NBER WORKING PAPER SERIES

TRIAGE PROTOCOL DESIGN FOR VENTILATOR RATIONING IN A PANDEMIC: INTEGRATING MULTIPLE ETHICAL VALUES THROUGH RESERVES

Parag A. Pathak
Tayfun Sönmez
Utku Unver
M. Bumin Yenmez

Working Paper 26951
http://www.nber.org/papers/w26951

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2020

We are grateful for input from several medical professionals. Nikhil Agarwal and Alex Rees-Jones provided helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Triage Protocol Design for Ventilator Rationing in a Pandemic: Integrating Multiple Ethical Values through Reserves
Parag A. Pathak, Tayfun Sönmez, Utku Unver, and M. Bumin Yenmez
NBER Working Paper No. 26951
April 2020
JEL No. D45,D47,I3

ABSTRACT

In the wake of the Covid-19 pandemic, the rationing of medical resources has become a critical issue. Nearly all existing triage protocols are based on a priority point system, in which an explicit formula specifies the order in which the total supply of a particular resource, such as a ventilator, is to be rationed for eligible patients. A priority point system generates the same priority ranking to ration all the units. Triage protocols in some states (e.g. Michigan) prioritize frontline health workers giving heavier weight to the ethical principle of instrumental value. Others (e.g. New York) do not, reasoning that if medical workers obtain high enough priority, there is a risk that they obtain all units and none remain for the general community. This debate is particularly pressing given substantial Covid-19 related health risks for frontline medical workers. In this paper, we analyze the consequences of rationing medical resources through a reserve system. In a reserve system, ventilators are placed into multiple categories. Priorities guiding allocation of units can reflect different ethical values between these categories. For example, a reserve category for essential personnel can emphasize the reciprocity and instrumental value, and another reserve category for general community can give higher weight to the values of utility and distributive justice. A reserve system provides additional flexibility over a priority point system because it does not dictate a single priority order for the allocation of all units. It offers a middle-ground approach that balances competing objectives. However, this flexibility requires careful attention to implementation, most notably the processing order of reserve categories, given that transparency is essential for triage protocol design. In this paper, we describe our mathematical model of a reserve system, characterize its potential outcomes, and examine distributional implications of particular reserve systems. We also discuss several practical considerations with triage protocol design.

Parag A. Pathak
Department of Economics, E52-426
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
ppathak@mit.edu

Tayfun Sönmez
Boston College
Department of Economics
140 Comm. Ave.
Chestnut Hill, MA, 02467
tayfun.sonmez@bc.edu

Utku Unver
Boston College
Department of Economics
140 Comm. Ave. Chestnut
Hill, MA, 02467
unver@bc.edu

M. Bumin Yenmez
Boston College
Department of Economics
140 Commonwealth Ave.
Chestnut Hill, MA 02467
bumin.yenmez@bc.edu

"If you fail to plan, you are planning to fail."

— Benjamin Franklin

1 Introduction

The Covid-19 pandemic has inspired renewed attention to how medical resources are rationed. As of the writing of this paper, intensive care unit (ICU) beds and ventilators are among the two most important scarce resources. In the United States, as of March 2020, there were between 68,000-85,000 ICU beds for the adult population. The total number of full-featured ventilators ranged between 72,000-82,000 including those in the Strategic National Stockpile, and there were roughly 100,000 additional partial-feature ventilators (Emmanuel et al., 2020). Operating ventilators also requires respiratory therapists and other staff, and a large fraction of existing ventilator capacity is already deployed during non-emergency times.[1] The U.S. Centers for Disease Control estimates that upwards of 21 million Americans will require hospitalization for Covid-19, and evidence from Italy suggests that 10-25% of the American population will require ventilators (Truog, Mitchell and Daley, 2020). While these forecasts evolve daily, Emmanuel et al. (2020) argue it is unclear how quickly supply can meet demand even with recent efforts to increase production of these critical resources.

How to implement a fair rationing system during a pandemic presents a complicated question rife with ethical concerns. Medical ethicists have established several principles for rationing scarce medical resources in the case of influenza. Many states have proposed frameworks or explicit rules during severe pandemics too.[2] These principles echo formal normative criteria in resource allocation problems in economics. For example, one principle is to treat people equally regardless of their socioeconomic status, age, or race. A second principle is to rely on the expected health outcomes measured through age, preexisting health conditions, or some notion of quality-adjusted life years. A third, sometimes conflicting, principle is to help the worst off or the sickest. Yet another principle considers a patient's instrumental value, which protects frontline health workers. Whatever principle is adopted, there is widespread consensus that any rationing protocol should be transparent, applied consistently, and not be left to the discretion of frontline health workers, who may be under severe pressure (Truog, Mitchell and Daley, 2020).

Contested debates on these ethical principles have motivated several in the medical community to propose a **priority point system**. In a priority point system, an explicit formula specifies the order in which resources should be rationed. Although our analysis applies to a variety of medical resources, we use ventilators to simplify terminology.

The 2015 New York State Task Force on Ventilator Allocation is an example of a priority point system. It has influenced the protocols of several states. The system recommends that

[1] The New York State Ventilator Guidelines in 2015 state that during non-emergency normal conditions, 85% of ventilators are used, leaving only 15% available for new cases (Zucker et al., 2015, p. 27).

[2] Guidelines include Maryland (Daughtery-Biddison et al., 2017), Michigan (Michigan, 2012) Minnesota (Vawter et al., 2010), New York (Zucker et al., 2015), North Carolina (Tong and Devlin, 2007), Ontario (Christian et al., 2006), South Carolina (Ball and Schneider, 2009), the Swiss Influenza Pandemic Plan (Iten and Strupler, 2018), and Utah (DoH, 2018).

after certain patients are excluded, there should be a priority order for patients who experience respiratory failure based on estimated mortality risk and such patient's priority status should be re-evaluated every 48 hours based on changes in health status (Zucker et al., 2015). In the Task Force's proposed priority system, there are four categories: Red, Yellow, Blue, or Green. Red patients have the highest priority because they are most likely to recover with the use of ventilators. Yellow patients are next, defined as those who are very sick and have intermediate or uncertain likelihood of survival with ventilators. Next are Blue category patients who have low likelihood of survival with ventilators and are excluded if there is not enough supply, but are allocated if there are ventilators remaining. Green patients are the lowest priority because they are in no need of ventilator support, and therefore do not receive one even if there is excess supply. At any priority level, there is the potential that one priority group could completely exhaust the remaining available resources. In cases of excess demand upon remaining resources by a given priority group, New York and other proposals recommended random allocation – a lottery – among equal-priority patients (Zucker et al., 2015; Emmanuel et al., 2020).

While a priority system does provide a well-defined protocol, we argue that it does not offer the flexibility to accommodate competing objectives and values. This need is widely acknowledged by task forces commissioned to develop these systems, and it has led several groups to propose a *multi-principle* approach, see, e.g., Daughtery-Biddison et al. (2017). For example, a Johns Hopkins study examining public perceptions of rationing principles such as equal treatment, allocation to those who stand the gain the most, and instrumental value summarized (Biddison et al., 2013):

> Both groups felt strongly that no single principle could adequately balance the competing aims and values triggered by allocation decisions. Some felt that a combination of principles should be used. Both groups suggested alternative strategies, such as organ transplantation allocation criteria as a model or adopting a tiered approach by applying different principles at different stages in process.

In this paper, we examine properties of a **reserve system** for rationing medical resources in pandemics. In a reserve system, patients are identified as members of particular groups – e.g., young or old, frontline health worker or not, very sick or sick. Group membership can overlap: a young patient could also be a frontline health worker. In a reserve system, reserve sizes (i.e., the quantity of resources reserved) are set for each group, and within a reserve group, priority can be based on an explicit score or random assignment.[3] A reserve system provides additional flexibility over a priority point system because priority points need not apply to the allocation of all ventilators. For example, a fraction is reserved for frontline health workers, while the rest is unreserved for all community members. Frontline health workers are always first in line in their reserve category, and the remaining priorities can be the same. This system balances their interests against other ethical goals.

Reserve systems are widespread in resource allocation settings outside of medicine when there are conflicting objectives.[4] The key idea of a reserve system is to divide the total supply into

[3] As we explain below, reserves can either be soft, in which no units are left unassigned, or hard.

[4] Examples include school assignment (Dur et al., 2018; Dur, Pathak and Sönmez, 2019), affirmative action (Sönmez and Yenmez, 2019*a*,*b*), and immigration visa allocation (Pathak, Rees-Jones and Sönmez, 2020*a*).

several categories, and consider allocation for these smaller number of units separately. Specific objectives can be realized within these categories, using explicit priorities or randomization. The advantage of a reserve system is in its flexibility. A priority point system obscures trade-offs between different principles because it aggregates several different considerations into a single priority score. It is even possible that one principle might dominate other principles unexpectedly depending on how scores are scaled.

It is important to note that we are agnostic on what the reserve types or sizes should be. Instead, our aim is to inform the debate on how a reserve system can be used to balance competing objectives, and provide a route forward in several high-stakes debates on rationing.

We first illustrate the power of a reserve system by relating it to the debate on whether frontline health workers should obtain priority for vital medical resources. Under Michigan's guidelines, essential personnel are prioritized for these resources (Michigan, 2012). Recently, Emmanuel et al. (2020) also advocated for prioritizing frontline health workers offering the following rationale:

> ...priority for ventilators recognizes their assumption of the high-risk work of saving others, and it may also discourage absenteeism.

On the other hand, the Minnesota Pandemic Working group argued against prioritizing frontline health workers, suggesting they should be treated the same way as all other patients (Vawter et al., 2010). One reason they offer is

> ... it is possible that they [key workers] would use most, if not all, of the short supply of ventilators; other groups systematically would be deprived access.

The New York State Task Force also struggled with this topic, recognizing the need to provide "insurance" for frontline health workers. However, they ultimately decided that access should only depend on a patient's clinical factors and not their occupation. There are presently deep disagreements on these issues as several states, medical associations, and hospitals develop their own guidelines (Fink, 2020). A reserve system provides a resolution to this conundrum, because only a subset of ventilators are reserved for frontline health workers and the rest are unreserved and open for any group that is eligible for the resources. Prioritization within these two categories can be based on a lottery or explicit criteria like sequential organ failure assessments (SOFA) or comorbidities. A reserve system would accommodate both concerns.

