplant centers to ignore lower health patients, the fact that the predictive statistical model itself, which is purely a function of candidate and donor characteristics, down-ranks transplant centers with perfect 1-year graft records is concerning for the reliability of the model.

J Kidney quality threshold change behavioral hypothesis

J.1 Can kidney quality selection theory explain pre-transplant survival differences?

In this section we turn to an alternative incentive-based explanation. We consider again a forward looking agent model as described in the previous section, but assume reservation survival thresholds are invariant to differences in the kidney offer rate. Instead, the setup assumes only that candidates' kidney quality acceptance thresholds change, and allow this change to be heterogeneous depending on underlying unobserved health. Because there is uncertainty around what size a health shock a candidate will receive each period, agents' (candidate and/or transplant centers) decisions must balance a trade-off. At time t, a candidate would like to select the kidney at time t or later which ensures the longest post-transplant longevity while acknowledging that their probability of succumbing to a negative health shock increases the longer they remain on the waitlist. This optimal stopping problem is the object of study in Agarwal et al. (2021).

In general, facing a higher kidney offer rate, rational candidates will become more selective about the quality of the kidney they are willing to accept. This change can be modelled as a rise in the kidney quality threshold in response to an increase in the kidney offer rate. It will also lead to an average decrease in the acceptance to offer ratio. A priori, however, it is unclear whether the acceptance to offer ratio will decrease more for higher or lower health candidates. Facing a higher offer rate, higher health candidates can reliably wait longer for a high quality kidney than lower health candidates. However, if the marginal returns to a higher quality kidney are larger for lower health candidates, meaning they will in expectation benefit more from an equal time wait than higher health candidates, then lower health candidates may have a larger decrease in their acceptance to offer ratio.

For the empirical patterns in Figure 1 of the introduction to arise in the forward looking agent model described in this section, the second stated dynamic must be more prevalent. Facing a higher offer rate, lower health candidates become relatively more selective about which kidneys to accept compared to higher health candidates. This explanation can operate alongside the lifestyle adjustment hypothesis but can also explain the pre-transplant results in the absence of any moral hazard.

J.2 Can kidney quality selection theory explain survival differences of early transplanted?

At first glance, the results in Figure 4a appear to corroborate with the kidney quality selection hypothesis. More specifically, if the relative acceptance to offer ratio is lower

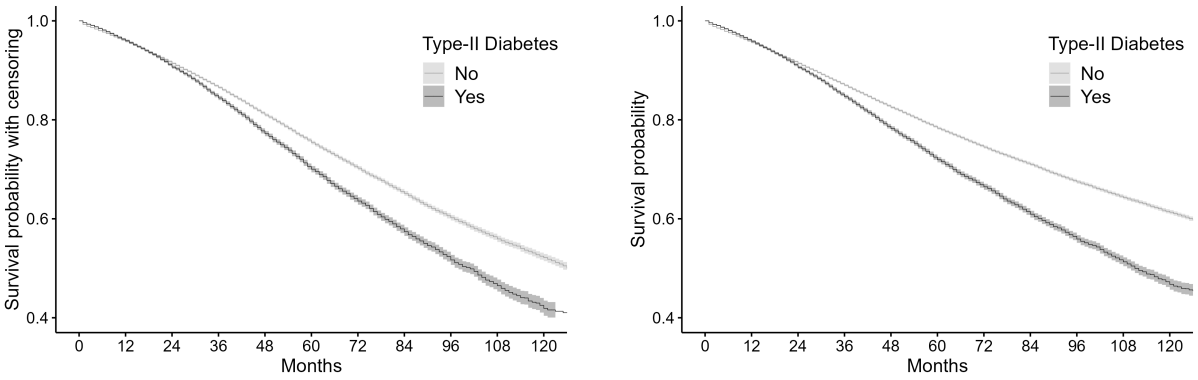
for poor health candidates in the AB blood group than in the O blood group, then the distribution of health among candidates accepting a transplant in the first months on the waitlist should be more skewed towards healthier candidates in the AB-blood type group. Relatively more healthy candidates receiving a transplant in the AB-blood regime should in turn lead to higher survival among those transplanted early in the AB-blood type regime.

