Explainable Affirmative Action

NICK ARNOSTI, CARLOS BONET, AND JAY SETHURAMAN

We study Prioritized Selection Problems in which an organization is presented with a set of individuals, and must choose which subset to accept. The organization makes a selection based on a priority ranking of individuals as well as other observable characteristics. We study *outcome based selection rules*, which are defined by a collection of feasible selections and a greedy processing algorithm.

Our first contribution is to argue that outcome based selection rules are uniquely *explainable*. We support this claim with two characterization results. First, these rules are the only rules that are *monotonic*, *priority non-bossy*, and *lower invariant*. Second, given a collection of feasible selections, the greedy processing rule chooses the only one that *respects priorities*.

Our second contribution is to demonstrate that some existing algorithms (such as "minimum guarantee" reserves) implement outcome based selection rules, while others (such as "over and above" reserves) do not. We also define a new algorithm for reserve matching ("maximal reserves"), which selects an outcome that priority dominates the selections made by slot-specific priority algorithms.

Finally, we connect these ideas with the Chilean Constitutional Assembly election. In this election, candidates were ranked by the number of votes received, and feasible selections had to allocate the correct number of seats to each party and ensure gender parity. We show that the rule that was implemented is the outcome based selection rule associated with these constraints.

1 INTRODUCTION

We study *Prioritized Selection Problems*, in which an organization is presented with a set of individuals, and must choose which subset to accept. We assume that the organization has a complete priority ranking of individuals. This ranking may be determined by various factors, including date of application, test scores or other measures of merit, a lottery, or some combination of these. The organization also has objectives or constraints that are not reflected in the priority ranking, but influence the final selection. Examples of such settings include:

- **Chilean Constitutional Assembly Elections.** In 2021, Chile held an election to choose 155 representatives to rewrite its constitution. Candidates were prioritized by votes received. However, the elected representatives from each district had to include an equal number of men and women, and include an appropriate number of members of each party (based on that party's vote share in the district).
- **US Immigrant Visa Programs.** Both H1B and Diversity Visas are allocated by lottery. However, some H1B visas are reserved for applicants with advanced degrees from US institutions. Furthermore, the Diversity Visa lottery includes upper quotas on the number of individuals selected from each country and region.
- **Affordable Housing Allocation.** People who apply for affordable housing through New York City's "HousingConnect" website are placed in a uniformly random order. However, half of the units in each building are reserved for residents of the local community district, additional units are set aside for people with disabilities and city employees, and each unit has eligibility criteria based on household size and income.
- **School Assignment.** Specialized high schools in New York City admit students on based on a test score. The "Discovery Program" reserves seats at these schools for students from disadvantaged backgrounds [Faenza et al., 2022]. Priority for university admissions is determined by exam scores in many countries. In India, positions are reserved for people from disadvantaged castes, people with disabilities, women, and people from the local state [Baswana et al., 2019]. In Brazil, positions are reserved for students who attended public high school, are low-income, or belong to a disadvantaged minority group [Aygun and Bo, 2019].

- **Civil Service Positions.** In India, priority for government positions is determined by standardized exam scores. Many positions are reserved for applicants from historically marginalized groups, including women and "backwards classes." [Aygün and Turhan, 2022, Sönmez and Yenmez, 2022].

We refer to goals unrelated to the priority ranking as "diversity constraints" or "affirmative action objectives."

It is not obvious how to best *combine* the priority ranking and affirmative action objectives to reach a final selection. In practice, there are many ad hoc methods for doing so, but these algorithms frequently produce outcomes that can be difficult to explain to participants. For example,

- **New York Affordable Housing:** Applicants from outside the community district are often passed over for lower-ranking applicants from the community district, even if the community preference constraint is not binding.
- **New York Specialized High Schools:** high scoring students who are eligible for the Discovery Program are sometimes assigned to their second or third choice, while seats in their preferred schools are offered to lower-scoring students [Faenza et al., 2022].
- **Brazilian University Admissions**: cutoff scores for applicants claiming multiple privileges are frequently *higher* than corresponding cutoffs for applicants claiming a subset of those privileges [Aygun and Bo, 2019].
- **Indian Civil Service:** Numerous lawsuits document cases where the algorithm selected a woman from a majority group instead of another woman from a backwards class *with a higher exam score* [Sönmez and Yenmez, 2022].

Each of these examples features an apparent violation of priority that is difficult to justify on the basis of the diversity constraints. Individually, each example provides an interesting research opportunity, and indeed the papers mentioned above generally provide "solutions" (new algorithms) for the particular context they study. Collectively, however, these examples illustrate the challenge of implementing affirmative action and diversity policies effectively.

Our primary contribution is to introduce a family of rules which we argue are uniquely *explainable*. These rules avoid the priority violations described above, ensuring that any violations are necessitated by the diversity constraints. Furthermore, we provide an axiomatic characterization that shows that these are the only rules that respond in "explainable" ways to changes in the priority order. We study when explainable rules can be implemented without losing efficiency, and establish that the algorithm used to elect members of Chile's recent Constitutional Assembly belongs to our family of rules. We now elaborate on each contribution.

1.1 Axiomatizing Explainability

Given a set of individuals $\mathcal{I}$, a *selection* is a subset of these individuals, and a *selection rule* φ is a function mapping each priority order $\succ$ to the resulting selection $\varphi(\succ)$. We consider three natural properties[1] for a selection rule to satisfy, which constrain how the rule responds to changes in the priority list.

- **Monotonic**: Improving a selected individual's priority never harms that individual.
- **Lower invariant**: Whether an individual is selected depends only on the priorities of higher-priority individuals, and not on those of lower-priority individuals.
- **Priority Non-bossy**: If a change in an individual's priority does not change that individual's outcome, then it does not change the final selection.

[1]Similar properties have been studied for matching and assignment problems under different names in the literature; we discuss these in Section 3.2.

	Selection Rule 1	Selection Rule 2
Algorithmic Description	Hire top woman, then Hire top remaining candidate.	Hire top candidate, then Hire top remaining woman.
Outcome Description	Selections must have 2 people, at least 1 woman, and must respect priorities.	✗

Fig. 1. Outcome-based selection rules provide a new way to describe existing procedures. For example, Theorem 2 establishes that the algorithm used to select members of Chile's Constitutional Assembly is equivalent to an outcome based rule, despite having a very different description. More simply, Selection Rule 1 implements a "minimum guarantee" reserve, and is equivalent to the outcome-based rule associated with the constraints that selections must contain at least one woman and at most two people. Other rules, however – such as Selection Rule 2 (an "over and above" reserve) – do not have any outcome-based description.

Theorem 1 establishes that a selection rule satisfies all three of these properties if and only if it is *outcome-based.* Outcome-based selection rules are defined by a collection of feasible selections $\mathcal{F} \subseteq \{0,1\}^{\mathcal{I}}$ along with a greedy processing rule presented in Algorithm 1. This rule considers individuals in priority order, and commits to selecting each individual if it is feasible to do so while honoring the commitments to higher-priority individuals.

We choose the name "outcome-based" selection because we believe that it is much easier for the public and policymakers to form opinions about what *outcomes* are acceptable, rather than the *algorithms* that should be used to find these outcomes. In many of the applications above, both priorities and feasible outcomes are clearly specified, but the algorithm for combining these considerations is either unspecified or receives very little attention. For example, Gonczarowski et al. [2019] describe a matching system in which

> *Each PMA [Pre-Military Academy] defines a set of "populations" that it cares about (e.g., based on gender, religiousness, city of origin, belonging to certain minority groups). For each such population the PMA is allowed to define a maximum quota as well as minimum target. In addition to ranking the candidates...*

They go on to note that this information does not necessarily pin down the PMAs selection ("There are several issues that need be specified before this becomes fully formal"), but that "virtually none of the PMAs showed any desire to dig into these issues." Relatedly, Pathak et al. [2022] demonstrate experimentally that although changing the processing order for different reserve categories often has significant consequences, people do not understand this fact.

To illustrate that the feasible selections $\mathcal{F}$ and the priority order $\succ$ do not obviously pin down the final selection, we present the following example.

Example 1. *There are four individuals, with* $1 \succ 2 \succ 3 \succ 4$. *The feasible selections are* $\mathcal{F} = \{\{1,4\},\{2,3\}\}$.

In this case, it is not clear whether $\{1,4\}$ or $\{2,3\}$ should be chosen. However, we believe that $\{1,4\}$ is easier to explain. We formalize this argument as follows. We say that a feasible selection *respects priorities* if, for every individual who is not selected, it is impossible to select this individual without displacing some higher-priority individual who is currently selected. Proposition 1 establishes that there is always a unique feasible selection that respects priorities ({1, 4} in Example 1), and it can be found by the greedy procedure described in Algorithm 1.

In essence, we can offer individuals 2 and 3 the following explanation for why they were not selected: selecting them would require displacing the higher-priority individual 1. No such simple explanation is available if we choose $\{2, 3\}$. In that case, we must explain to individual 1 why they were not chosen despite having the top priority. Our explanation will be complicated by the fact that had the bottom of the ranking been slightly different (with individual 4 ahead of individual 3), we presumably would have selected $\{1, 4\}$. Individual 1 may feel that the precise order of lower-priority agents should be irrelevant, motivating our definition of lower invariance.

1.2 Special Cases: Reserves and Quotas

Having defined our family of outcome-based selection rules and characterized them as uniquely explainable, we show how these rules can capture diversity and affirmative action considerations that arise in practice. We offer two ways to define the feasible selections: *reserves* and *quotas*.