A second debate about priority systems involves whether a certain group can be excluded from treatment as opposed to merely being placed in the lowest priority group. In Alabama's proposed plan, individuals with mental disabilities or dementia were to be considered "unlikely candidates for ventilator support" during rationing. Washington state guidelines recommended that hospital patients with "loss of reserves in energy, physical ability, cognition and general health" be switched to outpatient or palliative care (Fink, 2020). The Office for Civil Rights in the U.S. Department of Health and Human Services is currently investigating the legality of these rules, and authorities are in the process of re-evaluating their systems. If certain groups are prohibited from receiving ventilators, it violates a fundamental principle of non-discrimination. That is, it violates the idea that every patient, no matter his or her circumstances, should have some hope of obtaining a ventilator. In a reserve system, if a portion of ventilators are set aside for all, even if that allotment is small, and random allocation is used for that segment, every patient could possibly obtain a ventilator. In contrast, in a priority system coupled with excess demand for available resources by the higher-priority groups – even without any explicit exclusion of certain types of individuals – there will be some patients who would never be treated during a shortage.

An important medical precedent for reserve systems comes from deceased donor kidney allocation where the number of potential recipients far exceeds the number of available kidneys. Until 2014, the U.S. Organ Procurement and Transplantation Network (OPTN) used a priority system to allocate deceased donor kidneys (OPTN, 2014). In this system, after establishing medical compatibility, patients were ordered by priority type, and then priority points, with ties broken by waiting time. Agarwal et al. (2019) describe that in the NYC-area system, priority was given first to patients with a perfect tissue-type match, then to patients from the local procurement organization, and finally to patients in close geographic proximity. In 2014, the OPTN system changed to include a reserve (Israni et al., 2014). In the reserve, 20% of the highest quality kidneys –defined as those with the best donors based on the Kidney Donor Profile Index (KDPI)– are reserved for adults with the highest 20% expected post-transplant survival score (EPTS). A priority points system is used for the remaining kidneys. The OPTN offered a utilitarian justification for the reserve:

> The EPTS score is designed to achieve better longevity matching. The candidates with the top 20% EPTS scores will receive offers for kidneys from donors with KDPI scores of 20% or less before other candidates at the local, regional, and national levels of distribution.[5]

While the advantage of a reserve system is in its flexibility, it requires careful implementation. In some cases, implementation is straightforward. For kidneys, specific kidneys are reserved based on their quality through a formula, and there is a natural order of awarding these kidneys based on their arrival time. That is, whenever a high-quality kidney that qualifies for the 20% reserve arrives, priority is given to patients in the preferential treatment group for these types of kidneys. However, for rationing objects like ventilators or ICU beds, which are typically seen as homogenous, the design of reserve systems is more challenging. In particular, in a system that reserves a fraction of units for various groups, the processing order of reserve categories is important. In fact, there are now several instances of reserve systems with systematic implementation errors because of these complexities.[6]

[5] The New Kidney Allocation System (KAS) Frequently Asked Questions, Organ Procurement and Transplantation Network. Available at `https://optn.transplant.hrsa.gov/media/1235/kas_faqs.pdf`.

[6] The US Customs and Immigration has a reserve system for the H1-B immigration visas, and since 2005, it has deployed four different allocation mechanisms, sometimes unintentionally (Pathak, Rees-Jones and Sönmez, 2020*a*). Boston Public Schools had a reserve system in its school assignment plan in which half of each school's seats were reserved for students from the neighborhood. In 2013, after realizing that how reserves are processed dramatically influences their impact, the city abandoned the reserve after significant public debate (Pathak and Sönmez, 2013*b*; Dur et al., 2018). India's public universities and jobs use reserve systems for affirmative action. Since there are many traits eligible for reserves, faulty implementation of reserves has generated countless litigation and confusion (Sönmez and Yenmez, 2019*a*,*b*; Pathak and Sönmez, 2019). Finally, Pathak, Rees-Jones and Sönmez

Since a reserve system can be designed in several ways and the previous examples show that their properties can be counter-intuitive, in this paper, we characterize the entire class of reserve system policies satisfying three normative principles, which are defined precisely below. A major insight of our analysis is that when implementing a reserve system, processing a reserve category earlier is detrimental to its beneficiaries.

Section 2 of the paper discusses several aspects of reserve design motivated by medical rationing. Readers not interested in these mathematical details can skip the formal model in Section 3. The paper concludes in Section 4.

2 Design Considerations for Triage Protocol

2.1 Priority Point System vs. Reserve System

In a typical triage protocol, all units are allocated with a single priority order, often determined with a point system. The Sequential Organ Failure Assessment (SOFA) score features prominently in many systems. A consensus conference developed the SOFA score in 1994 to objectively measure organ dysfunction or failure over time in individual patients and in groups of patients with sepsis (Vincent et al., 1996). SOFA is part of the South Carolina Pandemic Influenza Ethics Task Force's (Ball and Schneider, 2009, Appendix 5) recommendation to use the White et al. (2009) priority point system. Their point system is given in Table 1.

Table 1. Illustration of Multi-Principle Strategy for Ventilator Allocation

Principle	Specification	Point System			
		1	2	3	4
Save the most lives	Prognosis for short term survival (SOFA score)	SOFA < 6	SOFA 6-9	SOFA 10-12	SOFA > 12
Save the most years of life	Prognosis for long term survival (Medical assessment of comorbidities)	No comorbidities	Minor comorbidities	Major comorbidities	Severe comorbidities; likely death within 1 year
Life-cycle principle	Prioritize those who have had the least chance to life through life's stages	Age 12-40	Age 41-60	Age 61-74	Age > 74

Notes: SOFA stands for sequential organ failure assessment. Multi-principle point system developed by White et al. (2009)

This system is based on integrating multiple ethical principles: saving the most lives, saving the most years of life, and the life-cycle principle. Within each priority group, points are assigned based on a formula. To understand how it works, consider a hypothetical patient with a SOFA score of seven. She would be awarded two points. If she patient has no comorbidities and is between 61-74 years old, she obtains four more points, yielding a total of six. A patient with a

(2020*b*) shows that a large number of participants in an incentivized experiment have incorrect beliefs about some properties of reserve systems.

lower total point score has a higher priority for a ventilator than a patient with a higher total point score.[7]

A priority points system tries to accommodate different ethical goals using a single score. However, there are some principles that are challenging or even impossible to represent in this way. The challenge arises from the need to norm or scale different and potentially-unrelated ethical principles into one dimension. These challenges are not unlike the usual challenges associated with any aggregation of social alternatives into a single ordering involving multiple inputs.

Moreover, there are some ethical principles that do not have cardinal representations that can be used as priority points. One notable example is group-based policies, like those in Europe for public office and company boards. The European Union has proposed balanced participation of women and men in political and public decision making by requiring that at least 40% of public offices are held by women and at least 40% are held by men (Dittmar, 2018; Rankin, 2020). There are similar policies for publicly traded countries in several countries (Lee, 2014). Such rules cannot be implemented in a priority point system. For medical rationing, it is not far-fetched that there is a future pandemic so devastating that it threatens a large enough portion of the human race. In such a case, it may be necessary to have a similar symmetric female and male reservation category, which cannot be done with priority points.

The problems associated with a single score are particularly visible in public debates on whether essential workers such as frontline medical personnel should be prioritized in ventilator allocation. There is a strong case for their claim based on the principles of reciprocity and instrumental values. While most state task forces consider this possibility for their pandemic ventilator allocation guidelines, states like New York and Minnesota ultimately did not embrace it due to concerns about the extreme scenarios where no units may remain for general community members (Zucker et al., 2015; Vawter et al., 2010). In contrast, in states like Michigan priority is given to essential workers. And, indeed, one of the main recommendations in Emmanuel et al. (2020) is prioritization of frontline health workers for allocation of these vital resources.

Several task force reports illustrate that there is no consensus on this particular question. As a result, two polar opposite policies have been adopted. A reserve system offers a solution to this dilemma: prioritize essential workers for ventilator allocation only for a subset of the ventilators.

2.2 Reserve System Parameters

The main parameters of a reserve system are a division of all units into multiple segments, referred to as *reserve categories*, and specification of a *priority order* of the patients in each of these categories. For some (or all) of the reserve categories, there can also be *exclusion criteria*, based on the nature of the medical resource that is being rationed along with the clinical assessment of the patient. The priority order of patients for each category incorporates this information. The reserve categories can differ based on the groups to receive higher priority and the ethical principle invoked.

[7]The South Carolina protocol does not specify a tie-breaker between two patients with the same total points.

The main idea is to use the associated priority order – which embeds ethical principles – when allocating units in each reserve category. Importantly, the priority order need not be the same between reserve categories. In some applications,

1. units are heterogeneous and specific units are reserved for specific reserve categories,
2. they are perishable and have to be allocated as soon as they arrive, and
3. they sequentially arrive given a natural process.

In these applications the allocation process is uniquely determined by these three conditions: Allocate each unit as soon as it arrives to the highest priority agent in its reserve category. The allocation of deceased donor kidneys in the U.S. is one of these applications.

In many applications, however, one or more of these conditions fail, and there are several mechanisms to implement the reserve system. Our application of interest, allocation of vital medical resources during a pandemic, is one of these applications. While these resources may become available sequentially, they are often homogenous and non-perishable. A reserve system can be implemented in those applications through several mechanisms. The outcome of these mechanisms can differ considerably depending on the choice of the mechanism.

The parameters of a reserve system can be modified for different medical resources. Emmanuel et al. (2020) emphasizes that "prioritization guidelines should differ by intervention and should respond to changing scientific evidence." In particular, vaccines and tests may rely on different ethical principles than ventilators and ICU beds. Clearly, decisions about reserve sizes and reserve categories are up to the medical community and other stakeholders. However, we next describe some potentially important reserve categories.

2.3 Potential Reserve Categories

Any reserve category has to be based on well-established ethical principles. One of these principles is *reciprocity* towards persons who accept risk for the common good of saving lives. Reciprocity is also closely related to the principle of *instrumental value*, which gives priority to those who can save others, or gives priority to those who have saved others in the past (Emmanuel et al., 2020; White and Lo, 2020). These are referred to prospective and retrospective instrumental values. We consider two such reserve categories, which we refer to as the *instrumental value category* and the *good samaritan category*.

2.3.1 Instrumental Value Category

The instrumental value category provides some form of priority to essential personnel such as frontline health workers during the pandemic. One of the primary reasons that several states do not prioritize essential personnel for ventilator allocations is because of the following concerns, articulated in 2015 New York State Ventilator Allocation Guidelines (Zucker et al., 2015):

> Expanding the category of privilege to include all the workers listed above may mean that only health care workers obtain access to ventilators in certain communities. This

> approach may leave no ventilators for community members, including children; this alternative was unacceptable to the Task Force.

Limiting priority allocation of ventilators to essential personnel for only a subset of ventilators is a natural compromise, compared to the two extreme policies that either provide it for all units (eg. Michigan) or for none of the units (eg. New York State and Minnesota).