Pushing the argument a step further, though, we find some complicated implications of this hypothesis. First, this conclusion rests on one of two assumptions concerning information acquisition. If we assume candidates are the main decision makers of which kidney to accept, then we must assume candidates immediately integrate the offer rate into their decision making process, which seems questionable. In addition, even if we assume clinicians in transplant centers are the main decision makers, the decision does not appear to be welfare improving from the point of view of candidates. This is noticeable in Figure 4b, where arguments for selective acceptance of transplants do not translate into notable differences in overall survival.

Beyond these intangible arguments, data driven arguments can also be seen in Appendix B Figure A1 columns 5-6. These present descriptive statistics for the O and AB blood type candidates receiving a transplant in the first three months on the waitlist. Based on the age and diabetes covariates, selection on observed variables operates in the opposite manner to that suggested in the previous section's kidney quality selection hypothesis. AB blood candidates accepting a transplant in the first three months on the waitlist are, on average, older and more likely to have diabetes upon entering the waitlist.

K Heterogeneous effects using Type-II diabetes as a proxy for poor health

In this appendix I use a binary indicator for type-II diabetes as a proxy for low health. As shown in Figure A10a-A10b, candidates with Type-II diabetes have a lower pre-treatment survival and a lower overall survival than other candidates on the waitlist.



(a) Pre-Transplant Survival

(b) Survival

Figure A10: Survival of candidates on the kidney transplant waitlist depending on Type-II diabetes status

Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B. Observations: $N_{t-IIDiab} = 104,490$, $N_{not-IIDiab} = 39,805$.

K.1 Heterogenous effects on survival

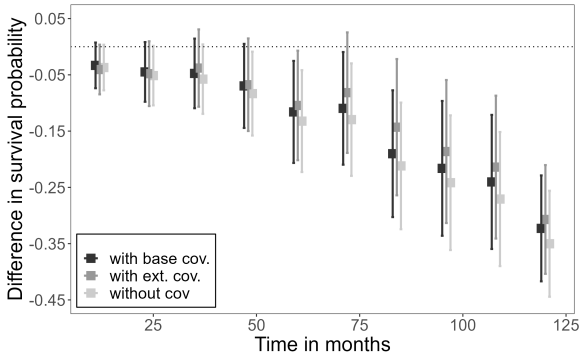
Given these different survival rates, I estimate the heterogenous effects model of section 5.4 including a binary variable $\mathrm{W} = 1$ for candidates with type-II diabetes (low health), and $\mathrm{W} = 0$ otherwise (high health),

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{W}_i + \beta_2 \mathrm{Z}_i \cdot (1 - \mathrm{W}_i) + \beta_3 \mathrm{Z}_i \cdot \mathrm{W}_i + \beta \mathrm{X} + u_i \tag{6}$$

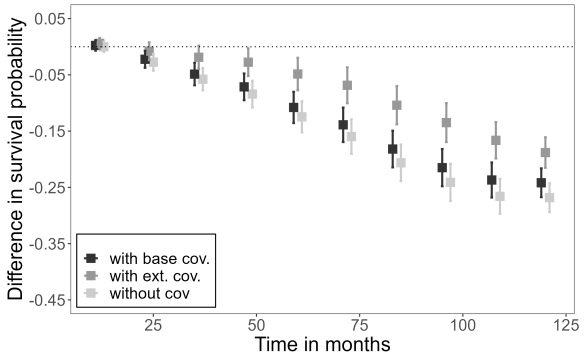
The parameters of interest in this model are $\beta_1 = \mathbb{E}[\mathrm{Y}_i | low, AB] - \mathbb{E}[\mathrm{Y}_i | high, AB]$, $\beta_2 = \mathbb{E}[\mathrm{Y}_i | high, O] - \mathbb{E}[\mathrm{Y}_i | high, AB]$, and $\beta_3 = \mathbb{E}[\mathrm{Y}_i | low, O] - \mathbb{E}[\mathrm{Y}_i | low, AB]$. Estimates for these effects are presented over time in Figures . As in the main text, Figures 4a, conditions the analysis on candidates receiving a transplant within the first three months on the waitlist. Figure 4b present differences in overall survival between O and AB blood types.