When using reserves, the feasible selections are defined by a set of positions and a compatibility graph G indicating which individuals may be matched to which positions. We consider several ways to interpret the graph G. *Hard reserves* require that compatibility constraints are always strictly enforced. *Soft reserves* allow the selection of incompatible individuals so long as all compatible individuals have already been selected. We formulate two interpretations of this requirement: soft *maximal reserves* require that feasible selections must be a superset of the individuals matched in some maximal matching M of G, while soft *maximum reserves* require that M be a maximum matching. While maximum reserves have been studied before (see for example Sönmez and Yenmez [2022]), to our knowledge we are the first to define maximal reserves.

Proposition 2 establishes that all three interpretations induce a matroid, implying that there exists a feasible selection that priority dominates all others. Proposition 3 shows that finding this selection can be done in polynomial time for maximum and hard reserves, but is NP-hard for maximal reserves. This latter result is interesting, and stems from the fact that in this context, we do not have an independence oracle to tell us whether a given set of applicants is a subset of a feasible selection.

Quotas, meanwhile, are specified by (i) a finite set of traits T; (ii) for each individual a subset of traits that the individual possesses; and (iii) upper and lower bounds (quotas) on the number of selected individuals with each trait. Proposition 4 establishes that for this family of feasible selections: (i) determining the existence of a feasible selection is NP-complete, and (ii) even if the set of feasible selections is nonempty, there may be no feasible selection that priority dominates all others. However, if the traits form a hierarchy, the resulting feasible sets induce a matroid (and this condition is “necessary" in a maximal domain sense). In cases where traits do not form a hierarchy – such as that considered by Gonczarowski et al. [2019] – our outcome based rule could improve upon existing solutions by guaranteeing selection of an undominated feasible outcome.

1.3 Electing the Chilean Constitutional Assembly

In 2021, the country of Chile convened an assembly to propose a new constitution. Members of this assembly were selected by a special election. Each district elected a set of representatives whose parties needed to reflect the district's vote totals. Furthermore, representatives needed to be equally balanced between male and female. These constraints do not induce a matroid, so in general there will not be a selection that priority dominates all others. The algorithm used by Chile begins by ignoring the gender constraints, and then corrects imbalances by replacing the lowest vote-getters from the over-represented gender. Interestingly, we show that this algorithm is equivalent to the outcome-based rule associated with the party and gender constraints (Theorem 2). This result could bolster the legitimacy of the election: each candidate who was not elected can be assured that in order to select them while complying with gender and party constraints, a candidate with more votes would have to be rejected.

2 RELATED WORK

Our paper contributes to a large and growing literature on assignment and matching problems with distributional constraints. There are too many papers to discuss each one in detail, so we divide our discussion into two clusters: those where individuals belong to disjoint groups, and those where affirmative action targets overlapping groups. Within each cluster, we focus our discussion on the most related work.

2.1 Disjoint Types

Most of the early matching work on distributional constraints and affirmative action assumes that individuals can be partitioned into types, and that constraints are specified with respect to the number of selected individuals of each type. We note that our uses of the words "reserve" and "quota" are somewhat different from their typical use in prior work. Specifically, many papers use "quota" to mean *maximum* quotas, and "reserve" to mean *minimum* quotas. By contrast, we use "quotas" to refer to minimum and maximum quotas. In settings where types do not overlap, the distinction between reserves and minimum quotas is of little consequence. However, with overlapping types, reserving spots for certain groups of people induces a very different mathematical structure than setting a minimum quota (for example, the number of reserved positions cannot exceed the number of selected individuals, whereas the sum of minimum quotas can).

Abdulkadiroğlu and Sönmez [2003] and Abdulkadiroğlu et al. [2005] assume students belong to disjoint types, and that there are type-specific maximum quotas. Kojima [2012] considers a setting with only two student types (minority and majority), and shows that implementing upper quotas on majority applicants can end up hurting all minority applicants. Hafalir et al. [2013] also consider a model with two types of students, and argue that minority reserves (minimum quotas) result in more efficient outcomes than majority (maximum) quotas. A related model of Ehlers et al. [2014] attempts to achieve diversity goals by imposing both upper and lower quotas on the number of students of each type. Not surprisingly, if these bounds are treated as hard constraints, a good assignment of students to schools may not exist; so these authors interpret these upper and lower bounds as "soft" constraints that should be satisfied "as much as possible" and design an algorithm to find a fair and non-wasteful assignment of students to schools.

Doğan and Yildiz [2023] and Abdulkadiroglu and Grigoryan [2022] also provide characterization results for prioritized selection problems with constraints. Like our work, these papers assume existence of a baseline priority ranking, in addition to other diversity objectives. However, both papers impose much more structure than we do on these diversity objectives. In particular, both papers assume that reserves and quotas are defined over non-overlapping subsets of applicants, whereas we allow more general feasibility constraints.

Doğan and Yildiz [2023] assume that applicants belong to one of two types, "minority" or "majority", and some of their axioms leverage this structure (i.e. requiring that changes to the type of an applicant can only affect the outcome in specific ways). Their "monotonicity in priority improvements" axiom is implied by (weaker than) the combination of our "monotonicity" and "priority non-bossy" axioms. As a result, if we model their setting using our framework, their axioms identify a set of selection rules, only some of which are outcome-based according to our definition. In particular, an "over and above" implementation of minority reserves satisfies all of their properties, but is priority bossy.

Abdulkadiroglu and Grigoryan [2022] assumes that agents can be partitioned into disjoint types, and each type has a corresponding upper quota and lower quota ("reserve"). This can be seen as a special case of our model with quotas. Because types are disjoint, their feasible outcomes induce a matroid, implying the existence of a single selection that priority dominates all others. They provide an explicit algorithm (the regular reserves rule) for identifying this selection.

Recent work by Pathak et al. [2022] loosely relates to our motivation of explainability. It uses a survey of over 1,000 people to demonstrate that the effects of changing the processing order of different positions are poorly understood. This work helps to reinforce our belief that people are generally better at understanding and forming opinions about final outcomes, rather than the procedures that generate them.

2.2 Overlapping Types

An important shortcoming of these early models is the assumption that agents are partitioned into disjoint types. In practice, affirmative action policies often target groups whose memberships overlap. A number of recent papers consider this case (as do we). Most of these papers focus on a specific application, and describe feasibility constraints, axioms, and algorithms suitable for that domain. For this reason, it can be difficult to translate results from one paper to the next. We instead propose axioms and a greedy algorithm which can be applied regardless of the structure of the feasible sets.

A similar approach is taken by Hafalir et al. [2022]. They consider choice rules that determine a final selection based on (i) a set of applicants, and (ii) a priority ranking for all individuals. Fixing the priority ranking, these rules determine a choice function. Fixing the set of applicants, these rules determine a selection rule (as defined in this paper). The family of rules on which they focus is "Diversity Choice Rules." These rules rely on a function f mapping selections to a score indicating how "diverse" they are. For any set of applicants $\mathcal{I}$, their rule defines the feasible sets to be the most diverse subsets of $\mathcal{I}$ (that is, $\mathcal{F} = \arg\max_{S \subseteq \mathcal{I}} f(S)$), and selects among sets in $\mathcal{F}$ using a greedy processing algorithm. Thus, our outcome-based selection rules can be seen as a special case of their Diversity Choice Rule where the set of applicants $\mathcal{I}$ is fixed and known.

Although their paper considers a generalization of our outcome-based selection rules, our central questions are fairly different. They focus primarily on efficiency: that is, what conditions on f ensure that the feasible set $\mathcal{F}$ induces a matroid? While similar in spirit to our investigation in Section 4, the conditions they provide are based on discrete concavity notions, making them possibly more general but also much more abstract than those in our Propositions 2 and 5. By contrast, our focus is on explainability. We provide two characterization results showing that outcome-based rules are uniquely explainable, regardless of whether the underlying constraint set induces a matroid.

Several papers present results that are closely related to our findings in Section 4. Gonczarowski et al. [2019] study a setting with lower and upper quotas, and identify the importance of having the traits form a hierarchy (which they refer to as 'laminar'). However, in order for their algorithms to be guaranteed to find a priority dominant selection, their Theorems 1 and 2 require further restrictions on the set of traits with a nonzero lower quota. A result from Yokoi [2017], reproduced as our Proposition 5, establishes that these restrictions are unnecessary. When traits do not form a hierarchy, Gonczarowski et al. [2019] propose a computationally simple algorithm that does not respect priorities, while we propose finding the selection that respects priorities (which is NP-hard in the worst case).

Aygün and Turhan [2022] and Sönmez and Yenmez [2022] study affirmative action policies in India. Their characterization results and ours are quite different: whereas they focus on axioms motivated by the Indian Constitution, we attempt to formalize the notion of explainability. This difference leads us in different directions: their vertical reservations are constitutionally mandated, whereas vertical reservations are priority bossy, and thus ruled out by our axioms.

However, the analysis of overlapping horizontal reservations in Sönmez and Yenmez [2022] has a close connection to our work. Their "one to all" convention (mentioned only briefly) is equivalent to our minimum quotas. Their "Meritorious Horizontal Choice Rule" is equivalent to our greedy selection procedure, when feasibility is defined by our soft maximum reserves. Despite their equivalence, the descriptions of these algorithms are quite different. Their Meritorious Choice Rule takes two passes (one to fill reserved spots and another to fill "open" spots), whereas our algorithm operates in a single pass. Both descriptions have value: their is easier to turn into working software, while ours makes it easier to explain to each applicant why they were or were not accepted.