The priority order of patients differs between reserve categories. Therefore, a reserve emphasizing instrumental value could place more weight on frontline health workers than other reserve categories. In a reserve system, the extent to which policymakers prioritize instrumental value is flexible. It could be the primary ethical consideration and give absolute priority to frontline health workers, or it could be a secondary factor with a role of a tie-breaker. One example bases priority on clinical criteria only (such as SOFA score) for a more inclusive *community* category. For the instrumental value category, the priority could use the same clinical criteria as in the community category, but give essential personnel absolute priority for these units. The clinical criteria would be used as a tie-breaker in this category.

Another possibility is using the clinical criteria as the main criteria, and simply using frontline health worker status as a tie-breaker. Both methods are valid in a reserve system, although with different degrees of preferential treatment for frontline health workers. In fact, this example is similar to an idea explored in Pittsburgh's system (White et al., 2020). Although there is no reserve, the priority score uses SOFA score and comorbidities as in the system of South Carolina presented in Section 2.1. Life-cycle and instrumental considerations shape the tie-breakers:

> In the event that there are 'ties' in priority scores/categories between patients and not enough critical care resources for all patients with the lowest scores, life-cycle considerations should be used as the first tiebreaker, with priority going to younger patients. We recommend the following categories: age 12-40, age 41-60, age 61-75, older than age 75. We also recommend that individuals who are vital to the acute care response be given priority, which could be operationalized in the form of a tiebreaker.

The Pittsburgh system illustrates that preferential treatment for medical personnel can be modest: it is only a tie-breaker where the primary ethical considerations are based on utilitarian values. The Pittsburgh system also suggests the potential need for a reserve system, given that its designers hesitated recommending a single tie-breaker. Instead, the guideline suggests utilization of either the age or essential personnel status as a tie-breaker. Clearly, the designers saw utilitarian values as more justified than either the life-cycle principle or the instrumental value principle. On the other hand, the designers hesitated to make a definitive choice between these two ethical values in the event of tie-breaking. In contrast, a reserve system offers flexibility to use one of these tie-breakers in one of the categories and the other in another category.

2.3.2 Good Samaritan Reciprocity Category

In contrast to the potential instrumental value category presented in Section 2.3.1, the *good samaritan reciprocity category* formulated in this section may be more controversial.

Emmanuel et al. (2020) recommend priority for frontline health workers because it provides incentives that come with their high-risk responsibilities:

> . . . but giving them priority for ventilators recognizes their assumption of the high-risk work of saving others, and it may also discourage absenteeism.

Consider a hypothetical good samaritan reciprocity category in a triage protocol that reserves a small fraction of ventilators to those who have saved lives through their past good samaritan acts. These could be participants for clinical trials on vaccine or treatment development, altruistic donors who have donated their kidneys to a stranger, or people who have donated large quantities of blood over the years. Good samaritan status can also be provided for compatible patient-donor pairs who voluntarily participate in kidney exchange even though they do not have to, and save another patient's life who was incompatible with his/her donor. Sönmez, Ünver and Yenmez (2020) show that 180 additional kidney patients can receive living donor transplants, for every 10 percent of compatible pairs who participate in kidney exchange. A state task force can determine which acts "deserve" a good samaritan status.

Even though a triage rationing protocol will hopefully be rarely deployed, the mere existence of a modest reserve of this nature may mitigate more persistent and ongoing crises in other healthcare domains through the incentives it creates.

2.3.3 Unreserved Category

The next category we consider is an *unreserved category*, which is open to all patients who are in need of the medical unit. Fink (2020) cites a British researcher of the 2009 H1N1 flu pandemic cautioning that when a group of patients are excluded, "at the end you have got a society at war with itself. Some people are going to be told they don't matter enough." Exclusion criteria are an important axis of the contentious debate in rationing plans. Several state protocols such as Indiana, Kansas, South Carolina, and Tennessee exclude certain groups of patients (Guterl, 2020).

White and Lo (2020) have challenged exclusion criteria stating:

> This violates the principle of justice because it applies additional allocation criteria to some patients but not others, without making clear what is ethnically different about the patients that would justify doing so. Categorically excluding patients will make many feel their lives are "not worth saving," which may lead to perceptions of discrimination.

They argue that "morally irrelevant" considerations, such as sex, race, religion, intellectual disability, insurance status, wealth, citizenship, social status, or social connections should not used in rationing. Likewise, the Director of the Office for Civil Rights at the US Department of Health and Human Services recently rejected rationing protocols that do not protect the equal dignity of every human life, dismissing such protocols as "ruthless utilitarianism" (Fink, 2020).

Disabilities advocates have been particularly active in voicing opposition to rationing plans based solely on efficient resource allocation based on quality-adjusted life year considerations. Ne'eman (2020) argues that a crisis situation does not justify abandoning principles of nondiscrimination and that nondiscrimination should be seen as an end in and of itself. He also posits that provisions that exclude certain groups can undermine overall trust in the medical system "based on a well-founded fear of being sacrificed for the greater good."

A reserve system offers a potential path forward between these different points of view. In an unreserve category, a portion of the ventilators can be reserved for all patient types, no matter their characteristics. A random lottery ensures non-discrimination at least for this subset of ventilators.

2.4 Reserve Category Processing Sequence

In Section 3, we characterize all mechanisms that implement a reserve system that satisfy three basic properties. This characterization shows that any reserve system that satisfies three simple normative principles can be represented as an application of the celebrated deferred acceptance algorithm (Gale and Shapley, 1962). Conversely, any mechanism obtained in this way, satisfies these three basic properties as well. However, not all reserve systems in this class have a natural interpretation.

Our focus therefore is on a subset of these mechanisms, called **sequential reserve matching rules**, which simply process reserve categories in sequence. These mechanism were first introduced formally in Kominers and Sönmez (2016) in a more general environment with heterogenous units and multiple terms of allocation. Although not life-and-death situations, reserves are widespread in real-life applications including implementation of affirmative action policies in school choice in Boston (Dur et al., 2018), Chicago (Dur, Pathak and Sönmez, 2019), implementation of reservation policies in India (Sönmez and Yenmez, 2019*a*,*b*), and the allocation of immigration visas in the U.S. (Pathak, Rees-Jones and Sönmez, 2020*a*). As shown in these studies, the processing order of reserve categories is a key parameter with significant distributional implications.

To explain intuitively why processing order is important, imagine a simple scenario in which there are 60 ventilators. A medical ethics committee decides that there are two important principles: equal treatment of equals and prioritizing essential medical personnel. Based on their view, they define a reserve category for essential medical personnel, which reserves 50% of ventilators for them. Within this reservation, there is random allocation via lottery. The remaining 50% of ventilators are unreserved and open to all patients, including essential personnel. These are also allocated via lottery. Suppose that there are 60 essential personnel who need a ventilator and 60 other patients who do as well. If the essential personnel reserve category is processed first, then all 30 ventilators in this category are allocated to them. If a new lottery is drawn for the remaining 30 unreserved ventilators, since there are 60 other patients and 30 remaining essential personnel, the lottery results in 20 unreserved ventilators ($2/3 \times 30 = 20$) allocated to other patients and 10 unreserved ventilators ($1/3 \times 30 = 10$) to essential personnel. The total number of ventilators awarded to essential personnel is 40 and the total number awarded to other patients is 20. On the other hand, if the reserve category is processed after the unreserved units, then the 30 unreserved units will be split evenly between the essential personnel and other patients. The 30 remaining ventilators are all reserved and allocated to essential personnel. This results assigning 45 ventilators to essential personnel and 15 ventilators to other patients. Thus in this simple example, the choice of the processing sequence of categories is a matter of life or death for 5 essential medical personnel and 5 members of the general community.

As this simple example illustrates, our application to triage protocol design is yet another setting where reserve processing matters. Understanding the implications of reserve category processing order is especially critical given the emphasis on transparency in this application. Parallel to results in Dur et al. (2018) Dur, Pathak and Sönmez (2019), and Pathak, Rees-Jones and Sönmez (2020*a*), we show in Theorem 2 that processing a reserve category earlier in the process is detrimental to its beneficiaries. As such, the choice of reserve category processing sequence is a key design parameter, often playing a role similar in order of magnitude to adjustments in reserve size (Dur et al., 2018; Dur, Pathak and Sönmez, 2019).

2.5 Dynamics and Reassigning Ventilators

For many medical resources such as vaccines, tests, or single-use drug treatment, a reserve system would only need to be applied to a set of patients once. However, for non-perishable goods like ventilators, the reserve system would need to define time windows for periodic clinical assessments for patients who are already allocated ventilators. For instance, the New York State protocol calls for a periodic clinical assessments at 48 and 120 hours for patients who are on ventilator therapy (Zucker et al., 2015, p. 62). The protocol recommends that the assessment is conducted by an official committee as opposed to the patient's attending physician.

After a patient is re-evaluated, there are several possible steps. The patient may no longer need a ventilator and be taken off. The patient may continue to need a ventilator and not have any change in their reserve category. Or the patient may continue to need a ventilator and have a change in their reserve category status. This could happen, for instance, if their SOFA score deteriorates. In particular, it is possible that a patient is removed from a community reserve, but the ventilator is still allocated to the patient because she is eligible for another type of reserve category. In such a situation, this may result in the removal of treatment for another patient. However, this principle accords with statements in the New York State task force:

> ... removing a ventilator from a patient who worsens or does not improve so that another patient with a strong likelihood of survival may have an opportunity for treatment helps support the goal of savings the greatest number of lives in an influenza pandemic where there are a limited number of ventilators.

For this updated set of patient types and needs, and total ventilator supply (which also may change over time), the reserve system could apply. Beyond observing that a reserve system is flexible enough to handle these dynamics, we have nothing new to add to the discussion about the appropriate time window.

3 Formal Model and Analysis

While our primary application is pandemic rationing of scarce medical resources, in this section we provide a more general model which has several other applications including affirmative action in school choice, college admissions, assignment of government positions, and diversity in immigration visa allocation. Despite this generality, the terminology is tailored to our main application in the interest of being concrete.

There is a set of **patients** $\mathcal{I}$ and q identical **medical units** to allocate. There is a set of **reserve categories** $\mathcal{C}$. For each category $c \in \mathcal{C}$, r_c units are **reserved** so that $\sum_{c\in\mathcal{C}} r_c = q$. For each category $c \in \mathcal{C}$, there is a linear **priority order** π_c over patients and $\emptyset$. For each category $c \in \mathcal{C}$ and patient $i \in \mathcal{I}$, if $i \; \pi_c \; \emptyset$ we say that patient i is **eligible** for category-c units.

A **rationing problem** is a tuple $\langle \mathcal{I}, \mathcal{C}, (\pi_c)_{c\in\mathcal{C}}, q, (r_c)_{c\in\mathcal{C}} \rangle$.