K.2 Testing differential probability of transplant at $t = 1$

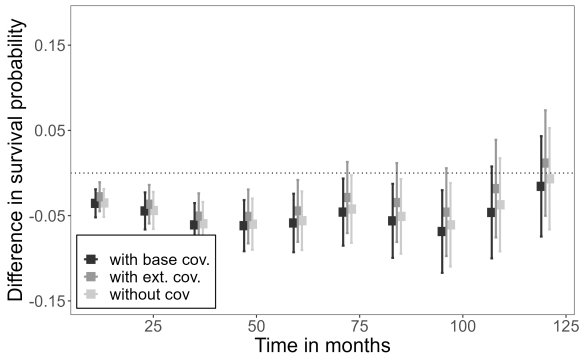
We can further use an adjusted version of the heterogenous model from the previous subsection to test the institutional incentive and kidney quality change hypotheses of selection. To do so, the evaluation replaces the outcome Y_i with a transplant indicator



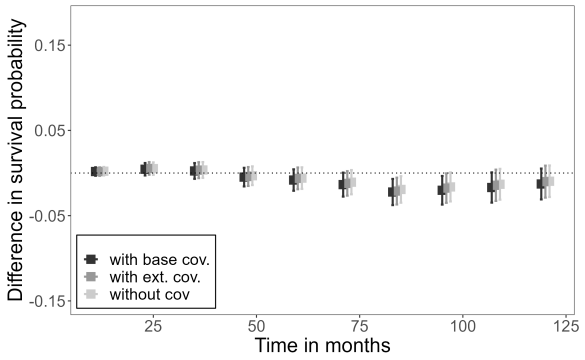
(a) Survival for early transplanted, β_1



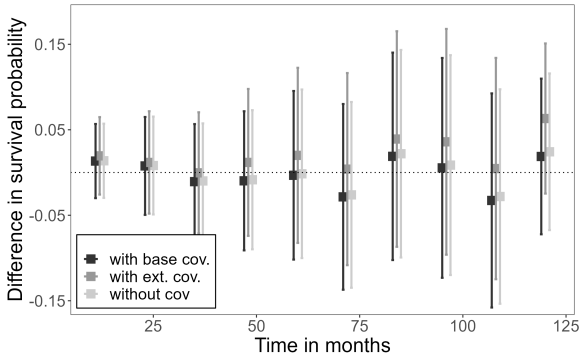
(b) Overall Survival, β_1



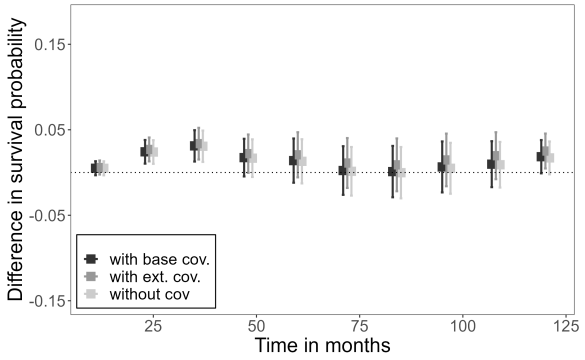
(c) Survival for early transplanted, β_2



(d) Overall Survival, β_2



(e) Survival for early transplanted, β_3



(f) Overall Survival, β_3

Figure A11: Difference in survival between AB and O blood types over time

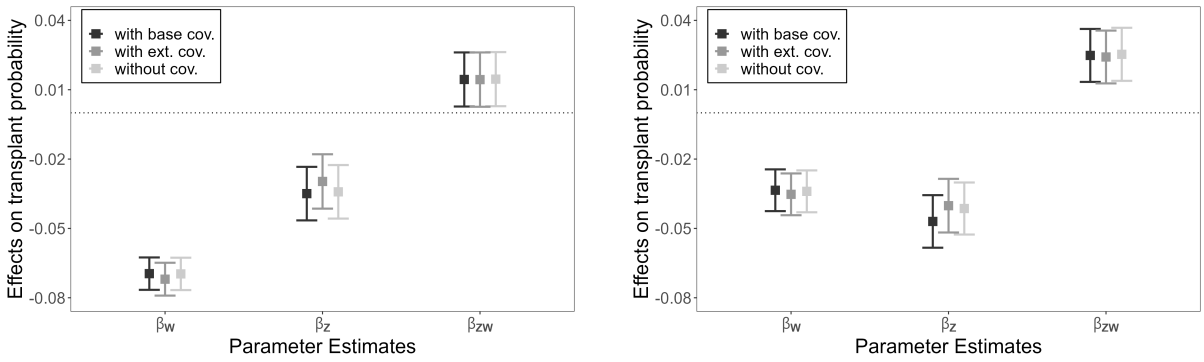
Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data described in Appendix B.
(*without cov*) includes no covariates, (*with base cov*) includes offer covariates: age bracket dummies, (*with ext. cov*) includes all variables presented in Appendix Table A1. Observations (a,c,e): $N_O = 3757$, $N_{AB} = 955$. Observations (b,d,f): $N_O = 125324$, $N_{AB} = 9697$.

