Mechanisms for Online Organ Matching

Nicholas Mattei
IBM T.J. Watson Research Center
n.mattei@ibm.com

Abdallah Saffidine
UNSW Sydney
abdallah.saffidine@gmail.com

Toby Walsh
UNSW Sydney, Data61, TU Berlin
toby.walsh@data61.csiro.au

Abstract

Matching donations from deceased patients to patients on the waiting list account for over 85% of all kidney transplants performed in Australia. We propose a simple mechanisms to perform this matching and compare this new mechanism with the more complex algorithm currently under consideration by the Organ and Tissue Authority in Australia. We perform a number of experiments using real world data provided by the Organ and Tissue Authority of Australia. We find that our simple mechanism is more efficient and fairer in practice compared to the other mechanism currently under consideration.

1 Introduction

Kidney disease costs the Australian economy billions of dollars every year in medical costs, welfare payments, and lost income. On 1st February 2016, there were 1083 people on the waiting list for a kidney transplant. Over the course of 2016, 1448 people received a kidney from a deceased donor, whilst a further 265 people received a kidney from a living donor. Just 44 of these came from paired exchanges. Paired exchange represents less than 3% of all transplants. Deceased donors provide the majority of transplanted organs, and will do so till we have xenotransplants or can grow organs.

Matching deceased donors to people on the waiting list is becoming increasingly difficult as road safety improves. In 1989, the mean age of deceased donors was 32 years. In 2014, this had increased to 46 years. Medicine has also advanced so that older organs can now be successfully transplanted. The oldest donor in 1989 was 69. In 2014, it was 80. We have therefore started to work with the Organ and Tissue Authority, the statutory body in Australia that allocates organs to develop a new mechanism that explicitly takes account of the age of the organs and of the recipients for the first time. The mechanism currently in use will offer an old organ to a young patient. A lot of time is then wasted as transplant surgeons reject offered organs till the older organ is accepted for a lower ranked and hence older patient.

It has been decided that the quality of the organ is to be measured by the Kidney Donor Patient Index (KDPI). This is an integer from 0 to 100 that is calculated from the age of the donor, their diabetic status, cause of death and other factors. A donor with a KDPI of X% has an expected risk of graft failure greater than X% of all kidney donors. The quality of the recipient is to be measured by the Expected Post-Transplant Survival (EPTS) score. This is also an integer from 0 to 100 that is calculated from the age of the recipient, their diabetic status, the number of prior organ transplants and their time on dialysis. A recipient with a lower EPTS is expected to have more years of graft function from high-longevity kidneys compared to candidates with higher EPTS scores. The goal of our collaboration with the Organ and Tissue Authority is to design a new mechanism that is fair and efficient, matching organs so that their KDPI is as close as possible to the EPTS score of their allocated patient.

This work fits into a broader research programme to develop models and mechanisms for resource allocation problems that reflect the richness and complexity of the real world [Walsh, 2015; Aziz *et al.*, 2016]. One of the fundamental features of the organ matching problem is that it is online. We must match organs as they arrive, before we know what organs or patients will arrive in the future. At the end of the year, we could find the best allocation in polynomial time by computing a maximum weight matching over the weighted bipartite graph. However, we cannot wait till the end of the year; we must allocate and transplant organs as they arrive. There are many other domains where we allocate resources in a similar online manner. A food bank might start allocating and distributing food to charities as soon as it is donated. There may be neither the time nor warehouse space to store the food. An observatory might start allocating time on an expensive telescope before all requests have come in. This work offers a case study in how we can efficiently and fairly solve *online* allocation problems [Albers, 2003]. Insights from this research may prove valuable in many domains.

2 Mechanisms

We propose a very simple mechanism that minizes the difference between KDPI and EPTS.

MIN: Allocate an arriving organ to a compatible patient that minimizes $|KDPI - EPTS|$, tie-breaking by time on the waiting list and then randomly.

We will compare this with the mechanism proposed by the Organ and Tissue authority before we started the collaboration, but not yet implemented.