Given a set of patients $I \subseteq \mathcal{I}$, a **matching** $\mu : I \to \mathcal{C} \cup \{\emptyset\}$ is a function that identifies the unit assignment of each patient in I such that, for every category $c \in \mathcal{C}$, $\left|\mu^{-1}(c)\right| \leq r_c$. For any patient $i \in I$, $\mu(i) = \emptyset$ means that the patient does not receive a unit and $\mu(i) \in \mathcal{C}$ means that the patient receives a unit reserved for category $\mu(i)$. Let $\mathcal{M}[I]$ denote the set of matchings for a given set of patients $I \subseteq \mathcal{I}$. Let $\mathcal{M} = \cup_{I\subseteq\mathcal{I}} \mathcal{M}[I]$ be the set of all matchings.

For a set of patients $I \subseteq \mathcal{I}$, for any matching $\mu \in \mathcal{M}[I]$ and any subset $I' \subseteq I$, let $\mu(I')$ denote the set of patients in I' who are matched in μ. More formally,

$$\mu(I') = \big\{ i \in I' : \mu(i) \in \mathcal{C} \big\}.$$

A **matching rule** φ is a function that selects a matching for each set of patients. Formally, it is a function $\varphi : 2^{\mathcal{I}} \to \mathcal{M}$ such that, for any set of patients $I \subseteq \mathcal{I}$, $\varphi[I] \in \mathcal{M}[I]$.

In real-life applications of our model, it is important to allocate units to qualified individuals without wasting any, and following the priorities for these units.

We next formulate these simple principles through three axioms:

Definition 1 *For a given set of patients $I \subseteq \mathcal{I}$, a matching $\mu \in \mathcal{M}[I]$ is **individually rational** if, for any $i \in I$, and $c \in \mathcal{C}$,*

$$\mu(i) = c \quad \Longrightarrow \quad i \; \pi_c \; \emptyset.$$

*A matching rule φ is **individually rational** if, for any set of patients $I \subseteq \mathcal{I}$, matching $\varphi[I]$ is individually rational.*

Our first axiom formulates the idea that individuals should only receive those units for which they are qualified. That is, units should be awarded only to eligible individuals. For ventilator rationing, any patient who is eligible for one category must also be eligible for any category. And if a patient is ineligible for all categories, then this patient can be dropped from the set of patients. Hence, individual rationality always holds for our main application.

Definition 2 *For a given set of patients $I \subseteq \mathcal{I}$, a matching $\mu \in \mathcal{M}[I]$ is **non-wasteful** if, for any $i \in I$ and $c \in \mathcal{C}$,*

$$i \; \pi_c \; \emptyset \quad \text{and} \quad \mu(i) = \emptyset \quad \Longrightarrow \quad \left|\mu^{-1}(c)\right| = r_c.$$

*A matching rule φ is **non-wasteful** if, for any set of patients $I \subseteq \mathcal{I}$, matching $\varphi[I]$ in non-wasteful.*

Our second axiom formulates the idea that no unit should be wasted. That is, if a unit is idle, there should not be any individual who is both unmatched and qualified for the unit. For ventilator rationing, each patient is eligible for all units, and therefore non-wastefulness in this context corresponds to either matching all the units or all the patients.

Definition 3 *For a given set of patients $I \subseteq \mathcal{I}$, a matching $\mu \in \mathcal{M}[I]$* ***respects priorities*** *if, for any $i, i' \in I$, and $c \in \mathcal{C}$,*

$$\mu(i) = c \;\; \textit{and} \;\; \mu(i') = \emptyset \quad \Longrightarrow \quad i \; \pi_c \; i'.$$

A matching rule φ ***respects priorities*** *if, for any set of patients $I \subseteq \mathcal{I}$, matching $\varphi[I]$ respects priorities.*

Finally our last axiom formulates the idea that for each category, the units should be allocated based on the priority order of individuals in this category.

It is important to emphasize that in many real-life guidelines for ventilator allocation all these axioms are implicit. Our formulation differs from these real life guidelines, only in its flexibility to allow for heterogeneity in priority orders guiding allocation of individual units.

3.1 Main Characterization Result

In this section, we explain how the deferred-acceptance algorithm of Gale and Shapley (1962) can be used to construct matching rules. We, furthermore, use this construction to characterize rules that satisfy desirable properties.

For every rationing problem, consider the following hypothetical two-sided matching market. On one side of the market, we have a set of patients $I \subseteq \mathcal{I}$. On the other side, we have categories $\mathcal{C}$. The capacity of category c is the number of units reserved for it, r_c, and category c is endowed with the linear order π_c.

Each patient has a preference ranking over categories $\mathcal{C}$ and the outside option of being unmatched, denoted by $\emptyset$. For each patient, category $c \in \mathcal{C}$ is ranked higher than the outside option if the patient is eligible for the patient. The remaining categories, for which the patient is not eligible, are ranked lower than the outside option.

Given a set of patients $I \subseteq \mathcal{I}$ and a preference ranking of patients in I, the deferred-acceptance algorithm (DA) produces a matching as follows.

Deferred Acceptance Algorithm (DA)

Step 1: Each patient in I applies to her most preferred category among categories for which she is eligible. Suppose that I_c^1 is the set of patients who apply to category c. Category c tentatively accepts applicants with the highest priority according to π_c until all patients in I_c^1 are chosen or r_c units are allocated, whichever comes first, and permanently rejects the rest. If there are no rejections, then stop.

Step k: Each patient who was rejected in Step $k-1$ applies to her next preferred category among categories for which she is eligible, if such a category exists. Suppose that I_c^k is the union of the set of patients who were tentatively accepted by category c in Step $k-1$, and the set of patients who just proposed to category c. Category c tentatively accepts patients in I_c^k with the highest priority according to π_c until all patients in I_c^k are chosen or r_c units are allocated, whichever comes first, and permanently rejects the rest. If there are no rejections, then stop.

For a fixed preference ranking of patients, this algorithm produces a matching for every set of patients. Therefore, DA induces a matching rule, which we call the **DA-induced matching rule**. DA-induced matching rules form a set of rules parametrized by the preference ranking of patients. The following result shows that this class gives us all matching rules satisfying desirable properties.

Theorem 1 *A matching rule is respecting priorities, non-wasteful, and individually rational if, and only if, it is a DA-induced matching rule.*

3.2 Sequential Reserve Matching Rules

In many applications with reserve-specific priorities, institutions may process reserve categories sequentially and allocate units associated with each category one at a time using its priority order. Formally, an **order of precedence** $\triangleright$ is a linear order over $\mathcal{C}$. For any two categories $c, c' \in \mathcal{C}$,

$$c \triangleright c'$$

means that category-c units are to be allocated before category-c' units. In this case, we say category c **has higher precedence** than category c'. Let Δ be the set of all orders of precedence.

Consider an order of precedence $\triangleright \in \Delta$. The **sequential reserve matching rule induced by** $\triangleright$, $\varphi_{\triangleright}$, is a matching rule defined as follows:

> Suppose categories are ordered under $\triangleright$ as
>
> $$c_1 \triangleright c_2 \triangleright \ldots \triangleright c_{|\mathcal{C}|}.$$
>
> For a given a set of patients $I \subseteq \mathcal{I}$, matching $\varphi_{\triangleright}[I]$ is found sequentially in $|\mathcal{C}|$ steps:
>
> **Step 1:** The highest priority r_{c_1} category-c_1-eligible patients in I according to π_{c_1} are matched with category-c_1 units. If there are less than r_{c_1} eligible patients in I than all of these eligible patients are matched with category-c_1 units. Let I_1 be the set of patients matched in Step 1.
>
> **Step k:** The highest priority r_{c_k} category-c_k-eligible patients in $I \setminus \cup_{k'=1}^{k-1} I_{k'}$ according to π_{c_k} are matched with category-c_k units. If there are less than r_{c_k} eligible patients in $I \setminus \cup_{k'=1}^{k-1} I_{k'}$ then all of these eligible patients are matched with category-c_k units. Let I_k be the set of patients matched in Step k.

Sequential reserve matching rules are individually rational, non-wasteful, and respect priorities. Thus, our characterization implies that there is an equivalent DA-induced rule to each of these.

Proposition 1 *Every sequential reserve matching rule is equivalent to a DA-induced rule.*

3.3 Comparative Statics for Sequential Reserve Matching Rules

In this section, we focus on a subclass of rationing problems and application of sequential reserve matching rules for these problems.

One way to utilize priorities of a category is to give a certain class of patients systematically higher priority than the rest.

To this end, we associate a **beneficiary group** with each category. In any set of patients $I \subseteq \mathcal{I}$, let $I_c \subseteq I$ be the set of beneficiaries of category $c \in \mathcal{C}$. All patients in the beneficiary group of a category are eligible for the category, i.e., $i \; \pi_c \; \emptyset$ for $i \in \mathcal{I}_c$. Furthermore, a category prioritizes patients in its beneficiary group: for a category $c \in \mathcal{C}$,

$$i \in \mathcal{I}_c \quad \text{and} \quad i' \in \mathcal{I} \setminus \mathcal{I}_c \quad \Longrightarrow \quad i \; \pi_c \; i'.$$

In particular, we will designate one category such that every patient is its beneficiary: there exists some $u \in \mathcal{C}$ such that $\mathcal{I}_u = \mathcal{I}$. We refer to category u as **unreserved** category and categories in $\mathcal{C} \setminus \{u\}$ as **reserve** categories with a slight abuse of terminology.

Each patient $i \in \mathcal{I}$ is beneficiary of at most one reserve category and this category is denoted by $\tau(i) \in \big(\mathcal{C} \setminus \{u\}\big) \cup \{\emptyset\}$ such that $\tau(i) = \emptyset$ means that patient i is no beneficiary of any reserve. Given a set of patients $I \subseteq \mathcal{I}$, let I_G be the set of patients who are not in any beneficiary group other than unreserved units:

$$I_G = I \setminus \cup_{c \in \mathcal{C} \setminus \{u\}} I_c.$$

We refer to patients in I_G as **general-community patients**.

In many applications, all priority orders of categories are determined using a single-priority order that we will refer to as π: For a category $c \in \mathcal{C}$,

$$\begin{pmatrix} i, i' \in \mathcal{I}_c \\ \text{or} \\ i, i' \in \mathcal{I} \setminus \mathcal{I}_c \end{pmatrix} \quad \text{and} \quad i \; \pi \; i' \quad \Longrightarrow \quad i \; \pi_c \; i'.$$

As a result, the unreserved category's priority order is identical to π as $\mathcal{I}_u = \mathcal{I}$:

$$\pi_u = \pi.$$

We refer to tuple $\langle \mathcal{I}, \mathcal{C}, \tau, \pi, q, (r_c)_{c \in \mathcal{C}} \rangle$ as a **rationing problem induced by a single priority order** π. In the rest of this section, we consider such problems.