D_i, where $\mathrm{D}_i = 1$ if a candidate received a transplant in the first three months on the waitlist, and $\mathrm{D}_i = 0$ otherwise. More specifically, the model estimates the probability of receiving a transplant in the first months on the waitlist using the following model,

$$\mathrm{D}_i = \beta_0 + \beta_Z \, \mathrm{Z}_i + \beta_W \, \mathrm{W}_i + \beta_{ZW} \, \mathrm{Z}_i \cdot \mathrm{W}_i + \beta_X \, \mathrm{X} + u_i \tag{7}$$

β_0 represents the probability of transplant in the first three months on the waitlist for AB-type candidates without type-II diabetes. β_Z represents the difference in transplant

probability for O-types without type-II diabetes. β_W is the difference in transplant probability for AB-types with type-II diabetes relative to AB-types without type-II diabetes. β_{ZW} measures any additional difference in the transplant probability for O-types with type-II diabetes relative O-types without type-II diabetes.



(a) Old kidney allocation system

(b) New kidney allocation system

Figure A12: Heterogenous effects of blood type and type-II diabetes on probability of transplant within first 3 months on the waitlist

Includes candidates entering waitlist after 01 June 2002, durations truncated at 01 December 2014. Based on selected sample of Scientific Registry of Transplant Recipients data. Selected sample is described in section 4.
(*without cov*) includes no covariates, (*with base cov*) includes offer covariates: age bracket dummies pre-2015, and dummies for all EPTS variables post-2015. (*with ext. cov*) includes all variables presented in Appendix Table A1. Observations (a): $N_{O:no-Tr} = 93,922$, $N_{O:Tr} = 35,989$, $N_{B:no-Tr} = 28,432$, $N_{B:Tr} = 10,303$. Observations (b): $N_{O:no-Tr} = 93,922$, $N_{O:Tr} = 35,989$, $N_{A:no-Tr} = 50,005$, $N_{A:Tr} = 31,637$.

The hypothesis in the kidney quality change hypothesis predicts that β_Z should be negative since, given the lower kidney offer rate, the probability of transplant should also be lower for O-blood types. It also predicts that β_W should be negative if unhealthier candidates with type-II diabetes are more selective about which kidney to choose and therefore delay their decision to accept a kidney. And, finally, the behavioral hypothesis put forward also predicts that β_{ZW} will be positive if, given the lower offer rate in their regime, unhealthy O-type candidates with type-II diabetes are relatively less selective about which kidney to accept than unhealthy AB-type candidates with type-II diabetes. All of these predictions are supported by the results presented in Figure A12.

We can, however, also show with simple numerical examples that the results in Figure A12 can arise solely under the institutional selection hypothesis assumptions of section 5.2. For the scenario in which clinicians are more likely to offer a kidney to higher health candidates than lower health candidates,these results replicate the patterns estimated in Figure A12 without any further assumptions. In the second scenario in which clinicians do not offer more kidneys to higher health candidates but lower health candidates are more likely to reject offered kidneys, an additional distributional assumption is necessary. To replicate coefficient patterns in Figure A12, the distribution of offered kidneys in terms of quality must be right skewed, meaning there are more undesirable than desirable kidneys. Evidence for this phenomenon is well known and outlined in (Agarwal et al., 2018, 2020).

Expanding on the above, we wish to show that the same coefficient signs for β_Z, β_W and β_{ZW} in Figures A12 can arise for the first scenario in which clinicians are more

likely to offer a kidney to higher health candidates than to lower health candidates. The consequence of this skewed distribution can be illustrated with the kidney acceptance numbers in Table A3 for the relevant case in which the probability of accepting a kidney is increasing in unobserved health (top frame). β_Z, the difference in the probability of accepting a kidney between O-blood and AB-blood type candidates with high health (no type-II diabetes) is given by $1h/5h - 3h/5h = -2/5$. β_W, the difference in the probability of accepting a kidney between low health (type-II diabetes) and high health (no type-II diabetes) candidates in the AB-blood type regime is also for the stylised numbers given by $1h/5h - 3h/5h = -2/5$. β_{ZW}, the additional difference between low health (type-II diabetes) and high health (no type-II diabetes) candidates in the AB-blood type regime is given by $(0l/5l - 1l/5l) - (1h/5h - 3h/5h) = 1/5$. These results replicate the patterns estimated in Figure A12.