Cembrano et al. [2021] also discuss the Chilean Constitutional Convention, and compare the selection rule used there to many alternatives. However, they do not provide any characterization of this rule. By contrast, we show that this procedure finds the unique selection that respects priorities.

Imamura and Kawase [2022] consider a setting with multiple institutions, in which only certain matchings are feasible. The existence of feasibility constraints across institutions implies that in their case, a standard serial dictatorship may fail to find certain Pareto efficient matchings. By contrast, in our setting, every Pareto efficient allocation is found by greedy processing applied to some priority order. Like our work, they note that greedy selection with feasibility constraints may sometimes require solving NP-complete problems.

3 EXPLAINABLE SELECTION RULES

In a *selection problem* there is a set of applicants $\mathcal{I}$, from which a decision maker has to select a subset $S \subseteq \mathcal{I}$. There is a strict priority order $\succ$ over $\mathcal{I}$. We denote by $\mathbf{P}$ the set of all priority orders over $\mathcal{I}$. A *selection rule* is a mapping $\varphi : \mathbf{P} \to 2^{\mathcal{I}}$ that determines a selection for every priority order. We let $\varphi(\succ)$ denote the selection of rule φ for priority order $\succ$.

3.1 Respecting Priorities

Our first characterization result is for cases where there is an explicitly defined collection of feasible selections $\mathcal{F} \subseteq 2^{\mathcal{I}}$. This may capture capacity constraints, diversity goals or other restrictions imposed by law or by the designer. In this section, our goal is to define a selection rule that (i) always returns a feasible selection and (ii) follows the priority order "as much as possible." To formalize this second objective, we introduce a method to compare feasible selections.

Definition 1. *Selection S'* **priority dominates** *selection S if (and only if) $S' \neq S$ and there exists an injective function ψ from S to S', such that for each $i \in S$, $\psi(i) \succeq i$.*

An equivalent definition of priority domination is as follows: S' priority dominates S if and only if S' selects at least as many "top k priority" individuals for *every* $k \in \mathbb{N}$. If $1 \succ 2 \succ 3 \succ 4$, then $\{1,3\}$ priority dominates $\{1,4\}$, $\{2,3\}$, and $\{2,4\}$. However, neither of the selections $\{1,4\}$ and $\{2,3\}$ priority dominate the other.

Priority domination induces a natural partial order on sets of applicants, which has appeared in other papers under different names. For example, Turhan and Aygun [2023] refer to this as "merit-based domination", while Sönmez and Yenmez [2022] use the name "Gale domination." Gale [1968] simply calls a selection that priority dominates all others "optimal." Abdulkadiroglu and Grigoryan [2022] offer an alternative (stronger) definition of "priority domination".[2]

If our goal is to respect the priority order as much as possible, then at a minimum, we should choose a feasible selection that is not priority dominated. However, this does not identify a unique outcome if there are multiple undominated feasible selections, as in the case of Example 1. To resolve this question, we propose the following criteria.

[2]Their definition (which we will call AG-domination) is that S' AG-dominates S if every agent in $S' \setminus S$ has higher priority than every agent $S \setminus S'$. This definition has the counter-intuitive property of not being transitive: if $1 \succ 2 \succ 3 \succ 4$, then $\{1,3\}$ AG-dominates $\{1,4\}$ and $\{2,3\}$, both of which AG-dominate $\{2,4\}$, but $\{1,3\}$ does not AG-dominate $\{2,4\}$.

ALGORITHM 1: Outcome Based

Input: Set of applicants $\mathcal{I}$, feasible selections $\mathcal{F}$ and priority order $\succ$, with $i_1 \succ i_2 \succ \cdots \succ i_n$.

$S := \emptyset$;

foreach $k = 1, \ldots, n$ **do**

  if $\exists\, S' \in \mathcal{F} : S \cup \{i_k\} \subseteq S'$ **then**

    $S = S \cup \{i_k\}$

  end

end

return S

Definition 2. *A feasible selection $S \in \mathcal{F}$* **respects priorities** *if, for each unselected applicant $i \notin S$, it is impossible to select them without displacing a higher-priority applicant. Formally, there is no feasible selection $S' \in \mathcal{F}$ such that $\{i\} \cup \{j \in S : j \succ i\} \subseteq S'$.*

Respecting priorities is intended to capture the absence of "justified envy" [Abdulkadiroğlu and Sönmez, 2003]. In simple settings with only a global capacity constraint, individual i has justified envy for j if i has higher priority than j, yet j is selected while i is not. In settings with richer constraints, there are several plausible definitions for what it means for i's envy for j to be justified. One might say that i's envy is only justified if replacing j with i maintains feasibility. However, if the feasible sets are $\{1, 2\}$ and $\{3, 4\}$ and the priority order is $1 \succ 2 \succ 3 \succ 4$, then this interpretation would imply that selecting $\{3, 4\}$ does not produce any justified envy. We instead take the position that i's envy for j should be considered justified whenever it is possible to take i without changing the outcome of any individuals with higher priority than i (but perhaps changing the outcome for j and many other lower-priority individuals).

Delacrétaz et al. [2023] model refugee assignment as a matching problem in which each locality has feasibility constraints defined by a multi-dimensional knapsack problem. If we define our feasible sets $\mathcal{F}$ as they do, then our definition of respecting priorities coincides with theirs, which they call "interference-freeness."

In addition, we point out that our definition of respecting offers a "simple" explanation for each individual's outcome. Suppose that an individual i complains that they were not selected, while some lower priority individuals were. If we have chosen a selection that respects priorities, then we can tell i that these lower priority individuals were not the reason that i was not selected. Rather, i was not selected because any feasible selection including i would require the displacement of some individual with a *higher* priority.

Applying this concept to Example 1 with $1 \succ 2 \succ 3 \succ 4$ and $\mathcal{F} = \{\{1, 4\}, \{2, 3\}\}$, the selection $\{1, 4\}$ respects priorities, while the selection $\{2, 3\}$ does not. We interpret this as saying that if we choose $\{2, 3\}$, there is no "simple" explanation for why individual 1 was not selected. The following (straightforward) result establishes that there is always a unique selection that respects priorities.

Proposition 1. *For any feasible set $\mathcal{F}$ and any priority order $\succ$, let $OB_{\mathcal{F}}(\succ)$ denote the outcome of Algorithm 1. Then:*

(i) $OB_{\mathcal{F}}(\succ)$ is not priority dominated by any other feasible selection.
(ii) $OB_{\mathcal{F}}(\succ)$ is the unique feasible selection that respects priorities.

Proof. Fix $\mathcal{F}$ and $\succ$. Let S be any feasible selection other than $OB_{\mathcal{F}}(\succ)$. Let i be the highest priority individual who receives a different outcome under $OB_{\mathcal{F}}(\succ)$ and S. By assumption, S and $OB_{\mathcal{F}}(\succ)$ make identical decisions on individuals with higher priority than i: $\{j : j \succ i\} \cap S = \{j : j \succ i\} \cap OB_{\mathcal{F}}(\succ)$. Note that if $i \in S$, then $i \in OB_{\mathcal{F}}(\succ)$ by definition: when $OB_{\mathcal{F}}(\succ)$ considered individual i, S was a feasible selection containing i and all higher-priority individuals that had already been selected, so $OB_{\mathcal{F}}(\succ)$ must have selected i. Because S and $OB_{\mathcal{F}}(\succ)$ result in different outcomes for i, it must be that $i \in OB_{\mathcal{F}}(\succ)$ but $i \notin S$. This immediately implies that (i) S does not priority dominate $OB_{\mathcal{F}}(\succ)$, and (ii) S does not respect priorities, as $OB_{\mathcal{F}}(\succ)$ is feasible and selects i along with all higher-priority individuals selected by S. □

3.2 A Characterization: Monotonicity, Lower Invariance, and Priority Non-Bossiness

In practice, selection rules need not be defined through a set of feasible outcomes. Often, they are defined by a description of an algorithm for considering and selecting applicants. In such cases, the feasible outcomes are often defined only implicitly. As one example, we consider the cases of "over and above" and "minimum guarantee" reserves [Pathak et al., 2022].

Example 2. *There are four individuals, 1, 2, 3, 4. Individuals 1 and 3 belong to a minority group. We can select at most two individuals, and wish to reserve at least one seat for minorities.*
The "over and above" selection rule φ^{OA} first selects the highest-priority individual, and then selects the highest priority minority that remains. The "minimum guarantee" selection rule φ^{MG} first selects the highest-priority minority, and then selects the highest priority individual that remains.

Notice that each of these rules are defined without explicitly identifying a set of feasible selections. Two natural questions are,

(1) Can an outcome-based selection rule (with suitably chosen $\mathcal{F}$) reproduce the behavior of φ^{OA} and φ^{MG}?
(2) If not, why not? What properties do outcome-based rules have, that other rules do not?

It turns out that φ^{MG} is equivalent to the outcome based rule associated with the feasible selections

$$\mathcal{F} = \{\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{3,4\}\}.$$

(Note that $\{2,4\}$ is never selected, as neither 2 nor 4 is a minority.) Meanwhile, we claim that **no** choice of $\mathcal{F}$ allows $OB_{\mathcal{F}}$ to reproduce the behavior of φ^{OA}. To see this, note that any such $\mathcal{F}$ must contain each of the five pairs listed above, as each of those pairs can be produced by some priority ordering. But when the priority order is $1 \succ 2 \succ 3 \succ 4$, φ^{OA} selects $\{1,3\}$, even though this selection is priority dominated by $\{1,2\}$.