For such problems, sequential reserve matching rules are a very natural alternative. However, the choice of the order of precedence may have important welfare implications. Intuitively speaking, the earlier a category is processed, the worse off are the patients in its beneficiary group. We formalize this result as follows:

Theorem 2 *Consider rationing problem induced by a single priority order π. Fix a set of patients $I \subseteq \mathcal{I}$ and a reserve category $c \in \mathcal{C} \setminus \{u\}$. Let $\rhd, \rhd' \in \Delta$ be two orders of precedence such that*

$$c_1 \rhd c_2 \rhd \ldots c_{k-1} \rhd c_k \rhd c_{k+1} \ldots \rhd c_{|C|}$$

and

$$c_1 \triangleright' c_2 \triangleright' \dots c_k \triangleright' c_{k-1} \triangleright' c_{k+1} \dots \triangleright' c_{|C|}$$

where $c_k = c$, that is, $\triangleright'$ is obtained from $\triangleright$ by only changing the order of c with its immediate predecessor category. Let $\mu = \varphi_{\triangleright}[I]$ and $\mu' = \varphi_{\triangleright'}[I]$.

Then,

$$\mu'(I_c) \subseteq \mu(I_c),$$

that is, every beneficiary of reserve category c matched by the sequential reserve rule under $\triangleright'$ is matched also under $\triangleright$.

As the following example shows, patients from other categories do not necessarily prefer the order of precedence $\triangleright'$ over $\triangleright$.

Example 1 *Let $\triangleright$ over $\triangleright'$ be two precedence orders such that*

$$u \triangleright c' \triangleright c$$

and

$$u \triangleright' c \triangleright' c'.$$

The priority order π is given as

$$i_1 \ \pi \ i_2 \ \pi \ i_3 \ \pi \ i_4.$$

Furthermore, suppose that there is only one medical unit reserved for each category and the beneficiary groups are $I_c = \{i_1\}$ and $I_{c'} = \{i_2, i_4\}$ and all patients are eligible for all categories.

Under $\triangleright$ the set of matched patients is $\{i_1, i_2, i_3\}$. Under $\triangleright'$ the set of matched patients is $\{i_1, i_2, i_4\}$. Therefore, when category c is processed before category c', i_3, who is a general-community patient is made worse off, and i_4, who is in the beneficiary group of c', is made better off.

3.4 Competing Interests Under Sequential Reserve Matching Rules

Theorem 2 states that the later a reserve category is processed the better for its members. This result motivates a closer look at sequential reserve matching rules induced by the following four classes of orders of precedence:

Reserves First Δ^{rf}: For any precedence $\trianglerighteq \in \Delta^{rf}$, any reserve $c \in \mathcal{C} \setminus \{u\}$ has higher precedence than unreserved units.

Under a reserves-first order of precedence, unreserved units are processed after reserve units. When there is a single reserve category, the resulting sequential reserve matching rule is uniquely defined, and first introduced by Hafalir, Yenmez and Yildirim (2013).

Reserves Last Δ^{rl}: For any precedence $\bar{\triangleright} \in \Delta^{rl}$, each reserve $c \in \mathcal{C} \setminus \{u\}$ has lower precedence than the unreserved category.

Under a reserves-last order of precedence, unreserved units are processed before reserved units.

Reserve-c Optimal Δ^c: For any precedence $\triangleright^c \in \Delta^c$, for any $c \in \mathcal{C} \setminus \{u\}$, for all $c' \in \mathcal{C} \setminus \{c, u\}$, reserve category c' has higher precedence than the unreserved category, which itself has higher precedence than reserve category c.

Under a reserve-c optimal order of precedence, all reserve units except those for reserve category c are processed prior to the unreserved units, whereas all reserve-c units are processed subsequent to the unreserved units.

Reserve-c Pessimal Δ_c: For any precedence $\triangleright_c \in \Delta_c$, for any $c \in \mathcal{C} \setminus \{u\}$, reserve category c has higher precedence than the unreserved category, which itself has higher precedence than each reserve category c' for all $c' \in \mathcal{C} \setminus \{c, u\}$.

Under a reserve-c pessimal order of precedence, all reserve-c units are processed prior to the unreserved units, which are processed prior to all other reserve units.

We obtain the sharpest results for these subclasses when we have reserve priority orders with an additional structure. We say that a priority profile $(\pi_c)_{c\in\mathcal{C}}$ has **soft caps** if for any $c \in \mathcal{C}$ and any $i \in \mathcal{I}$,

$$i \; \pi_c \; \emptyset.$$

Under a soft-caps rationing problem all patients are eligible for all categories.

We say that a priority profile $(\pi_c)_{c\in\mathcal{C}}$ has **hard caps** if for any $c \in \mathcal{C} \setminus \{u\}$ and any $i \in \mathcal{I} \setminus \mathcal{I}_c$,

$$\emptyset \; \pi_c \; i,$$

and for all $i \in \mathcal{I}$,

$$i \; \pi \; \emptyset.$$

Under a hard-caps rationing problem only beneficiary group of a reserve category is eligible for those reserve units.

We have the following results for hard-caps problems under sequential reserve matching rules induced by the four subclasses of orders of precedence defined above.

Theorem 3 *Consider a hard-caps rationing problem induced by a single priority order π. Fix a set of patients $I \subseteq \mathcal{I}$. Let $\underline{\triangleright} \in \Delta^{rf}$, $\overline{\triangleright} \in \Delta^{rl}$, $\overline{\mu} = \varphi_{\overline{\triangleright}}[I]$, $\underline{\mu} = \varphi_{\underline{\triangleright}}[I]$, and $\mu \in \mathcal{M}[I]$ be any matching that is respecting priorities, individually rational, and non-wasteful.*

Then,

$$\overline{\mu}(I_G) \subseteq \mu(I_G) \subseteq \underline{\mu}(I_G).$$

That is, of all matching rules that are respecting priorities, individually rational, and non-wasteful, a sequential reserve matching rule produces

- the best possible outcome under any reserves-first order of precedence, and
- the worst possible outcome under any reserves-last order of precedence

for general-community patients in a set inclusion sense.

Theorem 4 *Consider a hard-caps rationing problem induced by a single priority order π. Fix a set of patients $I \subseteq \mathcal{I}$ and a reserve category $c \in \mathcal{C} \setminus \{u\}$. Let $\triangleright_c \in \Delta_c$, $\triangleright^c \in \Delta^c$, $\mu_c = \varphi_{\triangleright_c}[I]$, $\mu^c = \varphi_{\triangleright^c}[I]$, and $\mu \in \mathcal{M}[I]$ be any matching that is respecting priorities, individually rational, and non-wasteful.*

Then,

$$\mu_c(I_c) \subseteq \mu(I_c) \subseteq \mu^c(I_c).$$

That is, of all matching rules that are respecting priorities, individually rational, and non-wasteful, a sequential reserve matching rule produces

- the best possible outcome under any reserve-c optimal order of precedence, and
- the worst possible outcome under any reserve-c pessimal order of precedence

for beneficiaries of reserve c in a set inclusion sense.

3.5 Related Theoretical Literature

Our study of reserve systems contributes to a large literature in matching market design focused on distributional issues. Several studies have examined allocation in the presence constraints such as minimum-guarantee reserves (or lower quotas), upper quotas, and regional quotas. An incomplete list of related work includes Abdulkadiroğlu (2005), Biro et al. (2010), Kojima (2012), Budish et al. (2013), Hafalir, Yenmez and Yildirim (2013), Westkamp (2013), Ehlers et al. (2014), Echenique and Yenmez (2015), Kamada and Kojima (2015), Kamada and Kojima (2017) Kamada and Kojima (2018), Aygün and Turhan (2016), Aygün and Bo (2016), Bo (2016), Dogan (2016), Kominers and Sönmez (2016), and Fragiadakis and Troyan (2017).

Our paper also introduces the triage protocol design problem into the market design literature. By considering a concrete real-world resource allocation problem, we contribute to the study of formal properties of specific allocation processes in the field and the study of alternative mechanisms. Studies in this vein include those on entry-level labor markets (Roth, 1984; Roth and Peranson, 1999), school choice (Balinski and Sönmez, 1999; Abdulkadiroğlu and Sönmez, 2003; Pathak and Sönmez, 2008, 2013*a*), spectrum auctions (Milgrom, 2000), kidney exchange (Roth, Sönmez and Ünver, 2004, 2005), internet auctions (Edelman, Ostrovsky and Schwarz, 2007; Varian, 2007), course allocation (Sönmez and Ünver, 2010; Budish, 2011), cadet-branch matching (Sönmez and Switzer, 2013; Sönmez, 2013), assignment of airport arrival slots (Schummer and Vohra, 2013; Schummer and Abizada, 2017) and refugee resettlement (Jones and Teytelboym, 2017; Delacrétaz, Kominers and Teytelboym, 2016; Andersson, 2017).

4 Conclusion

Because of the anticipated and ongoing shortage of key medical resources such as ventilators and ICU units during the Covid-19 pandemic, several leaders in the medical ethics community have made important recommendations regarding medical rationing. These recommendations reflect compromises between several ethical principles – maximizing lives, maximizing life-years, life-cycle considerations, instrumental values, reciprocity, protecting to the sickest, and non-discrimination. The priority point system, which aggregates all ethical dimensions into a single score, has become the norm. A reserve system offers additional flexibility and can simultaneously balance competing objectives in ways that a priority point system cannot.

Reserve systems can remedy challenges associated with prioritizing frontline health workers, an issue that has befuddled the medical ethics community. In the wake of the Covid-19 pandemic, Emmanuel et al. (2020) and White and Lo (2020) argue that these essential personnel should be prioritized, but several state standards express the concern that if frontline health workers are given priority, they could exhaust total supply and leave no ventilators for patients from the general community. By setting aside a fraction of ventilators for essential personnel as in a reserve system, not only can these concerns be overcome, but the ethical principles of reciprocity and instrumental valuation can also be reflected in the rationing system.

The theory of reserve systems in this paper is based on three simple normative criteria of non-wastefulness, individual rationality, and respect for priority. Non-wastefulness means no ventilator stays idle while there are patients who are in need. Individual rationality means if there are exclusion criteria for certain reserve categories, they have to be respected. For ventilators, it is natural to rule out category-specific exclusion criteria. Hence, in our application of triage protocol design, this criteria holds automatically. Respect for priority means that allocation of units in a given category has to follow the priority order of patients in that category. To the extent the priority order of patients represents the ethical principles guiding the allocation, this is also a natural principle. It is important to emphasize that these are the same principles that are embraced by the existing triage systems. What differs in our analysis is that we allow for different ethical principles to guide allocation decisions across different units.