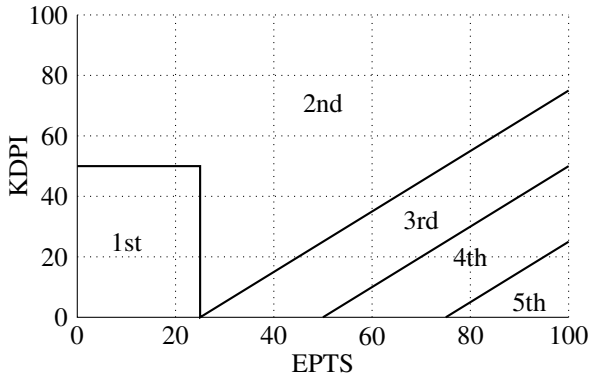


Figure 1: The BOX mechanism: the ordering is induced by a lexicographical scoring function with five key regions.

BOX**:** This mechanism ranks patients according to a lexicographical scoring function. The most important terms in the scoring function ensure a match between compatible types. The least important terms tie-break according to features like time on the waiting list. The middle term orders matches as: (1) $KDPI \leq 50$ and $EPTS \leq 25$; (2) $EPTS - 25 < KDPI$; (3) $EPTS - 25 \leq KDPI < EPTS - 50$; (4) $EPTS - 50 \leq KDPI < EPTS - 75$; (5) $EPTS - 75 \leq KDPI$.

The ordering can be viewed graphically (Figure 1). It favours matches in a rectangular box to the bottom left, then above the upper left diagonal, and then towards the lower right diagonal.

3 Online Organ Matching

Online organ matching occurs in a series of time steps. At each time step, either a set of organs arrives, or a set of patients arrive or depart. Organs are matched as soon as they arrive, whilst arriving patients are simply added to the waiting list. In practice, organs tend to arrive in pairs as each deceased donor typically donates both their kidneys. Patients depart the waiting list when they are matched to an arriving organ, receive a transplant elsewhere, become sick or die. In Australia, the size of the waiting list remains roughly constant over time. In practice, many more patients are on dialysis but are not put on the waiting list due to health and age issues.

We employ a simple model of organ matching where each arriving organ has an associated blood type, KDPI and state while each patient has an associated blood type, EPTS and state. The EPTS of a patient slowly increases over time as they spend longer on dialysis. We only permit matches of a compatible type (e.g. organs coming from a donor of blood type O can be matched to recipients of any blood type, whilst organs coming from a donor of blood type AB can be matched only to recipients of blood type AB). We focus on compatibility by blood type, but in the real world this consists of additional factors (e.g. HLA type, for a more nuanced discussion see the data and simulators from `www.srtr.org/`). We will consider mechanisms that match only within a state, as well as mechanisms that match nationwide. A secondary goal of our work is to measure the benefits of matching across state boundaries, currently under consideration by the Organ and Tissue Authority.

We define the quality of a matching of an organ to a patient in two ways. At the level of the individual, the *utility* of a match is simply $100 - KDPI$. A patient wants the youngest possible organ. At the level of society, the *welfare* of a match is $100 - |KDPI - EPTS|$. To maximize the benefit to society of a limited supply of organs, we want the $KDPI$ to be as close as possible to $EPTS$. As we consider randomized mechanisms, we can compute such measures as utility and welfare both ex post and ex ante.

4 Desirable Axioms for Allocation

An axiomatic study of algorithms is a corner stone of game theory [Maschler *et al.*, 2013] and computational social choice [Brandt *et al.*, 2016]. For instance, both the MIN and BOX mechanisms are **anonymous**: identical patients added to the waiting list at the same time are treated identically. We say a mechanism satisfies **participation** if a patient cannot increase their expected utility by joining the waiting list at a later date. Unfortunately, the online nature of organ matching means that no mechanism can satisfy participation.

Theorem 1 *No mechanism for online organ matching satisfies participation.*

Proof. Suppose two patients join the waiting list at the first time step, an organ with a KDPI of 50 arrives at the second time step, and one with a KDPI of 0 at the third. Both patients have an incentive to delay their participation and wait till after the first organ is allocated before joining the waiting list. □

A fundamental and desirable property of an allocation of organs is efficiency. In this setting all agents have the same utility value over the organs. If you lower the KDPI of the organ matched to one patient, you must inevitably increase the KDPI of an organ matched to some other patient. Therefore all allocations are efficient in terms of utility. We consider also efficiency in terms of welfare. An online allocation is **welfare efficient** if and only if there is no other online allocation in which one patient has greater welfare and none of the other patients have lower welfare. A mechanism is welfare efficient if and only if it only returns online allocations that are welfare efficient. Note that we limit our discussion to online allocations. Organs can only be matched to patients that are actually present on the waiting list as we cannot match to a patient yet to join the waiting list.