Thus, while choosing $\mathcal{F}$ gives the designer substantial flexibility (as we explore in Section 4), restricting to the family of outcome based rules does rule out certain selection rules commonly used in practice. One might therefore ask, why should a designer use a outcome based rule? What advantages do these rules offer?

We show that outcome based rules satisfy three desirable properties, and are in fact the *only* selection rules to satisfy all three properties. To define each of these properties, we introduce new notation. Given a priority order $\succ$ and a subset of individuals $S \subset \mathcal{I}$, we let $\succ_{-S}$ be the restriction of $\succ$ to $\mathcal{I} \backslash S$. In other words, for $i, j \in \mathcal{I} \backslash S$,

$$i \succ_{-S} j \iff i \succ j.$$

When $S = \{i\}$ is a singleton, we write $\succ_{-i}$ in place of $\succ_{-\{i\}}$.

Our first property is *monotonicity*, which states that improvements to an individual's priority should weakly improve their outcome.

Definition 3. *A selection rule φ is* **monotone** *if for any applicant $i \in \mathcal{I}$, and for every $\succ, \succ' \in \mathbf{P}$ such that (i) $\succ_{-i} = \succ'_{-i}$ and (ii) the priority of applicant i is higher in $\succ'$ than in $\succ$, then*

$$i \in \varphi(\succ) \Rightarrow i \in \varphi(\succ').$$

Monotonicity is a very natural property, which has appeared in many other papers under different names. Balinski and Sönmez [1999] refer to this as "respecting improvements", and this terminology has been adopted by many subsequent papers, including Sönmez and Switzer [2013], Kominers and Sönmez [2016], Kominers [2020], and Biró et al. [2023].

The second property is *lower-invariance*, which says that whether an individual is selected should depend only on the priority of *higher*-priority individuals, and not on the exact priority of lower-priority individuals.

Definition 4. *A selection rule φ is* **lower invariant** *if for any applicant $i \in \mathcal{I}$, and every $\succ, \succ' \in \mathbf{P}$ such that $\succ_{-\{j:i\succ j\}} = \succ'_{-\{j:i\succ' j\}}$,*

$$i \in \varphi(\succ) \Leftrightarrow i \in \varphi(\succ').$$

Properties similar to lower invariance have been studied in the matching literature. For instance, Ehlers and Klaus [2014] study an axiom called *truncation invariance*, which says that if any student truncates their preference report at the school they are matched to, it should not affect the final outcome. Hashimoto et al. [2014] proposed a similar axiom for randomized mechanisms termed *weak invariance*: they require an agent's share of an object to be independent of their ranking of less-preferred objects. In our setting, there are no preferences (as there is a single institution), and our lower invariance property concerns changes to the *priority order*. Lower-invariance is desirable because of its natural connection to explainability: if individual i is not selected, it is possible to provide i an explanation without referencing the priorities of lower-ranking applicants.

The final property is *priority non-bossiness*, which says that a change to an individual's priority which does not affect whether that individual is selected should also not affect who else is selected.

Definition 5. *A selection rule φ is* **priority non-bossy** *if for any applicant $i \in \mathcal{I}$, and for every $\succ, \succ' \in \mathbf{P}$ such that $\succ_{-i} = \succ'_{-i}$,*

$$\{i\} \cap \varphi(\succ) = \{i\} \cap \varphi(\succ') \Rightarrow \varphi(\succ) = \varphi(\succ').$$

We name this property "priority non-bossiness" because it bears some resemblance to the definition of non-bossiness given originally by Satterthwaite and Sonnenschein [1981]. That definition states that if a change to an agent's reported preference does not affect the allocation of that agent, then it should not affect any other agent's allocation. Priority non-bossiness imposes an analogous requirement for changes to an agent's priority (rather than their reported preferences).

We view the priority non-bossiness as capturing *consistency* of the selection rule. Suppose individual i's outcome is the same under two priority orderings that differ only in i's position in the priority ordering. If i is selected in *neither*, then it is natural to require the remaining selected individuals to be the same in both cases, as the relative ordering of the other individuals is identical in the two priority orderings; this is similarly natural if i is selected in *both*. This is precisely what priority non-bossiness requires. Indeed, this property is reminiscent of the commonly used axiom of *consistency* in resource allocation problems involving a varying population of individuals, which states that "An allocation rule is 'consistent' if the recommendation it makes for each problem 'agrees' with the recommendation it makes for each associated reduced problem, obtained by imagining some agents leaving with their assignments" [Thomson, 2012].

We are now ready to present our characterization of the family of outcome based rules.

Theorem 1.
For any $\mathcal{F} \subseteq 2^{\mathcal{I}}$, the selection rule $OB_{\mathcal{F}}$ is monotone, priority non-bossy and lower invariant.
If φ is a selection rule that is monotone, priority non-bossy and lower invariant, then there exists $\mathcal{F} \subseteq 2^{\mathcal{I}}$ such that $\varphi(\succ) = OB_{\mathcal{F}}(\succ)$ for every $\succ \in \mathbf{P}$.

Theorem 1 implies that any selection rule either has an outcome-based description, or fails to satisfy at least one of our three axioms.

Appendix A.2 proves that our axioms are independent by providing examples of rules that fail to satisfy one axiom while satisfying the other two. For example, the over and above rule φ^{OA} as defined in Example 2 is monotonic and lower invariant, but priority bossy. Meanwhile, optimization-based approaches that assign higher values to high-priority individuals and try to maximize total value are monotonic and priority non-bossy, but not lower invariant.

4 SPECIAL CASES OF INTEREST: RESERVES AND QUOTAS

Outcome based selection rules can implement many different policies, depending on how the feasible sets $\mathcal{F}$ are defined. In this section, we consider several procedures that have been proposed in the literature and/or used in practice. We focus on two types of feasibility constraints, which we call *reserves* and *quotas*. There are two key contributions of this section.

First, we connect our ideas to the existing literature on reserves and matching with affirmative action considerations. Some of these procedures – including algorithms used to implement "vertical reservations" in India and "slot specific priority" algorithms introduced by Kominers and Sönmez [2016] and used for assigning students to schools in Chile – are not outcome based. However, others are. By demonstrating that these procedures have an outcome-based description, we provide a new way to explain the outcomes from these procedures to participants.

Our second contribution in this section is to introduce *maximal reserves*, which have not been studied in prior work. As we discuss below, the outcome-based selection rule associated with maximal reserves will satisfy the same feasibility criterion as the slot-specific priority algorithms do, but will choose a selection that priority dominates the selections made by these algorithms.

We note that our use of the terms "quotas" and "reserves" departs from their usage in prior work. Hafalir et al. [2013] and several follow-up papers use the term "reserves" for minimum targets on specific groups of applicants, and "quotas" for maximum targets. By contrast, we use the word "quotas" to refer to both minimum *and* maximum targets. The key distinction between a minimum quota and a reserve is that individuals count towards quotas of all groups to which they belong, while they can claim a seat reserved for any one of these groups. Sönmez and Yenmez [2022] also mention this distinction, and refer to our reserves as "one to one" reserves and our minimum quotas as "one to all" reserves.

4.1 Background On Matroids

In general, choosing the selection that respects priorities may come at a significant cost. For example, suppose that there are only two feasible selections: $\{1\}$, and $\mathcal{I}\backslash\{1\}$. Then if individual 1 is at the top of the priority order, respecting priorities requires that we select this individual, and reject all others. However, this example is somewhat contrived and unlikely to arise in practice.

In essence, $OB_{\mathcal{F}}$ greedily selects individuals, so long as doing so would not violate feasibility. Greedy algorithms are well-known to solve optimization problems when the underlying constraints define a matroid.[3] We now restate this result in our setting.

Definition 6. *The feasible selections $\mathcal{F}$* **induce a matroid** *if for any $I_1, I_2 \subseteq \mathcal{I}$ such that $I_1 \subseteq S_1 \in \mathcal{F}$, $I_2 \subseteq S_2 \in \mathcal{F}$, and $|I_1| > |I_2|$, there exists $i \in I_1 \backslash I_2$ and $S_3 \in \mathcal{F}$ such that $I_2 \cup \{i\} \subseteq S_3$.*

For those already familiar with matroids in other contexts, this definition says that our independent sets are all *subsets* of feasible selections.

Lemma 1. *If $\mathcal{F}$ induces a matroid, then for any priority order $>$, $OB_{\mathcal{F}}(>)$ priority dominates any other feasible assignment.*

[3]See https://en.wikipedia.org/wiki/Matroid#Greedy_algorithm.

While Lemma 1 gives a complete answer to when $OB_{\mathcal{F}}$ will produce a priority dominant selection, verifying whether $\mathcal{F}$ induces a matroid can be non-trivial. Therefore, we consider two structures for $\mathcal{F}$ from practice: reserves and quotas.