In our formal analysis, we characterize the entire class of reservation policies that satisfy three minimal principles though implementation of the deferred-acceptance algorithm. As such, we also provide a full characterization of affirmative action policies. In the context of triage protocol design, one of the main insights for implementation of reserve systems is that processing a reserve category earlier is detrimental to its beneficiaries. This result is especially important since transparency is essential for triage protocol design.

While our analysis has focused narrowly on triage protocol design, there are several other direct applications of our formal theoretical results. These include immigration visa allocation in the United States (Pathak, Rees-Jones and Sönmez, 2020*a*), affirmative action in school choice systems in Boston, Chicago, New York City and Chile (Dur et al., 2018; Dur, Pathak and Sönmez, 2019; Correa et al., 2019), affirmative action for public school and government positions in India (Aygün and Turhan, 2017; Baswana et al., 2018; Sönmez and Yenmez, 2019*a*,*b*), and diversity plans for college admissions in Brazil (Aygün and Bo, 2016). We leave explorations of these connections to future drafts of this paper.

We conclude by noting that it is our hope that the triage rationing protocol we have analyzed will only be necessary in exceptional circumstances. The flexibility that a reserve system offers comes with the cost of needing to pay attention to several implementation details. It is also important to precisely define what the ethical principles that govern rationing and make sure that these details are made transparent to all stakeholders. Since failure to receive a unit corresponds to certain death, understanding the difference between mechanisms and choosing them accordingly is of vital importance.

A Proofs

In this section, we present proofs of our results.

Proof of Theorem 1. **Sufficiency:** We first prove that any DA-induced matching rule is respecting priorities, non-wasteful, and individually rational. Let $I \subseteq \mathcal{I}$ be a set of patients and μ be the outcome of a DA-induced matching rule for I.

Non-wastefulness: Suppose that $i \; \pi_c \; \emptyset$ and $\mu(i) = \emptyset$. By construction, every patient $i \in I$ ranks every category for which she is eligible; other categories are not ranked. Furthermore, each category $c \in \mathcal{C}$ chooses as many eligible patients as possible using its priority ranking π_c up to its capacity r_c. Since patient $i \in I$ is unmatched and i is eligible for c, $\left|\mu^{-1}(c)\right| = r_c$. Hence, matching rule μ is non-wasteful.

Respecting priorities: Suppose that $\mu(i) = c$ and $\mu(i') = \emptyset$. For every category $c \in \mathcal{C}$, π_c is used to choose eligible patients at every step of DA. Therefore, $\mu(i) = c$ implies $i \; \pi_c \; \emptyset$. Since $\mu(i') = \emptyset$, then it must be either because $\emptyset \; \pi_c \; i'$ or because $i \; \pi_c \; i'$. In the first case, we get $i \; \pi_c \; i'$ as well because π_c is transitive. Therefore, μ respects priorities.

Individual rationality: Suppose that $\mu(i) = c$. Then i must rank c, which means $i \; \pi_c \; \emptyset$.

Necessity: We now prove that any matching rule with the stated properties is a DA-induced matching rule. Let $I \subseteq \mathcal{I}$ be a set of patients and μ be the outcome of a matching rule that satisfies the stated properties for I. We need to specify the preference ranking of patients in I such that μ is the outcome of DA when agents have those preferences.

Consider patient $i \in I$ such that $\mu(i) = c$ where $c \in \mathcal{C}$. Since μ is individually rational, i must be eligible for category c. Let i rank category c first.

Consider patient $i \in I$ be such that $\mu(i) = \emptyset$. Let i rank categories in any order such that only eligible categories are ranked above the empty set.

For any such preference profile, in DA, for every category $c \in \mathcal{C}$, patients in $\mu^{-1}(c)$ apply to category c first. Unmatched patients apply to their first-ranked categories. Since μ respects priorities, an unmatched patient who is eligible for a category $c \in \mathcal{C}$ has a lower priority than any patient in $\mu^{-1}(c)$. Furthermore, since μ is non-wasteful, $|\mu^{-1}(c)| = r_c$ as there are unmatched patients in this category. Therefore, all unmatched patients in μ are rejected at the first step of DA and the other patients are tentatively accepted.

The unmatched patients in μ continue to apply to the other categories at which they are eligible. Since μ respects priorities and is non-wasteful, these patients are rejected from all categories for which they are eligible. The outcome of DA is such that, for each category $c \in \mathcal{C}$, all patients in $\mu^{-1}(c)$ are matched with c. Every patient $i \in I$ such that $\mu(i) = \emptyset$ is unmatched at the end of DA. Therefore, μ is the outcome of DA with the constructed patient preferences.
■

Proof of Proposition 1. Let $\triangleright$ be a precedence order and $\varphi_{\triangleright}$ be the associated sequential reserve matching rule. We show that there is a DA-induced matching rule μ, which is equivalent to $\varphi_{\triangleright}$.

For each patient $i \in \mathcal{I}$, consider a preference ranking $\succ_i$ such that all categories for which she is ineligible are ranked below the empty set and, furthermore, two categories for which she is eligible are ranked as follows: for any $c, c' \in \mathcal{C}$,

$$c \succ_i c' \iff c \triangleright c'.$$

For each patient $i \in \mathcal{I}$, consider another preference ranking $\succ'_i$ such that all categories are ranked above the empty set and, furthermore, for any $c, c' \in \mathcal{C}$,

$$c \succ'_i c' \iff c \triangleright c'.$$

Note that the relative ranking of two categories for which i is eligible is the same between $\succ_i$ and $\succ'_i$. Since for any category c and any patient $i \in \mathcal{I}$ who is ineligible for category c, $\emptyset \ \pi_c \ c$, the outcome of DA when the preference profile is $(\succ_i)_{i\in\mathcal{I}}$ and $(\succ'_i)_{i\in\mathcal{I}}$ are the same. Furthermore, when the preference profile is $(\succ'_i)_{i\in\mathcal{I}}$, DA works exactly like $\varphi_{\triangleright}$ because each category c, according to $\triangleright$, sequentially considers the set of patients who are not matched yet, applies π_c to choose patients and rejects patients who are not chosen.

Therefore, we conclude that $\varphi_{\triangleright}$ is the same rule as DA when the patient preference profile is $(\succ_i)_{i\in\mathcal{I}}$. ■

Proof of Theorem 2. Using the sequential reserve matching rule, in both matchings, all matched patients and their matched categories are identical until the end of step k-2. Let $J \subseteq I$ refer to the set of patients that has not been matched until the end of step k-2 under the construction of either matching. Let $c' = c_{k-1}$, the category preceding $c = c_k$ under $\triangleright$.

Observe that if $r_c \geq |J_c|$, then $\mu(J_c) = \mu'(J_c) = J_c$ and we are done with the proof. Therefore, assume $r_c < |J_c|$.

Two cases are possible regarding category c':

If $r_{c'} \leq |J_{c'}|$, then under both constructions, steps k-1 and k will result in matching the same set of patients to the category $c_k = c$ only patients from J_c and category $c' = c_{k-1}$ only patients from $J_{c'}$ and the rest of both matchings will be same as well after step k.

Therefore, assume that $r_{c'} > |J_{c'}|$. Then $|J_{c'}|$ units of category c' are assigned to category-c' beneficiaries in μ and μ' and the rest of the units are assigned with respect to priority order π conditional on the eligibility for category c'. These units are treated similar to the unreserved units. The difference for category-c beneficiaries is that weakly higher priority patients in J_c are unassigned just before category-c' units are processed while determining μ with respect to μ'. This is because, at this point while determining μ', the r_c highest priority patients in J_c are already assigned category-c units. Thus, $\mu'^{-1}(\{c, c'\}) \cap J_c \subseteq \mu^{-1}(\{c, c'\}) \cap J_c$ and $\mu^{-1}(c') \setminus \mu'^{-1}(c') \subseteq J_c$.

Let $J^0 = J$. If the set inclusion is not strict $\mu = \mu'$ and the theorem is proven. Therefore assume that the set inclusion is strict.

For each $h = 0, 1, ...$, we construct the following chains iteratively given that set J^h is constructed in the previous step:

Step h:

Chain of patients and categories $i^1 - c^1 - i^2 - c^2 - \ldots - i^n$ with $c^\ell \neq c'$ for each ℓ and $i^1, i^2, \ldots, i^n \in J^h$ such that

$$\begin{array}{lll}
\mu'(i^1) = c' & & \mu(i^1) = c^1 \neq c' \\
\mu'(i^2) = c^1 & & \mu(i^2) = c^2 \neq c^1 \\
\quad\vdots & & \quad\vdots \\
 & 1. & \mu(i^n) \in \{c', \emptyset\} \\
\mu'(i^n) = c^{n-1} & \text{or} & \\
 & 2. & \mu(i^n) = c^n \neq c' \quad \text{and} \quad \Big(\mu'^{-1}(c^n) \setminus \mu^{-1}(c^n)\Big) \cap J^h = \emptyset
\end{array}$$

such that i^ℓ is the highest priority patient according to $\pi_{c^{\ell-1}}$ in $\Big(\mu'^{-1}(c^{\ell-1}) \setminus \mu^{-1}(c^{\ell-1})\Big) \cap J^h$ setting $c^0 = c'$ for each ℓ.

Then set $J^{h+1} = J^h \setminus \{i^1, i^2, \ldots, i^n\}$, and continue with step $h+1$.

We continue until all patients in $\mu^{-1}(c') \setminus \mu'^{-1}(c')$ are considered to start such a chain.

Consider $h = 0$:

First consider the first type of chain ending, $\mu(i^n) \in \{c', \emptyset\}$:

Observe that $\{i^1, \ldots, i^{n-1}\} \cap J_c^h = \emptyset$. As otherwise, the first category-c beneficiary patient in this sequence would be the highest priority patient according to π in $\big(\mu^{-1}(c') \setminus \mu'^{-1}(c')\big) \cap J^h$ and the sequence would have ended at that patient.

If $\mu(i^n) = c'$, then $i^n \in J_c^h$ and $i^1, \ldots, i^n$ are matched both in μ and μ'. On the other hand, if $\mu(i^n) = \emptyset$ again observe that $i^n \notin J_c^h$: As otherwise, she would be the highest priority unmatched patient in J_c^h when we fill c^{n-1}; as none of the patients $i^1, \ldots, i^{n-1}$ are in J_c^h, she should be matched in μ with c', leading to a contradiction.

Next consider the second type of chain ending, $\mu(i^n) = c^n \neq c'$ and $\big(\mu^{-1}(c^n) \setminus \mu'^{-1}(c^n)\big) \cap J^h = \emptyset$:

In this case all patients $i^1, \ldots, i^n$ are matched in μ and μ'.