Example 1 *Suppose we have one patient on the waiting list with an EPTS of 100, and an organ with a KDPI of 0 arrives and is matched to this patient. Suppose a new patient now arrives with an EPTS of 0, followed by an organ with a KDPI of 100, that is again matched with the only patient on the waiting list. This is the only possible online allocation so it is welfare efficient. But in an offline setting, this allocation is not welfare efficient. We could match the first patient with the second organ, and the second patient with the first organ, increasing the welfare of both patients.*

Another desirable property of online mechanisms is that they do not lead to **starvation**. In online settings this is typically handled by giving higher priority to agents the longer

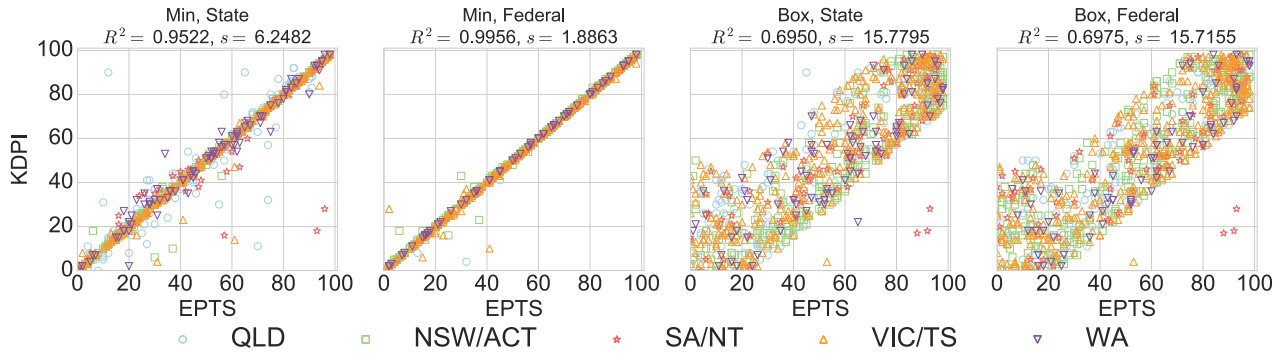


Figure 2: Visualization of the results of one run of our algorithm for both the MIN and BOX algorithms with both the State and Federal restrictions. MIN performs significantly better than BOX in terms of minimizing $|KDPI - EPTS|$ and the Federal exchange outperforms the State exchanges in ensuring better outcomes for all patients.

they wait. In all the proposed mechanisms, wait time is only used for tie-breaking between agents. It is possible in the worst case for patients never to be matched. However, we do not observe this in practice with the historical distribution of patients and organs. Both mechanisms under consideration may never match patients on the list.

5 Experiments

From an axiomatic perspective, online organ matching appears to be a challenging problem. In this section, we run some experiments using historical data to determine whether such properties are nevertheless achieved in practice. The Scientific Registry of Transplant Recipients publishes detailed models of expected donations and transplant survival times for the US Market, see `www.srtr.org/`. In what follows we generate patients and kidneys using some of the same techniques but tuned for the Australian market. We do not go into as much detail on post transplant survival as we are more concerned with the properties of the initial allocation. Subjecting this allocation to optimization under the expectation of future survival times is an interesting avenue for future work with some work beginning in the kidney exchange literature already [Dickerson and Sandholm, 2015].

Data Generation. Since we have only one historical data set, we first need to build a simulator that is as realistic as possible so we can run thousands of experiments with realistic data. We use two different sources of data to ensure that our model is based on real world assumptions. Our data for generation and testing comes from two sources, both provided by the Organ and Tissue Authority of Australia.

Public ANZDATA: Long run statistics published by `www.anzdata.org.au`. This aggregates details about kidney donation in Australia and New Zealand.

Research ANZDATA: Detailed data from 2010–2014, enabling us to create more fidelity for our random simulators. This data includes the waiting list at a single point in the year and information on all kidneys donated in 2010-2014. The distributions and quality of patients and kidneys in this dataset is designed to meet the random pattern observed over 10 years of historical data.

Cross checking data from the Australian Bureau of Statistics; the Public and Research sets; and Wikipedia, we determined the probability distribution over the *Blood Type* and *State* of donors and patients. We generate both patients and donors according to these statistical distributions. We generate the patients for the waiting list using the population proportions provided by both the historical data and national statistics; These distributions match to within 1%.