4.2 Reserved Positions

We now study several families of feasible selections which we collectively refer to as *reserved positions*. These families are described by a bipartite graph $G = (\mathcal{I},\ \mathcal{P},\ E)$, where each node of the left side correspond to an individual $i \in \mathcal{I}$, each node on the right side corresponds to a position $p \in \mathcal{P}$, and there is an edge $(i, p) \in E \subseteq \mathcal{I} \times \mathcal{P}$ only if individual i is eligible for position p. We refer to $G = (\mathcal{I},\ \mathcal{P},\ E)$ as the *compatibility graph*. A *matching* M of G is a subset of edges such that every node is incident to at most one edge in M. For any matching M of G, we denote by $\mathcal{I}_M$ the set of applicants that are incident to an edge in M:

$$\mathcal{I}_M = \{i \in \mathcal{I} :\ \text{exists a position } p \in \mathcal{P} \text{ such that } (i, p) \in M\}. \tag{1}$$

In some environments, compatibility constraints must be strictly enforced. We model these situations using what we call *hard reserves*. Formally, given a compatibility graph $G = (\mathcal{I},\ \mathcal{P},\ E)$, the feasible selections when using hard reserves are

$$\mathcal{F}^{hard}(G) = \{S \subseteq \mathcal{I} :\ \text{exists a matching } M \text{ of } G \text{ such that } \mathcal{I}_M = S\}. \tag{2}$$

Under hard reserves, a position may go unfilled if there is no unselected individual who is eligible for it. In many applications, it may be preferable to reduce waste by allowing someone who would otherwise not be eligible for this position to be assigned to it (Aygün and Turhan [2022] refer to this as "de-reserving" this position). To model such cases, we introduce two types of "soft" reserves: *maximum reserves* and *maximal reserves*. We define

$$\mathcal{F}^{maximum}(G) = \{S \subseteq \mathcal{I} : |S| \leq |\mathcal{P}| \text{ and } \exists \text{ a } \textit{maximum} \text{ matching } M \text{ of } G \text{ such that } \mathcal{I}_M \subseteq S\}. \tag{3}$$

$$\mathcal{F}^{maximal}(G) = \{S \subseteq \mathcal{I} : |S| \leq |\mathcal{P}| \text{ and } \exists \text{ a } \textit{maximal} \text{ matching } M \text{ of } G \text{ such that } \mathcal{I}_M \subseteq S\}. \tag{4}$$

Both of these definitions allow applicants to be matched to positions for which they are not eligible, so long as everyone who is eligible for that position is also selected. Both definitions also will ensure that if a total of r positions are reserved for a particular group of applicants $\mathcal{R} \subseteq \mathcal{I}$, then in every feasible selection, at least r of these applicants will be selected (or all of them, if $|\mathcal{R}| < r$).

The distinction between hard, maximum, and maximal reserves is illustrated in Figure 2. Observe that $\mathcal{F}^{maximum}(G) \subseteq \mathcal{F}^{maximal}(G)$ by definition, as every maximum matching is also maximal. Furthermore, any set in $\mathcal{F}^{hard}(G)$ is also a subset of some set in $\mathcal{F}^{maximum}(G)$. Therefore, for a given graph G, one can think of hard reserves as providing the "strictest" interpretation of the associated compatibility requirements, while maximum and maximal reserves are more permissive, with maximal reserves allowing the most flexibility.

Proposition 2. *For any $G = (\mathcal{I}, \mathcal{P}, E)$, the feasible selections $\mathcal{F}^{hard}(G)$, $\mathcal{F}^{maximum}(G)$ and $\mathcal{F}^{maximal}(G)$ defined by* (2), (3) *and* (4) *each induce a matroid.*

From Lemma 1, it follows that the selection from using maximal reserves will priority dominate the selection from using maximum reserves, which will priority dominate the selection from using hard reserves. However, in some cases, the selection obtained from maximal reserves may not be seen as complying with the spirit of the constraints encoded by G. Depending on the relative importance of compatibility constraints and the priority order, the designer can choose any of these three interpretations.

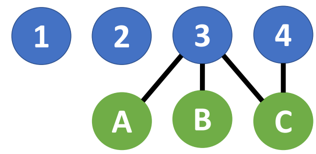

Fig. 2. An example graph G illustrating the difference between hard, soft maximum, and soft maximal reserves. Individuals are in blue and positions are in green. Hard reserves preclude matching individuals 1 and 2, and the final selection will be $\{3, 4\}$ for any priority order. Maximum reserves require that individuals 3 and 4 are selected, but will also select whichever of 1 and 2 have higher priority. Maximal reserves, meanwhile, allow the possibility of selecting $\{1, 2, 3\}$, as this can be justified by assigning individual 3 to position C (since $(3, C)$ is a maximal matching in G).

We now comment on connections between this result and others in the matching literature. Hard reserves induce the well-known "matching matroid" or "transversal matroid"[4].

The horizontal reserves in Sönmez and Yenmez [2022] correspond to our maximum reserves: their non-wastefulness condition implies that reserves are soft, and their condition of "maximally complying with reservations" further specifies maximum (rather than maximal) reserves. Their Proposition 2 states that the "meritorious horizontal choice rule" priority dominates all other feasible selections, which is analogous to our finding that maximum reserves induce a matroid.

Although their algorithm and ours lead to identical outcomes in this context, their descriptions are quite different. The meritorious horizontal choice rule proceeds in two passes: the first tries to fill as many reserve spots as possible with eligible applicants, and the second fills any remaining spots with the highest priority remaining applicants. Explaining to an applicant why he or she was not selected requires explaining the detailed working of the algorithm. We show that the same outcome can be found by an algorithm that makes a single pass. Our outcome-based description offers the following justification for every rejected applicant: "Our chosen selection fills k reserve slots. We determined that in order to select you, we would have to either (i) fill fewer reserve slots, or (ii) displace an applicant with higher priority than you."

To our knowledge, our work is the first to define and study maximal reserves. Some might say that this definition is "too permissive:" in the example in Figure 2, choosing applicants 1 and 2 over applicant 4 arguably goes against the spirit of the reserve policy. However, slot-specific priority algorithms as defined by Kominers and Sönmez [2016] often select matchings that are feasible under maximal reserves but *not* maximum reserves: in the example in Figure 2, if position C chooses first and $1 \succ 2 \succ 3 \succ 4$, then position C will choose individual 3 and positions A and B will choose individuals 1 and 2, resulting in a selection that is not feasible according to maximum reserves. Related algorithms have been proposed for Brazilian University admissions [Aygun and Bo, 2019] and are used for Chilean school assignment [Correa et al., 2021].

If these algorithms are viewed as acceptable for these applications, it indicates that perhaps maximal reserve feasibility is sufficient in practice. In this case, we believe that instead of using an ad-hoc selection procedure, it would be better to explicitly identify the set of feasible assignments and find a feasible selection that priority dominates all others (which exists by Proposition 2).

One barrier is that the problem of finding this selection is in general NP hard, which we show by reduction from the minimum maximal matching problem. By contrast, the priority-respecting assignment can be found in polynomial time for maximum and hard reserves.

[4]https://en.wikipedia.org/wiki/Matroid#Matroids_from_graph_theory.

Proposition 3. *The problem of evaluating $OB_{\mathcal{F}^{maximal}(G)}$ is NP-complete.*
For any graph G, the rules $OB_{\mathcal{F}^{maximum}(G)}$ and $OB_{\mathcal{F}^{hard}(G)}$ can be evaluated in polynomial time.

4.3 Quotas

In the second family of constraints we study, there are multiple groups of applicants. For each group, there is an upper and lower limit on the number of its members that can be selected. Formally, there is a finite set of traits T. For each trait $t \in T$, there is a set of applicants with the trait $\mathcal{I}_t \subseteq \mathcal{I}$, and an upper and lower quota $u_t, \ell_t \in \mathbb{N}$ on the number of selected applicants with trait t. We use $\mathcal{I}_T$ as shorthand for $\{I_t\}_{t \in T}$. Then, a selection is feasible if it satisfies both quotas for every trait. That is,

$$\mathcal{F}(\mathcal{I}_T, \boldsymbol{u}, \boldsymbol{\ell}) = \{S \subseteq \mathcal{I} : \text{ for every } t \in T,\ \ell_t \le |S \cap \mathcal{I}_t| \le u_t\}. \tag{5}$$

In contrast with the reserved positions defined in Section 4.2, the problem of determining whether there exists a feasible selection is NP-complete. Furthermore, even when a feasible selection exists, there may be no feasible selection that priority dominates all others.

Proposition 4. *Determining whether $\mathcal{F}(\mathcal{I}_T, \boldsymbol{u}, \boldsymbol{\ell})$ defined in* (5) *is empty is NP-complete.*
If $\mathcal{F}(\mathcal{I}_T, \boldsymbol{u}, \boldsymbol{\ell})$ is non-empty, there may be no feasible selection that priority dominates all others.

Our proof in the appendix establishes NP-completeness result even in special cases in which there is just a single global lower quota or a single global upper quota. These proofs reduce from the independent set problem and the set cover problem, respectively. Meanwhile, Example 3 shows that in general, there may be no feasible selection that priority dominates all others.

Example 3. *Consider an instance with three applicants $\mathcal{I} = \{1, 2, 3\}$ and two traits $T = \{A, B\}$. Let $\mathcal{I}_A = \{1, 2\}$, $\mathcal{I}_B = \{1, 3\}$ and $u_A = u_B = 1 = \ell_A = \ell_B$. In this instance, there are only two feasible selections: $\mathcal{F}(\mathcal{I}_T, \boldsymbol{u}, \boldsymbol{\ell}) = \{\{1\}, \{2, 3\}\}$. If $1 \succ 2 \succ 3$, then neither feasible selection priority dominates the other.*

4.3.1 Hierarchies. As Example 3 illustrates, under quotas there might be tension between respecting priorities and other goals, such as maximizing the number of selected individuals. We now provide conditions under which the feasible selection that respects priorities priority dominates all others.

Definition 7. *The traits $\mathcal{I}_T = \{I_t\}_{t \in T}$ form a* **hierarchy** *if for every $t, t' \in T$, either $\mathcal{I}_t$ and $\mathcal{I}_{t'}$ are disjoint ($\mathcal{I}_t \cap \mathcal{I}_{t'} = \emptyset$) or one contains the other ($\mathcal{I}_t \subset \mathcal{I}_{t'}$ or $\mathcal{I}_{t'} \subset \mathcal{I}_t$).*

This condition has been referred to as 'laminarity' by previous work, including Huang [2010], Fleiner and Kamiyama [2016], Yokoi [2017], and Gonczarowski et al. [2019], and its importance is widely recognized. In fact, this structure is both sufficient and "necessary" (in a maximal domain sense) for the existence of a feasible selection that priority dominates all others.