Thus, we showed that in such a chain no new patient from beneficiary group J_c^h is matched in μ' with respect to μ.

We iteratively continue over h, and each step, we end up with a chain of patients and categories such that no new patient from beneficiary group J_c^h is matched in μ' while she was not matched in μ.

This completes the proof of $\mu'(I_c) \subseteq \mu(I_c)$. ■

Proof of Theorem 3. We prove the result in two separate claims. By Theorem 1, μ is generated by a DA-induced matching rule. Note that in a hard-caps problem, general-community patients apply only to unreserved units in any DA-induced matching rule.

Claim 1: $\overline{\mu}(I_G) \subseteq \mu(I_G)$.

Proof. Note that $\overline{\mu}$ is the outcome of DA when each patient first applies to unreserved units. Therefore, $\overline{\mu}^{-1}(u)$ the set of r_u-highest priority patients according to π, which implies that $i \in \overline{\mu}(I_G)$ if, and only if, i is among the r_u-highest priority patients according to π in I. As a result, in any DA-induced matching rule outcome, $i \in \overline{\mu}(I_G)$ is matched with category u, which shows that $\overline{\mu}(I_G) \subseteq \mu(I_G)$.

Claim 2: $\mu(I_G) \subseteq \underline{\mu}(I_G)$.

Proof. Note that $\underline{\mu}$ is the outcome of DA when each beneficiary patient of a reserve first applies to that reserve before the unreserved units.

On the other hand, in DA matching rule with outcome μ, potentially some beneficiary patients of a reserve category first apply to unreserved units before their reserve.

We will show that $\left|\mu^{-1}(u) \setminus I_G\right| \geq \left|\underline{\mu}^{-1}(u) \setminus I_G\right|$.

Consider a patient $i \in \underline{\mu}^{-1}(u) \setminus I_G$. Suppose $\tau(i) = c$ for some reserve category c. She has a lower priority than at least r_c beneficiary patients of c according to π and is among the r_u highest priority patients in $I \setminus \cup_{c' \in \mathcal{C}} \underline{\mu}^{-1}(c')$ according to π. Thus, if $\mu(i) \neq u$, then either

- $\mu(i) = c$, which can only happen if one patient $j \in I_c$, such that (i) she has a higher priority than i according to π and (ii) $\underline{\mu}(j) = c$, applied to u and got matched with u in μ; or
- $\mu(i) = \emptyset$, which can only happen if one patient $j \in I_{c'} \setminus \mu^{-1}(u)$ with a higher priority than i according to π for some reserve category $c' \neq c$ applied to u and got matched with u in μ.

Thus, anytime such a patient i is rejected by u in the DA resulting with μ, some distinct patient $j \in \left(\mu^{-1}(u) \setminus \underline{\mu}^{-1}(u)\right) \setminus I_G$ is matched with u in μ.

This shows $\left|\mu^{-1}(u) \setminus I_G\right| \geq \left|\underline{\mu}^{-1}(u) \setminus I_G\right|$, which in turn shows $\mu(I_G) \subseteq \underline{\mu}(I_G)$ as in both DA rules generating either matching, general-community patients only apply to u. ■

Proof of Theorem 4. We prove the result in two separate claims. By Theorem 1, μ is generated by a DA-induced matching rule. Note that in a hard-caps problem, general-community patients apply only to unreserved units in any DA-induced matching rule. For a patient $i \in I$ who is eligible at category $c \in \mathcal{C}$, let $\text{rank}(i; I, \pi_c)$ denote the rank of i among patients in I according to π_c. For example, the patient with the highest priority has rank one.

Claim 1: $\mu_c(I_c) \subseteq \mu(I_c)$.

Proof. Note that $\mu_c(I_c)$ is generated by a reserve-c pessimal precedence order. Therefore,

$$\mu_c(I_c) = \widehat{I}_c \cup \widetilde{I}_c$$

where

$$\widehat{I}_c = \{i \in I_c : \text{rank}(i; I_c, \pi) \leq r_c\}$$

and

$$\widetilde{I}_c = \{i \in I_c \setminus \widehat{I}_c : \text{rank}(i; I \setminus \widehat{I}_c, \pi) \leq r_u\}.$$

Let $i \in \mu_c(I_c)$. Then there are two cases two consider. First consider the case that $i \in \widehat{I}_c$. Then if $\mu(i) = 0$, then there exists $i' \in \mu^{-1}(c)$ such that $i \; \pi_c \; i'$, which cannot happen because μ respects priorities. Therefore, $\mu(i) \neq \emptyset$ which is equivalent to $i \in \mu(I_c)$. This implies $\mu_c(I_c) \subseteq \mu(I_c)$.

Suppose that $i \in \widetilde{I}_c$. Suppose, for contradiction, that $\mu(i) = \emptyset$. Since μ is non-wasteful, $|\mu^{-1}(c)| = r_c$, and since μ respects priorities, for every $i' \in \mu^{-1}(c)$, $i' \; \pi_c \; i$. Likewise, $|\mu^{-1}(u)| = r_u$, and since μ respects priorities, for every $i' \in \mu^{-1}(u)$, $i' \; \pi_u \; i$. As a result, $\text{rank}(i; I, \pi) > r_c + r_u$, which is a contradiction to the construction of $\widetilde{I}_c$.Therefore, $\mu(i) \neq \emptyset$, which implies $\mu_c(I_c) \subseteq \mu(I_c)$.

Claim 2: $\mu(I_c) \subseteq \mu^c(I_c)$.

Proof. We show that $|\mu(I_c)| \leq |\mu^c(I_c)|$. The claim then follows because both μ and μ^c respect priorities. Since both μ and μ^c are non-wasteful, the inequality holds if, and only if, the number of category-c patients assigned to unreserved units in μ is less than the number of category-c patients assigned to unreserved units in μ^c because we are considering a hard-caps rationing problem.

For every category $c' \in \mathcal{C}$, let

$$\widehat{I}_{c'} = \{i \in I_{c'} : \text{rank}(i; I_{c'}, \pi) \leq r_{c'}.\}$$

The set of category-c patients matched to unreserved units in μ^c is

$$\widetilde{I}_c = \{i \in I_c : \text{rank}(i; I \setminus \cup_{c' \in \mathcal{C} \setminus \{c,u\}} \widehat{I}_{c'}, \pi) \leq r_u\}.$$

Because μ is the outcome of DA for a preference profile of patients, the set of category-c patients matched to unreserved units in μ is at most

$$\{i \in I_c : \text{rank}(i; I \setminus \cup_{c' \in \mathcal{C} \setminus \{c,u\}} \mu^{-1}(c'), \pi) \leq r_u\}.$$

The cardinality of this set is smaller than $|\widetilde{I}_c|$ because, by construction, $|\widehat{I}_{c'}| \geq |\mu^{-1}(c')|$ and every patient in $\widehat{I}_{c'} \setminus \mu^{-1}(c')$ has a higher priority according to π than every patient in $\mu^{-1}(c') \setminus \widehat{I}_{c'}$. ■

References

Abdulkadiroğlu, Atila. 2005. "College Admissions with Affirmative Action." *International Journal of Game Theory*, 33(4): 535–549.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93: 729–747.

Agarwal, Nikhil, Itai Ashlagi, Michael Rees, Paulo Somaini, and Daniel Waldinger. 2019. "An Empirical Framework for Sequential Assignment: The Allocation of Deceased Donor Kidneys." Working paper, MIT.

Andersson, Tommy. 2017. "Refugee Matching as a Market Design Application." Working paper.

Aygün, Orhan, and Bertan Turhan. 2016. "Dynamic Reserves in Matching Markets." *Working paper.*

Aygün, Orhan, and Bertan Turhan. 2017. "Large-Scale Affirmative Action in School Choice: Admissions to IITs in India." *American Economic Review*, 107(5): 210–13.

Aygün, Orhan, and Inacio Bo. 2016. "College Admission with Multidimensional Reserves: The Brazilian Affirmative Action Case." *Working paper, WZB Berlin Social Science Center*, 869.

Balinski, Michel, and Tayfun Sönmez. 1999. "A Tale of Two Mechanisms: Student Placement." *Journal of Economic Theory*, 84: 73–94.

Ball, Robert, and Phil Schneider. 2009. "South Carolina Prepares for Pandemic Influenza: An Ethical Perspective." South Carolina Pandemic Influenza Ethics Task Force.

Baswana, Surender, Partha Pratim Chakrabarti, Sharat Chandran, Yashodhan Kanoria, and Utkarsh Patange. 2018. "Centralized admissions for engineering colleges in India." forthcoming, *INFORMS Journal on Applied Analytics*.

Biddison, Elizabeth, Howard Gwon, Monica Schoch-Spana, Robert Cavalier, Douglas B. White, Timothy Dawson, Peter B. Terry, Alex John London, Alan Regenberg, Ruth Faden, and Eric S. Toner. 2013. "The Community Speaks: Understanding Ethical Values in Allocation of Scarce Lifesaving Resources during Disasters." *Annals of the American Thoracic Society*, 11: 777–783.

Biro, P., T. Fleiner, R.W. Irving, and D.F. Manlove. 2010. "The College Admissions Problem with Lower and Common Quotas." *Theoretical Computer Science*, 411(34-26): 3136–3153.

Bo, Inacio. 2016. "Fair Implementation of Diversity in School Choice." *Games and Economic Behavior*, 97: 54–63.

Budish, Eric. 2011. "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes." *Journal of Political Economy*, 119(6): 1061–1103.

Budish, Eric, Yeon-Koo Che, Fuhito Kojima, and Paul R. Milgrom. 2013. "Designing Random Allocation Mechanisms: Theory and Applications." *American Economic Review*, 103(2): 585–623.

Christian, M.D., L Hawryluck, R. S. Wax, T Cook, N. M. Lazar, M. S. Herridge, M. P. Muller, D. R. Gowans, W. Fortier, and F. M. Burkle. 2006. "Development of a triage protocol for critical care during an influenza pandemic." *Canadian Medical Association Journal*, 175: 1377–1381.

Correa, Jose, Rafael Epstein, Juan Escobar, Ignacio Rios, Bastian Bahamondes, Carlos Bonet, Natalie Epstein, Nicolas Aramayo, Martin Castillo, Andres Cristi,

and Boris Epstein. 2019. “School Choice in Chile.” *EC ’19*, 325–343. New York, NY, USA:ACM.

Daughtery-Biddison, Lee, Howard Gwon, Alan Regenberg, Monica Schoch-Spana, and Eric Toner. 2017. “Maryland Framework for the Allocation of Scarce Life-Sustaining Medical Resources in a Catastrophic Public Health Emergency.” https://www.law.umaryland.edu/media/SOL/pdfs/Programs/Health-Law/MHECN/ASR%20Framework_Final.pdf.