In order to simulate the arrival of kidneys over the course of a year we use a Poisson process to simulate the arrival times of patients to donate kidneys over the course of a year. Death is typically modeled as an independent random variable, e.g., the arrival of bodies to morgues is a standard case study in modeling random processes with Poisson processes. We use the long run average from both data sources of 340 donors a year. This gives us $\mu = 340/365$ and defines a probability distribution we can sample from for the number of patients arriving in a day, $P(x) = e^{-\mu} \cdot \mu^x / x!$. Each donor that arrives can donate one or two kidneys. The long run average from our datasets is 1.72 kidneys transplanted per deceased donor. Therefore, each patient that arrives donates one kidney with probability 0.28 and two with probability 0.72.

The waiting list in Australia remains about the same length over time with the historical average ≈ 1200 people. Everyone who receives a transplant is removed from the list and new individuals are added. On average 150 patients are removed from the waiting list randomly throughout the year for a number of reasons including death and off list transplant. We model removal from the list again as Poisson process with $\mu = 150/365$; we uniformly at random draw a patient from the set of patients to remove when necessary. In addition, the EPTS of patients degrades over the course of the year with about 180 people being removed from the waiting list every year because their EPTS has become too high. To simulate this, we "age" each patient by incrementing their EPTS every 30 days they are on the list by increasing it between 0 – 1 point, uniformly at random.

We generate EPTS and KDPI according to the distributions provided to us for the years 2010–2014. We generate a number between 0–100 in bands of 10 using the historic distributions. We then add 0–9 to this number uniformly at random.

We do this both for EPTS and KDPI. Our simulator, along with one for the US kidney exchange market, is available at `www.preflib.org` [Mattei and Walsh, 2013].

Experimental Treatments and Evaluation. For our experiments we generate an initial list of patients and kidneys and then simulate the arrivals, departures, and donations by stepping through a simulated 8 years. We repeat this process 1000 times to gain confidence in the statistics we report here [Cohen, 1995]. We use the first four years of data to burn in our simulator, so that the exchange has reached a steady state, and report statistics based on the latter four years of data. The same list of kidneys and patients (their order of arrival) are used for all of the treatments (both mechanisms and both state and federal exchanges). Hence, the difference in statistics are generated only by the particular matching strategy.

For all experiments we enforce basic blood type compatibility. That is, A can donate to A and AB, B to B and AB, AB to AB, and O can donate to anyone. We also differentiate between two treatments regarding the level the exchange takes place at. In a **Federal** exchange, a kidney can be transplanted to any state in Australia. We enforce that in the case of a tie, the kidney goes to the instate patient. In a **State** exchange all exchanges are run within states only. Organs not able to be transplanted in state due to type restrictions are then matched to any patient outside the state where the kidney originated.

Ideally we would transplant every organ into a patient such that KDPI and EPTS are equivalent. This gives us a notion of optimality that is defined by the line $y = x$. We use two error metrics to judge how well our points match this idealized line. Note that since we have regression through the origin here we must be careful with computing our statistics [Cohen, 1995]. Firstly, the **Coefficient of Determination (R^2)**, if we treat KDPI as a dependent variable w.r.t. ETPS then we can use R^2 to get an indication of the proportion of the variance of KDPI that is predictable from EPTS. If we want to ensure that KDPI and EPTS are perfectly correlated ($x = y$) then we would get an $R^2 = 1$. We compute: $R^2 = 1 - \frac{\sum_i (EPTS_i - KDPI_i)^2}{\sum_i (KDPI_i - \overline{KDPI})^2}$. Secondly, we use the **Standard Error of the Regression (s)**, since s is expressed in the same units as the independent variable (KDPI) it gives us an intuitive measure of how how much KDPI varies, on average, from EPTS. Since we are forcing our regression through the origin we only have one degree of freedom in the model giving $s = \sqrt{\frac{1}{n-1} \sum_i (KDPI_i - EPTS_i)^2}$.

6 Results and Discussion

Figure 2 shows the results for one run which are typical of all the runs. Here we can see that, no matter the state, the patient results are better for the MIN algorithm and Federal matching. MIN outperforms BOX by a statistically significant margin for both the R^2 measure and s for all 1000 instances that we tested. For the MIN algorithm, Federal significantly outperforms State, across all 1000 instances as well.