Proposition 5.
If the collection of sets $\mathcal{I}_T = \{\mathcal{I}_t\}_{t \in T}$ is a hierarchy, then for any upper and lower quotas $\boldsymbol{u}, \boldsymbol{\ell} \in \mathbb{N}^T$ such that the set of feasible selections is non-empty, the feasible sets $\mathcal{F}(\mathcal{I}_T, \boldsymbol{u}, \boldsymbol{\ell})$ induce a matroid. Therefore, the selection that respects priorities also priority dominates all other feasible selections.
If the collection of sets $\mathcal{I}_T = \{\mathcal{I}_t\}_{t \in T}$ is not a hierarchy, then there exist upper and lower quotas $\boldsymbol{u}, \boldsymbol{\ell} \in \mathbb{N}^T$ and a priority order $\succ$ such that no feasible selection priority dominates all others.

The first part (sufficiency) of Proposition 5 is proven as Lemma 3 in Yokoi [2017]. The second part (necessity) is proved in Appendix B.2.

This result implies that if quota categories form a hierarchy, the use of the greedy processing rule should be uncontroversial. However, in practice, categories may not form a hierarchy. Gonczarowski et al. [2019] describe one such example, and provide several heuristic algorithms, which may fail

Fig. 3. There are three parties (A, B, C), with four candidates from each party (2 men and 2 women). Votes for each candidate are shown in black, and the candidate's priority order (with 1 being highest priority) is shown in green. There are a total of 4 seats. Based on vote totals, $E_A = 2$, $E_B = E_C = 1$. In addition, the district must elect two men and two women. The Chilean algorithm selects candidates $\{1, 2, 3, 9\}$. This is undominated by any other feasible selection, but selects the low-priority candidate 9. Other undominated selections in this example are $\{1, 2, 4, 7\}$ and $\{2, 3, 4, 6\}$.

A		B		C	
2	130	1	140	3	124
4	118	6	85	7	80
5	111	8	33	10	18
9	23	11	14	12	14

to find a feasible selection. We believe that the outcome-based rule associated with their quota constraints could be a better solution. As evidence for this, our last section considers another application with quota categories that do not form a hierarchy (the Chilean Constitutional Assembly Election), and establishes that the election procedure that they used is outcome-based.

5 CHILEAN CONSTITUTIONAL ASSEMBLY ELECTION

In what follows, we describe the mechanism used in Chile to elect the members of the 2021 Constitutional Assembly. Each district held a separate election. In addition, there was an election to select 17 members of indigenous communities, which was conducted separately from the district elections. Below, we focus on what happened within each district.

The algorithm used to select candidates in Chile works as follows.[5]

ALGORITHM 2: Chilean Constitutional Mechanism

Apportionment. Determine the number of votes earned by candidates from each party. Based on this, determine the total number of elected representatives E_i from each party i. (Roughly speaking, a party with 30% of the votes will win 30% of the seats. For details on rounding issues, see Cembrano et al. [2021].)

Initial Selection. Tentatively select the E_i highest priority candidates from each party i. Let M and W denote the total number of men and women (across all parties) selected in this step.

Swap. While the number of elected men and women differ by more than one ($|M - W| > 1$), take the lowest vote-getter among the "majority" gender electees, and replace this candidate with the highest vote-getter (who is not currently selected) in the same party of the opposite gender. (If no opposite gender candidate exists in the same party, skip past this party, and find the lowest vote-getter among the majority gender electees belonging to another party).

The goals of this algorithm are to ensure that each party gets its "fair share" of representatives, and that the number of men and women elected from the district is as equal as possible.

The main result of this section is Theorem 2, which establishes that the rule implemented in Chile is equivalent to the outcome based selection rule in which priority is determined by vote totals (candidates who received more votes have higher priority), and the feasible sets $\mathcal{F}_{chile}$ are described using quotas. In particular, $\mathcal{F}_{chile}$ is the set of selections in which (i) exactly E_i candidates are elected from each party, and (ii) the number of elected candidates from each gender is at least $\lfloor N/2 \rfloor$ and at most $\lceil N/2 \rceil$, where N is the number of seats in the district.

THEOREM 2. *The selection made by Algorithm 2 and $OB_{\mathcal{F}_{chile}}(\succ)$ coincide.*

Note that in this case, the quota categories that define $\mathcal{F}_{chile}$ do not form a hierarchy: the set of candidates of a particular gender and candidates from a particular party overlap, but neither contains the other. In other words, in this setting, we should not expect one feasible selection to priority

[5]We slightly simplify reality in this description. In reality, each candidate belongs to a list, and each list is composed of different parties or independent candidates. Incorporating both lists and parties adds some complexity to the actual algorithm, which we do not delve into here. Instead, we describe a simplified procedure, which corresponds to what the algorithm would do if all members of a list belonged to the same party.

dominate all others. This is illustrated in Figure 3. In that example, the Chilean algorithm selects candidate 9, despite the fact that this candidate received only 23 votes. Unsuccessful candidates with more than 23 votes might demand an explanation for why candidate 9 was elected.

Theorem 2 provides a new way to explain the outcome of the election: anybody who was not elected could not have been elected without displacing a candidate who received more votes (or causing a violation of the diversity constraints). This fact is not obvious from the original description of the Chilean algorithm. We take the fact that Chile chose this rule – despite the potential downside that it may elect some candidates with very low vote totals – as (admittedly incomplete) evidence that even in settings where the feasible selections do not induce a matroid and the greedy processing rule does not select a priority dominating selection, outcome based rules may be appealing solutions due to their explainability.

REFERENCES

Atila Abdulkadiroglu and Aram Grigoryan. 2022. Priority-Based Assignment with Reserves and Quotas. (2022).

Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. 2005. The new york city high school match. *American Economic Review* 95, 2 (2005), 364–367.

Atila Abdulkadiroğlu and Tayfun Sönmez. 2003. School choice: A mechanism design approach. *American economic review* 93, 3 (2003), 729–747.

Orhan Aygun and Inacio Bo. 2019. College admission with multidimensional privileges: The Brazilian affirmative action case. *American Economic Journal: Microeconomics* (2019).

Orhan Aygün and Bertan Turhan. 2022. How to De-Reserve Reserves: Admissions to Technical Colleges in India. *Management Science* (2022).

Michel Balinski and Tayfun Sönmez. 1999. A Tale of Two Mechanisms: Student Placement. *Journal of Economic Theory* 84, 1 (1999), 73–94. https://doi.org/10.1006/jeth.1998.2469

Surender Baswana, Partha Pratim Chakrabarti, Sharat Chandran, Yashodhan Kanoria, and Utkarsh Patange. 2019. Centralized admissions for engineering colleges in India. *INFORMS Journal on Applied Analytics* 49, 5 (2019), 338–354.

Péter Biró, Flip Klijn, Xenia Klimentova, and Ana Viana. 2023. Shapley–Scarf Housing Markets: Respecting Improvement, Integer Programming, and Kidney Exchange. *Mathematics of Operations Research* 0, 0 (2023), null. https://doi.org/10.1287/moor.2022.0092 arXiv:https://doi.org/10.1287/moor.2022.0092

Javier Cembrano, Jose Correa, Gonzalo Diaz, and Victor Verdugo. 2021. Proportional apportionment: A case study from the Chilean Constitutional Convention. *Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO)* (2021).

Jose Correa, Natalie Epstein, Rafael Epstein, Juan Escobar, Ignacio Rios, Nicolás Aramayo, Bastian Bahamondes, Carlos Bonet, Martin Castillo, Andres Cristi, et al. 2021. School choice in Chile. *Operations Research* (2021).

David Delacrétaz, Scott Duke Kominers, and Alexander Teytelboym. 2023. Matching mechanisms for refugee resettlement. *American Economic Review* 113, 10 (2023), 2689–2717.

Battal Doğan and Kemal Yildiz. 2023. Choice with affirmative action. *Management Science* 69, 4 (2023), 2284–2296.

A. L. Dulmage and N. S. Mendelsohn. 1958. Coverings of Bipartite Graphs. *Canadian Journal of Mathematics* 10 (1958), 517–534. https://doi.org/10.4153/CJM-1958-052-0

Lars Ehlers, 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, C (2014), 648–683.

Lars Ehlers and Bettina Klaus. 2014. Strategy-Proofness Makes the Difference: Deferred-Acceptance with Responsive Priorities. *Math. Oper. Res.* 39, 4 (nov 2014), 949–966. https://doi.org/10.1287/moor.2014.0662

Yuri Faenza, Swati Gupta, and Zhang Xuan. 2022. Discovering Opportunities in New York City's Discovery Program: an Analysis of Affirmative Action Mechanisms. (2022).

Tamás Fleiner and Naoyuki Kamiyama. 2016. A matroid approach to stable matchings with lower quotas. *Mathematics of Operations Research* 41, 2 (2016), 734–744.

David Gale. 1968. Optimal assignments in an ordered set: an application of matroid theory. *Journal of Combinatorial Theory* 4, 2 (1968), 176–180.

Yannai A Gonczarowski, Noam Nisan, Lior Kovalio, and Assaf Romm. 2019. Matching for the Israeli "Mechinot" Gap-Year Programs: Handling Rich Diversity Requirements. *Proceedings of the ACM Conference on Economics and Computation* (2019).