Delacrétaz, David, Scott Duke Kominers, and Alexander Teytelboym. 2016. “Refugee Resettlement.” Working paper.

Dittmar, Pia. 2018. “Gender Quotas - Balancing Act Between Efficient Policies and Gap Fillers.” EU-Logos Athena.

Dogan, Battal. 2016. “Responsive Affirmative Action in School Choice.” *Journal of Economic Theory*, 165: 69–105.

DoH, Utah. 2018. “Utah Crisis Standards of Care Guidelines.” Utah Department of Health.

Dur, Umut, Parag Pathak, and Tayfun Sönmez. 2019. “Explicit vs. Statistical Preferential Treatment in Affirmative Action: Theory and Evidence from Chicago’s Exam Schools.” forthcoming, *Journal of Economic Theory*.

Dur, Umut, Scott Kominers, Parag Pathak, and Tayfun Sönmez. 2018. “Reserve Design: Unintended Consequences and The Demise of Boston’s Walk Zones.” *Journal of Political Economy*, 126(6).

Echenique, Federico, and M. Bumin Yenmez. 2015. “How to Control Controlled School Choice.” *American Economic Review*, 105(8): 2679–2694.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz. 2007. “Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords.” *American Economic Review*, 97(1): 242–259.

Ehlers, Lars, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. 2014. “School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds.” *Journal of Economic Theory*, 153: 648–683.

Emmanuel, Ezekial J., Govind Persad, Ross Upshur, Beatriz Thome, Michael Parker, Aaron Glickman, Cathy Zhang, Conor Boyle, Maxwell Smith, and James Phlips. 2020. “Fair Allocation of Scarce Medical Resources in the Time of Covid-19.” *New England Journal of Medicine*, March 23.

Fink, Sheri. 2020. “U.S. Civil Rights Office Rejects Rationing Medical Care Based on Disability, Age.” New York Times, March 29.

Fragiadakis, Daniel, and Peter Troyan. 2017. "Improving Matching under Hard Distributional Constraints." *Theoretical Economics*, 12(2): 863–908.

Gale, David, and Lloyd S. Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69: 9–15.

Guterl, Fred. 2020. "Who Should Doctors Save? Inside the Debate About How to Ration Coronavirus Care." *Newsweek*, April 3.

Hafalir, Isa E., M. Bumin Yenmez, and Muhammed A. Yildirim. 2013. "Effective Affirmative Action in School Choice." *Theoretical Economics*, 8(2): 325–363.

Israni, A. K., N. Salkowski, S. Gustafson, J. J. Snyder, J. J. Friedewald, R. N. Formica, X. Wang, E. Shteyn, W. Cherikh, D. Stewart, C. J. Samana, A. chung, A. Hart, and B. L. Kasiske. 2014. "New National Allocation Policy for Deceased Donor Kidneys in the United States and Possible Effects on Patient Outcomes." *Journal of the American Society of Nephrology*, 25(8): 1842–1848.

Iten, Anne, and Pascal Strupler. 2018. "Swiss Influenza Pandemic Plan." Federal Office of Public Health, Swiss Confederation.

Jones, Will, and Alexander Teytelboym. 2017. "The Local Refugee Match: Aligning Refugees' Preferences with the Capacities and Priorities of Localities." *Journal of Refugee Studies*, 31(2): 152–178.

Kamada, Yuichiro, and Fuhito Kojima. 2015. "Efficient Matching under Distributional Constraints: Theory and Applications." *American Economic Review*, 105(1): 67–99.

Kamada, Yuichiro, and Fuhito Kojima. 2017. "Stability Concepts in Matching with Distributional Constraints." *Journal of Economic Theory*, 168: 107–142.

Kamada, Yuichiro, and Fuhito Kojima. 2018. "Stability and Strategy-proofness for Matching with Constraints: A Necessary and Sufficient Condition." *Theoretical Economics*, 13(2): 761–794.

Kojima, Fuhito. 2012. "School Choice: Impossibilities for Affirmative Action." *Games and Economic Behavior*, 75(2): 685–693.

Kominers, Scott Duke, and Tayfun Sönmez. 2016. "Matching with Slot-specific Priorities: Theory." *Theoretical Economics*, 11(2): 683–710.

Lee, Alice. 2014. "Gender Quotas Worked in Norway. Why Not Here?" *The New Republic*, September 5.

Michigan. 2012. "Guidelines for Ethical Allocation of Scarce Medical Resources and Services During Public Health Emergencies in Michigan." Department of Community Health, Office of Public Health Preparedness.

Milgrom, Paul R. 2000. “Putting Auction Theory to Work: The Simultaneous Ascending Auction.” *Journal of Political Economy*, 108: 245–272.

Ne’eman, Ari. 2020. “I Will Not Apologize for My Needs.” *New York Times*, March 23.

OPTN. 2014. “Organ Procurement and Transplantation Network Policies.” Technical Report.

Pathak, Parag A., and Tayfun Sönmez. 2008. “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism.” *American Economic Review*, 98(4): 1636–1652.

Pathak, Parag A., and Tayfun Sönmez. 2013*a*. “School Admissions Reform in Chicago and England: Comparing Mechanisms by their Vulnerability to Manipulation.” *American Economic Review*, 103(1): 80–106.

Pathak, Parag A., and Tayfun Sönmez. 2019. “Implementation Issues in 10% Reservation.” *The Hindu.* Available at: https://www.thehindu.com/opinion/op-ed/implementation-issues-in-10-reservation/article27130396.ece, last accessed: March 24, 2020.

Pathak, Parag, Alex Rees-Jones, and Tayfun Sönmez. 2020*a*. “Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms.” NBER Working Paper 26767.

Pathak, Parag, Alex Rees-Jones, and Tayfun Sönmez. 2020*b*. “Reversing Reserves.” NBER Working Paper.

Pathak, Parag, and Tayfun Sönmez. 2013*b*. “Recommendation on Algorithm Processing, Public Testimony to EAC.” Available at: https://economics.mit.edu/files/16966.

Rankin, Jennifer. 2020. “EU revises plans for mandatory quotas of women on company boards.” The Guardian, March 5.

Roth, Alvin E. 1984. “The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory.” *Journal of Political Economy*, 92: 991–1016.

Roth, Alvin E., and Elliott Peranson. 1999. “The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design.” *American Economic Review*, 89: 748–780.

Roth, Alvin E., Tayfun Sönmez, and Utku Ünver. 2004. “Kidney Exchange.” *Quarterly Journal of Economics*, 119: 457–488.

Roth, Alvin E., Tayfun Sönmez, and Utku Ünver. 2005. “Pairwise Kidney Exchange.” *Journal of Economic Theory*, 125: 151–188.

Schummer, James, and Azar Abizada. 2017. “Incentives in Landing Slot Problems.” *Journal of Economic Theory*, 170: 29–55.

Schummer, James, and Rakesh V. Vohra. 2013. "Assignment of Arrival Slots." *American Economic Journal: Microeconomics*, 5(2): 164–185.

Sönmez, Tayfun. 2013. "Bidding for Army Career Specialties: Improving the ROTC Branching Mechanism." *Journal of Political Economy*, 121(1): 186–219.

Sönmez, Tayfun, and Bumin Yenmez. 2019*a*. "Affirmative Action in India via Vertical and Horizontal Reservations." Working paper.

Sönmez, Tayfun, and Bumin Yenmez. 2019*b*. "Constitutional Implementation of Vertical and Horizontal Reservations in India: A Unified Mechanism for Civil Service Allocation and College Admissions." Working paper.

Sönmez, Tayfun, and M. Utku Ünver. 2010. "Course Bidding at Business Schools." *International Economic Review*, 51(1): 99–123.

Sönmez, Tayfun, and Tobias Switzer. 2013. "Matching with (Branch-of-Choice) Contracts at the United States Military Academy." *Econometrica*, 81(2): 451–488.

Sönmez, Tayfun, Utku Ünver, and Bumin Yenmez. 2020. "Incentivized Kidney Exchange." forthcoming, *American Economic Review*.

Tong, Rosemarie, and Leah Devlin. 2007. "Stockpiling Solutions: North Carolina's Ethical Guidelines for an Influenza Pandemic." North Carolina Institute of Medicine, Department of Health and Human Services Division of Public Health.

Truog, Robert D., Christine Mitchell, and George Q. Daley. 2020. "The Toughest Triage – Allocating Ventilators in a Pandemic." *New England Journal of Medicine*, March 23.

Varian, Hal R. 2007. "Position Auctions." *International Journal of Industrial Organization*, 25(6): 1163–1178.

Vawter, Dorothy E., J. Eline Garrett, Karen G. Gervais, Angela Witt Prehn, Debra A. DeBruin, Carol A. Tauer, Elizabeth Parilla, Joan Liaschenko, and Mary Faith Marshall. 2010. "For the Good of Us All: Ethically Rationing Health Resources in Minnesota in a Severe Influenza Pandemic." Minnesota Pandemic Ethics Project, Minnesota Department of Health.

Vincent, J. L., R Moreno, J Takala, S Willatts, A Mendonca, and H Bruining. 1996. "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure." *Intensive Care Medicine*, 22(7): 707–710. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine.

Westkamp, Alexander. 2013. "An Analysis of the German University Admissions System." *Economic Theory*, 53(3): 561–589.

White, Douglas B., and Bernard Lo. 2020. "A Framework for Rationing Ventilators and Critical Care Beds During the Covid-19 Pandemic." *Journal of the American Medical Association*. doi:10.1001/jama.2020.5046.

White, Douglas B., Mitchell H. Katz, John M. Luce, and Bernard Lo. 2009. "Who Should Receive Life Support During a Public Health Emergency? Using Ethical Principles to Improve Allocation Decisions." *Annals of Internal Medicine*, 150(2): 132–138.

White, Douglas B., Mithcell Katz, John M. Luce, and Bernard Lo. 2020. "Allocation of Scarce Critical Care Resources During a Public Health Emergency." University of Pittsburgh, Department of Critical Care Medicine, March 26.

Zucker, Howard A, Karl P. Adler, Donald P. Berens, Rock Brynner, Karen Butler, Yvette Calderon, Carolyn Corcoran, Nancy Neveloff Dubler, Paul J. Edelson, Joseph J. Fins, Rev. Francis H. Geer, Samuel Gorovitz, Cassandra Henderson, Hassan Khouli, Fr. Joseph W. Koterski, Hugh Maynard-Reid, John Murnane, Karen Porter, Robert Swidler, Sally True, Stuart C. Sherman, Susia A. Han, and Valier Gutmann Koch. 2015. "Ventilator Allocation Guidelines." New York State Task Force on Life and the Law, New York Department of Health.