One concern with the Federal v. State treatments on the practical level is the flow of organs out of a state will exceed the flow of organs into a state. In effects, states do not want to lose organs to other states. When tracking organ flow we use the term inflow to mean organs that come into a particular state or type and outflow to mean organs that move to a different state or type. A flow between organ types means that, for instance, an O kidney was donated to an AB patient. Figure 3a illustrates the flow of organs between states and we can see that the majority of organs are exchanged within the state with the exception of Western Australia. The cost in terms of efficiency to the overall system, illustrated in Figure 2, is a much greater concern than organ flows.

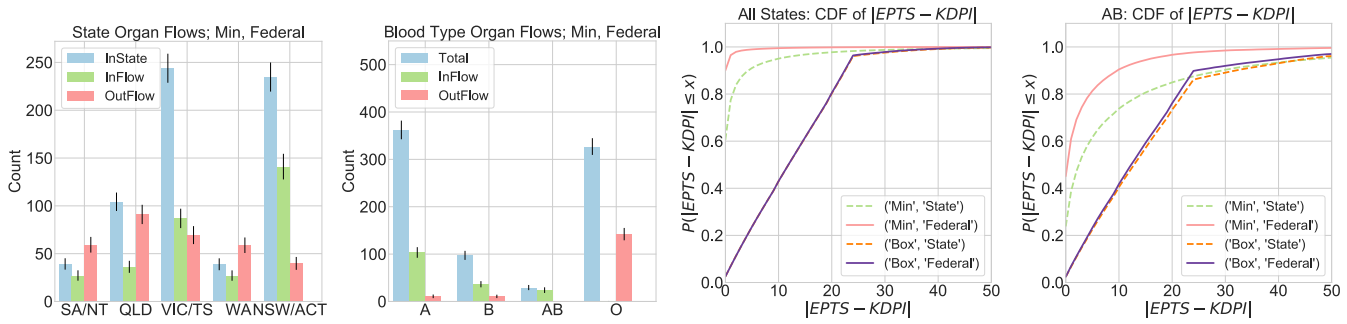
Time on List. One of the main concerns for patients is the amount of time that they expect to spend on the waiting list. To investigate this we computed the mean waiting time in days by EPTS, the results for all states and for NSW/ACT are broken out in Figure 4a. We have omitted variance/error bars for this graph as they completely overlap. In general, the variance is constant for each of the mechanisms and treatments with the MIN mechanism having a lower variance of about ± 200 days while the BOX mechanism has a higher variance of about ± 600 days. We observe that the variance for the BOX algorithm is strictly higher than the variance for the MIN algorithm across all treatments.

Looking at Figure 4a we see that the MIN mechanism has a lower mean waiting time versus BOX except for patients with very high EPTS. This gap closes for patients with higher EPTS, likely due to the fewer organs that are donated with very high EPTS. It is interesting to note that the State v. Federal question is roughly negligible for mean time on list. Consequently, Federal exchanges increase patient welfare without a significant impact on waiting time. The BOX has large equivalence classes between patients, illustrated in Figure 1. We conjectured that these large equivalence classes would give lower wait times as the mechanisms tie-break based on the time on the waiting list. However, this is not the behavior that we see in the data except for those high EPTS patients.

In Figure 4b we slice the mean waiting time data along the blood type axis. This reveals another interesting property of online organ matching: different blood types get treated very differently. Those patients with type O blood, the most common amongst the types, track very closely to the waiting times see in Figure 4a while those with type AB blood have strictly lower waiting times. This is due to the small number of AB patients that require transplants and their ability to accept kidneys of any type.

Efficiency. To investigate the efficiency of the various mechanisms in practice we investigate the number of Pareto swaps [Brandt *et al.*, 2016], i.e., between kidney/patient pairs such that the welfare of one of the pairs is increased while the other is not decreased. To do this we took, for each of the 1000 iterations, the complete set of transplants and checked to see for each kidney/patient if there exists another kidney/patient that arrives later in time that would enable a Pareto swap. Since one kidney/patient pair may be involved in multiple potential swaps, and a welfare efficient matching would only use at most one of these swaps *per patient*, we also compute the number of unique kidney/patient pairs that could be part of a Pareto improving swap. We also track the average increase in welfare over the set of potential swaps per instance. These values are reported in Table 1 over the 1000 trials.