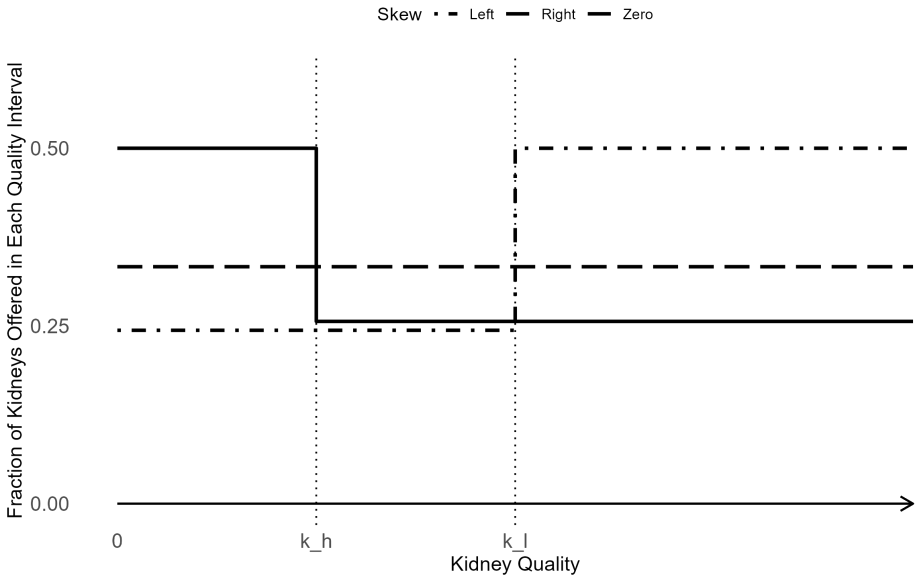


Figure A13: Distribution of offered kidney scenarios

We can also show that the same coefficient signs for β_Z, β_W and β_{ZW} can arise for the second scenario in which clinicians do not offer more kidneys to higher health candidates but lower health candidates are more likely to reject offered kidneys.

Consider the case in which two kidneys are offered in the O-blood regime and 4 are offered in the AB-blood regime. Assume as well for simplicity that clinicians randomly allocate these kidneys regardless of candidate health. As such, higher and lower health candidates can each expect to receive half the kidneys offered. Higher health candidates will further accept any kidney above the threshold k_h while lower health candidates are more selective only accepting kidneys above some threshold k_l. These thresholds can be thought as normalised with respect to the quality of kidneys offered. So, even if in absolute terms both higher and lower health candidates may have the same rejection threshold, the differing thresholds can reflect the lower average quality kidneys offered to

lower health candidates. Also assume that when multiple kidneys are offered, they are offered simultaneously.

Take first the case of a right skewed distribution as described in Figure A13. In expectation, O-type higher health candidates will accept $\frac{1}{2}$ offers while AB-type higher health candidates will accept $3 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$ offers. Similarly, O-type lower health candidates will accept $\frac{1}{4}$ offers while AB-type lower health candidates will accept $2 \cdot \frac{1}{4} \cdot \frac{3}{4} + \frac{1}{4} \cdot \frac{1}{4} = \frac{1.75}{4}$ offers. As a result, we would obtain $\beta_Z = \frac{2}{4} - \frac{3}{4} = -\frac{1}{4}$, $\beta_W = \frac{1.75}{4} - \frac{3}{4} = -\frac{1.25}{4}$ and $\beta_{ZW} = (\frac{1}{4} - \frac{2}{4}) - (\frac{1.75}{4} - \frac{3}{4}) = \frac{0.25}{4}$. As we can see, the coefficient patterns follow those in Figure A13. Similar calculations will show that in the zero-skew case effects are 0 and in the left-skewed case the coefficient patterns are inversed.

All in all, the effects on the probability of transplant in the first three months outlined above support both the institutional incentive and kidney quality change hypotheses of selection. The results also caution against drawing firm conclusions from static treatment models alone when it comes to separating theories of selection.