Isa E Hafalir, Fuhito Kojima, M Bumin Yenmez, and Koji Yokote. 2022. Design on Matroids: Diversity vs. Meritocracy. (2022).

Isa E Hafalir, M Bumin Yenmez, and Muhammed A Yildirim. 2013. Effective affirmative action in school choice. *Theoretical Economics* 8, 2 (2013), 325–363.

Tadashi Hashimoto, Daisuke Hirata, Onur Kesten, Morimitsu Kurino, and M. Utku Ünver. 2014. Two axiomatic approaches to the probabilistic serial mechanism. *Theoretical Economics* 9, 1 (2014), 253–277. https://doi.org/10.3982/TE1010 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE1010

Chien-Chung Huang. 2010. Classified stable matching. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*. SIAM, 1235–1253.

Kenzo Imamura and Yasushi Kawase. 2022. Efficient matching under general constraints. (2022).

Fuhito Kojima. 2012. School choice: Impossibilities for affirmative action. *Games and Economic Behavior* 75, 2 (2012), 685–693.

Scott Kominers. 2020. Respect for Improvements and Comparative Statics in Matching Markets. *Working Paper* 0, 0 (2020), null.

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

L. Lovász and M.D. Plummer. 2009. *Matching Theory*. AMS Chelsea Pub. https://books.google.de/books?id=OaoJBAAAQBAJ

J. Oxley. 2011. *Matroid Theory*. OUP Oxford. https://books.google.com/books?id=EckCyC3e_RgC

Parag A. Pathak, Alex Rees-Jones, and Tayfun Sonmez. 2022. Reversing Reserves. *Management Science* (2022).

Mark A Satterthwaite and Hugo Sonnenschein. 1981. Strategy-proof allocation mechanisms at differentiable points. *The Review of Economic Studies* 48, 4 (1981), 587–597.

Tayfun Sönmez and Tobias B Switzer. 2013. Matching with (branch-of-choice) contracts at the United States Military Academy. *Econometrica* 81, 2 (2013), 451–488.

Tayfun Sönmez and M. Bumin Yenmez. 2022. Affirmative Action in India via Vertical, Horizontal, and Overlapping Reservations. *Econometrica* (2022).

William Thomson. 2012. On The Axiomatics Of Resource Allocation: Interpreting the Consistency Principle. *Economics and Philosophy* 28, 3 (2012), 385–421. https://doi.org/10.1017/S0266267112000296

Bertan Turhan and Orhan Aygun. 2023. *Affirmative Action in India: Restricted Strategy Space, Complex Constraints, and Direct Mechanism Design*. Technical Report. Iowa State University, Department of Economics.

Mihalis Yannakakis and Fanica Gavril. 1980. Edge dominating sets in graphs. *SIAM J. Appl. Math.* 38, 3 (1980), 364–372.

Yu Yokoi. 2017. A generalized polymatroid approach to stable matchings with lower quotas. *Mathematics of Operations Research* 42, 1 (2017), 238–255.

A PROOFS FROM SECTION 3

A.1 Proof of Theorem 1

We divide the proof into two parts: Lemma 2 establishes that for any collection of feasible selections $\mathcal{F} \subseteq 2^{\mathcal{I}}$, the rule $OB_{\mathcal{F}}$ is monotonic, priority non-bossy and lower invariant. Lemma 3 shows that if a selection rule satisfies these three properties, then it must be equivalent to $OB_{\mathcal{F}}$ for some suitably chosen $\mathcal{F}$.

Lemma 2. *For any collection $\mathcal{F} \subseteq 2^{\mathcal{I}}$, the selection rule $OB_{\mathcal{F}}$ is monotone, priority non-bossy and lower invariant.*

Proof. It is clear from the definition of $OB_{\mathcal{F}}$ that it is monotone and lower invariant. Priority non-bossiness needs only to be shown for transpositions of adjacent agents in the preference order (as all changes to the position of individual i can be expressed as a sequence of transpositions). Suppose that i and j appear consecutively in the priority order, and let S be the set of applicants selected before these individuals are considered. We analyze the following cases.

- It is feasible (given the prior commitments to agents in S) to select both i and j.
  In this case, swapping positions has no effect for either agent: both will be selected.
- It is feasible to select either i or j, but not both.
  Swapping positions changes the outcome for both i and j.
- It is feasible to select i, but infeasible to select j.
  Swapping positions has no effect for either agent: i will be selected but j will not.
- It is feasible to select j, but infeasible to select i.
  Swapping positions has no effect for either agent: j will be selected but i will not.
- It is infeasible to select either i or j.
  Swapping positions has no effect for either agent: neither will be selected.

Now suppose that the swap does not affect agent i. By the case analysis above, it also must not affect agent j. Therefore, after $OB_{\mathcal{F}}$ has processed i and j, it has made the same set of commitments regardless of the order in which they appear. By the definition of $OB_{\mathcal{F}}$, this implies that the order in which i and j appear does not affect the decisions made for all subsequent agents, so $OB_{\mathcal{F}}$ is priority non-bossy. □

Lemma 3. *Let φ be a selection rule that is monotone, priority non-bossy and lower invariant. Then there exists $\mathcal{F} \subseteq 2^{\mathcal{I}}$ such that*

$$\varphi(\succ) = OB_{\mathcal{F}}(\succ) \text{ for any } \succ \in \mathbf{P}. \tag{6}$$

Proof. Define $\mathcal{F} = \{\psi(\succ)\}_{\succ \in \mathbf{P}}$ to be the image of ψ. Fix a priority ordering $\succ$, and without loss of generality, label agents so that agent 1 has highest priority, followed by 2, etc. Fix some k, and let $S = \{1, \ldots, k-1\} \cap \psi(\succ)$ be the subset of the first $k-1$ agents who are selected by ψ. To prove the result, it suffices to show that if ψ and $OB_{\mathcal{F}}$ make the same decision for agents $1, 2, ..., k-1$, then they also make the same decision for agent k.

We first show that $k \in \psi(\succ)$ implies $k \in OB_{\mathcal{F}}(\succ)$. If k is selected by ψ, then $\psi(\succ)$ is a feasible selection containing $\{k\} \cup S$, and therefore $OB_{\mathcal{F}}$ will select k.

The more interesting direction is showing that $k \in OB_{\mathcal{F}}(\succ)$ implies $k \in \psi(\succ)$. By definition of $\mathcal{F}$, if $k \in OB_{\mathcal{F}}(\succ)$ then there must exist a priority ordering $\succ_0$ such that $\{k\} \cup S \subseteq \psi(\succ_0)$. We will construct a sequence of priority orderings $\{\succ_i\}_{i=1}^{k}$ as follows: $\succ_i$ is constructed from $\succ_{i-1}$ by taking the i^{th} ranked individual (according to $\succ$), and moving that individual into position i. We claim that for $1 \le i \le k$, we have the following:

A) $\succ_i$ and $\succ$ produce the same outcome for individuals $1, 2, \ldots, k-1$.

B) Individual k is selected when the priority order is $\succ_i$.

These clearly hold for $i = 0$: by definition, $\psi(\succ_0)$ contains S, and if it contained any other agent in $\{1, \ldots, k-1\}$, then $OB_{\mathcal{F}}(\succ)$ would also include this agent, contradicting that ψ and $OB_{\mathcal{F}}$ agree on the first $k-1$ individuals.

Now suppose that these assumptions hold for $i < k - 1$. Note that:

I. $\succ_{i+1}$ produces the same outcome as $\succ$ for agent $i + 1$ (the first $i + 1$ positions of $\succ_i$ and $\succ$ are identical by construction, so this follows from lower invariance of ψ).

II. $\succ_i$ and $\succ_{i+1}$ produce the same outcome for *all* agents (by step I and inductive assumption A, they produce the same outcome for $i + 1$, and ψ is priority non-bossy).

In particular, this implies that the inductive assumptions hold for $i + 1$.

This inductive logic implies that agent $k \in \psi(\succ_{k-1})$. Therefore, when we construct $\succ_k$ from $\succ_{k-1}$, k must continue to be selected (by monotonicity of ψ). But then lower invariance implies that $k \in \psi(\succ)$, completing the proof. □

A.2 Examples Demonstrating Independence of Our Three Axioms

Example 4 (Monotone, lower-invariant but priority bossy). *Consider a setting where there are multiple minority groups that are disjoint, that is, each applicant can be a member of at most one minority. There are two types of seats: (i) unreserved seats that can be allocated to any applicant, and (ii) for each minority, there are reserved seats that can be allocated only to members of the minority. Let φ be the selection rule that, for any priority order $\succ$, generates a selection according to the following procedure:*

(1) Unreserved seats are allocated based on the priority order $\succ$.
(2) For each of the (mutually exclusive) minorities, positions that are set aside are allocated to the remaining members of the minority again based on the priority order $\succ$.

We start by showing that no selection rule $OB_{\mathcal{F}}$ can replicate the behavior of φ. Consider an instance with three applicants $\mathcal{I} = \{1, 2, 3\}$, one open seat and one seat reserved for members of the minority group $M = \{2, 3\}$. Let

$$1 \succ 2 \succ 3 \text{ and } 2 \succ' 1 \succ' 3.$$

From the definition of φ, we must have that $OB_{\mathcal{F}}(\succ) = \{1, 2\}$ so $\{1, 2\} \in \mathcal{F}$. In addition, it should be that $OB_{\mathcal{F}}(\succ') = \{2, 3\}$. However, selection $\{2, 3\}$ doesn't respects priorities $\succ'$ as $1 \succ' 3$ and $\{1, 2\} \in \mathcal{F}$.