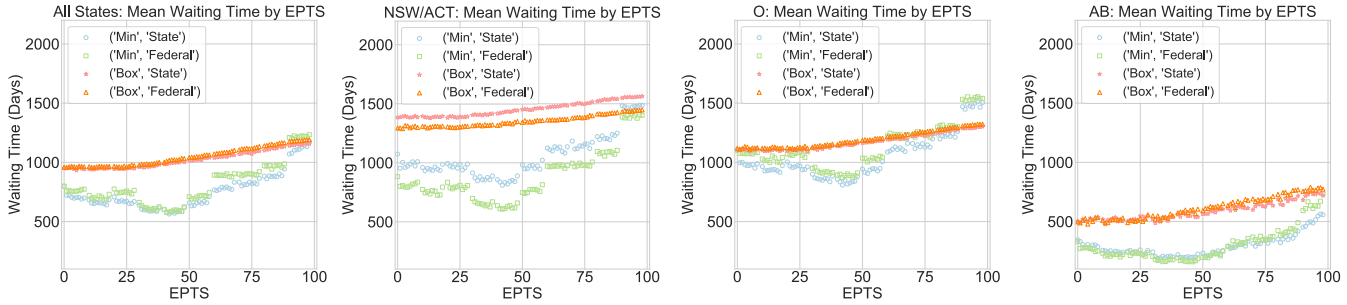
The metric we compute is a worst case one. We do not take into account the aging factor for the later swap. Hence, the



(a) The majority of organs are transplanted within the state of donation.

(b) With the MIN, Federal treatment, 99% of agents receive an organ where $|KDPI - EPTS| < 5$. Agents with a higher $|KDPI - EPTS|$ are almost all waiting for a rare AB organ.

Figure 3: In (a) we show the average over 1000 iterations of the flow of organs between states and blood types; error bars represent one standard deviation. In (b) we show the Cumulative distribution function (CDF) of $|KDPI - EPTS|$. The vertical axis shows the percentage of patients who have a $|KDPI - EPTS|$ below the value on the horizontal axis.



(a) Mean time on list for all states (left) and for NSW/ACT (right).

(b) Mean time on list for type O (left) and type AB (right).

Figure 4: In (a) we see that MIN gives lower average waiting times and that the difference between State and Federal exchanges is negligible. In (b) Blood type O is the most common and thus the waiting times track closely to the overall mean while the rarest blood type AB does not.

		Mean Number of Swaps		Mean Number of Patients		Mean Gain in Welfare	
MIN	Federal	3.90	(3.4)	3.59	(3.0)	0.1	(0.01)
	State	314.12	(53.9)	181.4	(22.6)	3.81	(0.85)
BOX	Federal	24,548.6	(2,233.1)	937.83	(30.87)	14.18	(0.34)
	State	23,555.6	(2,173.6)	934.62	(29.79)	14.18	(0.32)

Table 1: Mean number of Pareto improving swaps per patient, mean number of patients who could participate in a swap, and mean gain in welfare (standard deviation) for the four treatments. The MIN, Federal algorithm achieves the greatest efficiency and lowest frequency of opportunities for Pareto improving trades amongst the patients.

number of swaps reported is an upper bound and would be lower, in reality, if we aged the patient participating in the swap. Additionally, kidneys need to be matched on the same day they are donated so these swaps are not even realistically possible. In the experiments for MIN, there were no improving swaps possible on the same day (though there were some for BOX). We report improving swaps that could have happened on *any day* to give an upper bound on the worst case that could happen if we could store the kidneys.

The lower the numbers across the board in Table 1 for the MIN algorithm shows that it is doing much better in terms of welfare efficiency than BOX. It is encouraging to see that the Federal exchanges result in significantly fewer patients who would benefit from a Pareto swap. Additional evidence for the quality of the MIN algorithm can be seen in the CDF of $|KDPI - EPTS|$ over all 1000 runs seen in Figure 3b; the clear winner is MIN with the Federal treatment. There is a vanishingly small probability of observing a value of $|KDPI - EPTS|$ larger than 10. Turning to the right side of 3b, we see that, in fact, most of the efficiency loss is due to AB organ transplants. As there are so few AB organs, it is hard to match them in way that maximizes welfare.

Participation. We examined if patients can strategically delay entrance into the market to improve their outcome. To do this we took, for each of the 1000 iterations, the complete set of transplants made during that iteration and checked if a patient could receive a kidney for which they had higher utility by declaring a later arrival time. We took each patient p and checked if, in a future time step, a compatible kidney arrives of better quality for p. If so, we check if p (1) could have arrived before p' who received the better kidney and (2) p would have had priority greater than or equal to that of p' where priority is determined by $|KDPI - EPTS|$ for the MIN algorithm and by placement in the same equivalence class for

		Mean Num. Patients Delay Improves	Mean Num. of Improvements/Patient	Mean Max Improvement in KDPI
MIN	Federal	138.2 (20.5)	2.7 (0.34)	3.0 (1.25)
	State	269.3 (24.1)	2.5 (0.31)	9.2 (2.45)
BOX	Federal	960.5 (29.9)	169.5 (7.1)	50.9 (0.93)
	State	928.8 (31.6)	40.6 (2.1)	49.6 (1.00)

Table 2: The mean number of patients who can delay and improve their received organ, the mean number of improvements per patient, and the mean of the maximum improvement a patient could see for all 1000 iterations (standard deviation) for the four algorithmic treatments. The MIN, Federal algorithm often achieves allocations where KDPI=EPTS which affords few opportunities for strategic participation.

the BOX algorithm. If so than we say p could have delayed for an improvement in utility. This is a worst case measure, we assume that p has full knowledge of all the kidneys that arrive in the future and p wins tie breakers against p'.

Our results for participation are shown in Table 2, standard deviations over 1000 runs in parenthesis. The MIN algorithm is much less susceptible to manipulation from arriving later. This is due to the large equivalence classes that are crated by the BOX mechanism, which offer more opportunities for optimistic manipulation. Additionally, the magnitude of the improvement in KDPI of the transplanted kidney is enormous for the BOX mechanism, indicating that arriving later can significantly increase patient utility. Interestingly, we see an increase in the number of patients that could manipulate in the MIN, State treatment. This is likely due to the smaller pools of patients and donors which leads to an increased disparity between EPTS and KDPI of patient and donor, illustrated in Figure 2, leading to an increase in the delay opportunity.

7 Related Work

Online problems have been studied in computer science for decades, the primary application area is online scheduling [Albers, 2003]. Mechanism design concerns have been extensively addressed in the online scheduling literature [Porter, 2004] though online allocation has received less attention.

Online matching markets without money have been a common area of study in computer science. Gujar and Parkes [2010] study an online matching market where there is no money, much like organ allocation, though only one side of their market is dynamic. Another online matching market was studied by Bosek *et al.* [2014]. In their model the market incrementally increases in size and at each increase, a new maximum stable matching must be found. In our work we are not concerned with stability but fairness. Additionally, in online organ matching, unlike general online matching, all agents have a shared preference model and *both* sides of the market are dynamic. Finally, Aleksandrov *et al.* [2015] consider the online allocation problem faced by foodbanks. In the foodbank problem, each charity is allocated multiple indivisible goods, not just one in organ matching. Additionally, each charity can have arbitrary preferences, in organ matching these preferences are shared. Finally, in the foodbank problem, the charities being allocated items are fixed while in organ matching both sides are dynamic.

Kidney *exchanges* have been extensively studied in the economics [Roth *et al.*, 2005; 2004], medical [Montgomery *et al.*, 2006], and computer science literatures [Dickerson *et al.*, 2012]. Axiomatic and fairness concerns have received attention in the computer science literature in both theory and practice [Dickerson *et al.*, 2014]. Kidney exchanges have also been studied in online settings [Awasthi and Sandholm, 2009] and in predictive settings using machine learning techniques [Dickerson and Sandholm, 2015]. Additional work in computer science has also focused on strategyproof mechanisms for kidney exchange at the patient and hospital level [Ashlagi *et al.*, 2015] and on merging kidney exchanges with other organs such as lungs [Luo and Tang, 2015] and livers [Dickerson and Sandholm, 2016]. However, as we have pointed out, the majority of kidney donations are performed from deceased patients and the online version of the deceased donor has historically received little attention.

A related line of work is determining what factors should be included in the scores such as KDPI and EPTS. Bertsimas *et al.* [2013] study the problem of using data driven methods for finding national organ allocation policies. In general, they adopt a statistical approach that designs policies that work well on average based on historical data. By comparison, we have taken a more axiomatic approach. The Organ and Tissue Authority in Australia wish to use a new mechanism based solely on blood/tissue type and KDPI/EPTS. This prevents an approach like Bertsimas *et al.* [2013] where we compute weights for terms going into KDPI/EPTS.

8 Conclusions

We have proposed a new mechanism, MIN for the online matching of deceased organs to donors. We have compared this with the current mechanism, BOX under consideration by the Organ and Tissue authority of Australia. By running experiments on historic data, we find that the MIN algorithm outperforms the proposed BOX mechanism. There are several directions for future work. We would, for example, like to consider axiomatic properties like strategy-proofness, fairness with respect to blood types, and states as well as issues like egalitarian and utilitarian welfare. Additionally our empirical evaluation of post-transplant success could incorporate more complex features like those found in the SRTR.

Acknowledgements

Data61 is supported by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program. Abdallah Saffidine is the recipient of an ARC DECRA Fellowship (DE 150101351). Toby Walsh is supported by the European Research Council and by AOARD Grant FA2386-12-1-4056.

References

[Albers, 2003] S. Albers. Online algorithms: A survey. *Mathematical Programming*, 97(1-2):3–26, 2003.

[Aleksandrov *et al.*, 2015] M. Aleksandrov, H. Aziz, S. Gaspers, and T. Walsh. Online fair division: Analysing a food bank problem. In *Proc. 24th IJCAI*, 2015.

[Ashlagi *et al.*, 2015] I. Ashlagi, F. Fischer, I. A. Kash, and A. D. Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 91:284–296, 2015.

[Awasthi and Sandholm, 2009] P. Awasthi and T. Sandholm. Online stochastic optimization in the large: Application to kidney exchange. In *Proc. 21st IJCAI*, 2009.

[Aziz *et al.*, 2016] H. Aziz, C. Cahan, C. Gretton, P. Kilby, N. Mattei, and T. Walsh. A study of proxies for shapley allocations of transport costs. *JAIR*, 56:573–611, 2016.

[Bertsimas *et al.*, 2013] D. Bertsimas, V. F. Farias, and N. Trichakis. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, 61(1):73–87, 2013.

[Bosek *et al.*, 2014] B. Bosek, D. Leniowski, P. Sankowski, and A. Zych. Online bipartite matching in offline time. In *Proc. 55th FOCS Symposium*, pages 384–393, 2014.

[Brandt *et al.*, 2016] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[Cohen, 1995] P. R. Cohen. *Empirical Methods for Artificial Intelligence*. MIT Press, 1995.

[Dickerson and Sandholm, 2015] J. P. Dickerson and T. Sandholm. Futurematch: Combining human value judgments and machine learning to match in dynamic environments. In *Proc. 29th AAAI*, pages 622–628, 2015.

[Dickerson and Sandholm, 2016] J. P. Dickerson and T. Sandholm. Multi-organ exchange: The whole is greater than the sum of its parts. *JAIR*, 2016.

[Dickerson *et al.*, 2012] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Optimizing kidney exchange with transplant chains: Theory and reality. In *Proc. 11th AAMAS*, pages 711–718, 2012.

[Dickerson *et al.*, 2014] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Price of fairness in kidney exchange. In *Proc. 13th AAMAS*, pages 1013–1020, 2014.

[Gujar and Parkes, 2010] S. Gujar and D. Parkes. Dynamic matching with a fall-back option. In *Proc. 19th ECAI*, pages 263–268, 2010.

[Luo and Tang, 2015] S. Luo and P. Tang. Mechanism design and implementation for lung exchange. In *Proc. 24th IJCAI*, pages 209–215, 2015.

[Maschler *et al.*, 2013] M. Maschler, E. Solan, and S. Zamir. *Game Theory*. Cambridge University Press, 2013.

[Mattei and Walsh, 2013] N. Mattei and T. Walsh. Preflib: A library for preferences, HTTP://WWW.PREFLIB.ORG. In *Proc. 3rd ADT*, 2013.

[Montgomery *et al.*, 2006] R. A. Montgomery, S. E. Gentry, W. H. Marks, D. S. Warren, J. Hiller, J. Houp, A. A. Zachary, J. K. Melancon, W. R. Maley, H. Rabb, C. Simpkins, and D. L. Segev. Domino paired kidney donation: a strategy to make best use of live non-directed donation. *The Lancet*, 368(9533):419, 2006.

[Porter, 2004] R. Porter. Mechanism design for online real-time scheduling. In *Proc. 5th ACM-EC*, pages 61–70, 2004.

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

[Roth *et al.*, 2005] A. E. Roth, T. Sönmez, and M. U. Ünver. Pairwise kidney exchange. *Journal of Economic theory*, 125(2):151–188, 2005.

[Walsh, 2015] T. Walsh. Challenges in resource and cost allocation. In *Proc. 29th AAAI*, pages 4073–4077, 2015.