From the instance above, we can verify that φ is priority bossy. Notice that under both orders (i) applicant 2 is selected, and (ii) applicant 1 is preferred to applicant 3. However, the overall selections differ.

Example 5 (Monotone, priority non-bossy but not lower-invariant). *Consider an instance with three applicants $\mathcal{I} = \{1, 2, 3\}$. We define the selection rule φ as follows:*

(1) If 3 is not the worst priority applicant, then select applicant 3.
(2) Otherwise, select applicants 1 and 2.

We begin by showing that no selection rule $OB_{\mathcal{F}}$ can replicate the behavior of φ. Let

$$1 \succ 2 \succ 3 \text{ and } 1 \succ' 3 \succ' 2.$$

By (2) have that $OB_{\mathcal{F}}(\succ) = \{1, 2\}$ so $\{1, 2\} \in \mathcal{F}$. In addition, by (1) we have that $OB_{\mathcal{F}}(\succ') = \{3\}$. However, selection $\{3\}$ doesn't respects priorities $\succ'$ as $1 \succ' 3$ and $\{1, 2\} \in \mathcal{F}$.

It is easy to see that φ is priority non-bossy. Notice that the two feasible selections are disjoint. Hence, if the outcome of an applicant remains unchanged, then the overall selection must be the same.

We now show that φ is monotone. On one hand, if the priority of applicant 3 is improved, then she will be selected and weakly better. On the other hand, applicants $\{1, 2\}$ are only selected when they are *in the first two positions. In this situation, the only improvement feasible is to increase the priority of the applicant in the second position. Of course, this maintains the overall selection and this agent is weakly better.*

Finally, we see that φ is not lower-invariant. Notice that applicant 1 has the top priority in $\succ$ and $\succ'$. However, φ only selects applicant 1 under $\succ$.

Example 6 (Priority Non-bossy, lower-invariant but not monotone). *Consider an instance with two applicants $\mathcal{I} = \{1, 2\}$. Let the selection rule φ be defined as follows:*

(1) If applicant 1 has the top priority, select all applicants.
(2) Otherwise, select nobody.

We start by showing that no selection rule $OB_{\mathcal{F}}$ can replicate the behavior of φ. Let

$$1 \succ 2 \text{ and } 2 \succ' 1.$$

From (1) above, we must have that $OB_{\mathcal{F}}(\succ) = \{1, 2\}$ so $\{1, 2\} \in \mathcal{F}$. In addition, it should be that $OB_{\mathcal{F}}(\succ') = \emptyset$. Of course, selecting no applicants does not respects priorities $\succ'$ as applicant 2 has the first priority and $\{1, 2\} \in \mathcal{F}$.

From the instance above, we see that φ is not monotone. Notice that the priority of applicant 2 is better in $\succ'$ than in $\succ$ but it is only selected under $\succ$.

Clearly, φ is priority non-bossy. Note that φ either selects all applicants or nobody. Hence, if the status of an applicant remains unchanged so does the overall selection.

Finally, we show that φ is lower-invariant. This follows from the fact that the selection is determined by the applicant who has the top priority.

B PROOFS FROM SECTION 4

Proof of Lemma 1. Lemma 1 is now standard and goes back to Gale [1968]. We include a proof for the sake of completeness.

For the sake of contradiction, suppose that there exists a priority order $\succ$ and a feasible selection S, such that S is not priority dominated by $OB_{\mathcal{F}}(\succ)$.

For each selection S and integer $k \in \{1, \ldots, |S|\}$, we let $S^k \subseteq S$ be the subset containing the k highest priority individuals in S, that is,

$$S_k = \{i_1^S, \ldots, i_k^S\}.$$

Let $k^* \in \{1, \ldots, |S|\}$ be the highest-priority index for which

$$i_{k^*}^S \succ i_{k^*}^{OB_{\mathcal{F}}(\succ)}.$$

The existence of k^* is guaranteed as S is not dominated by $OB_{\mathcal{F}}(\succ)$. Consider $OB_{\mathcal{F}}(\succ)^{k^*-1}$ and S^{k^*}. Because $\mathcal{F}$ induces a matroid, there must exists $j \in S^{k^*} \setminus OB_{\mathcal{F}}(\succ)^{k^*-1}$ and $S' \in \mathcal{F}$ such that $OB_{\mathcal{F}}(\succ)^{k^*-1} \cup \{j\} \subset S'$. Then $OB_{\mathcal{F}}(\succ)$ doesn't respect priorities as $j \notin OB_{\mathcal{F}}(\succ)$ but can be selected without displacing any higher priority individual in $OB_{\mathcal{F}}(\succ)$.

□

B.1 Reserved Positions

Proof of Theorem 2. **Hard Reserves.** In the case of hard reserves, $\mathcal{F}^{hard}$ induces a transversal matroid [Oxley, 2011, Section 1.6]; we omit the standard proof of this result. (This is a special case of a gammoid (see https://en.wikipedia.org/wiki/Gammoid).

Maximal Reserves. Let S and S' be two feasible selections of maximum cardinality, that is, $|S| = |\mathcal{P}| = |S'|$. By the basis exchange property [Oxley, 2011, p.16-17], it suffices to show that for any $i \in S \setminus S'$ there is a $j \in S' \setminus S$ such that $S \setminus \{i\} \cup \{j\}$ is a feasible selection.

Suppose there is a maximal matching M of G such that $I_M \subseteq S$ such that $i \notin I_M$. In this case, agent i is a "free agent" in the sense that i can be replaced by *any* agent $j \in S' \setminus S$ to obtain another feasible selection. Thus we may assume that i is matched in *every* maximal matching M of G with $I_M \subseteq S$. Among all such matchings, choose one that maximizes I_M. Let M' be a maximal matching of G such that $I_{M'} \subseteq S'$.

Consider the graph consisting of the edges of M and M' only. Each node in this graph has degree at most 2: those with degree 2 are matched in both M and M'; those with degree 1 are matched in exactly one of M or M'. Note that i has degree 1 in this graph (it is matched in M but not in M'). Let p be the position assigned to i in M. Note that p must be assigned to some agent $k \in M'$ and hence S': otherwise M' would not be a maximal matching, as (i, p) can be added to M'. We now consider all the possibilities:

- If $k \notin S$, then $S \setminus \{i\} \cup \{k\}$ is a feasible selection: the matching M'' obtained from M by dropping (i, p) and adding (k, p) is a maximal matching with $I_{M''} \subseteq S \setminus \{i\} \cup \{k\}$.
- If $k \in S$, then k must be in M as well: for otherwise, k is a "free agent" in S; the matching M'' obtained from M by dropping (i, p) and adding (k, p) is a matching with $I_{M''} \subseteq S$ in which i is a "free agent" and $|I_{M''}| = |I_M|$. By our assumption M'' is not maximal as it does not match i, which implies that M'' can be made larger; but this contradicts our initial choice of M.

To summarize: if $k \notin S$, then we can replace i with k to complete the argument; if $k \in S$, then $k \in M$, in which case k is matched to some other position p'; this position p' must be matched to some $k' \in S'$. Repeating the same argument, we can replace i with k' if $k' \notin S$, and extend it yet again otherwise. The proof is complete by noting that the path originating at node i must end at some node $\hat{k}$ with degree 1 in M'.

Maximum Reserves. Let S and S' be two feasible selections of maximum cardinality, that is, $|S| = |\mathcal{P}| = |S'|$. As before, it suffices to show that for any $i \in S \setminus S'$ there is a $j \in S' \setminus S$ such that $S \setminus \{i\} \cup \{j\}$ is a feasible selection.

Suppose there is a maximum cardinality (or simply, maximum) matching M of G such that $I_M \subseteq S$ such that $i \notin I_M$. In this case, agent i is a "free agent" in the sense that i can be replaced by *any* agent $j \in S' \setminus S$ to obtain another feasible selection. Thus we may assume that i is matched in *every* maximum matching M of G with $I_M \subseteq S$. Let M' be a maximum matching of G such that $I_{M'} \subseteq S'$.

Consider the graph consisting of the edges of M and M' only. Each node in this graph has degree at most 2: those with degree 2 are matched in both M and M'; those with degree 1 are matched in exactly one of M or M'. Note that i has degree 1 in this graph (it is matched in M but not in M'). Let p be the position assigned to i in M. Note that p must be assigned to some agent $k \in M'$ and hence S': otherwise M' would not be a maximal matching, as (i, p) can be added to M'. We now consider all the possibilities:

- If $k \notin S$, then $S \setminus \{i\} \cup \{k\}$ is a feasible selection: the matching M'' obtained from M by dropping (i, p) and adding (k, p) is a maximum matching with $I_{M''} \subseteq S \setminus \{i\} \cup \{k\}$.
- If $k \in S$, then k must be in M as well: for otherwise, k is a "free agent" in S; the matching M'' obtained from M by dropping (i, p) and adding (k, p) is a matching with $I_{M''} \subseteq S$ in which i is a "free agent" and $|I_{M''}| = |I_M|$, which is also a maximum matching.

To summarize: if $k \notin S$, then we can replace i with k to complete the argument; if $k \in S$, then $k \in M$, in which case k is matched to some other position p'; this position p' must be matched to some $k' \in S'$. Repeating the same argument, we can replace i with k' if $k' \notin S$, and extend it yet again otherwise. The proof is complete by noting that the path originating at node i must end at some node $\hat{k}$ with degree 1 in M'.

□

B.1.1 Soft Reserves.

Proof of Proposition 3. We start by proving the hardness result for maximal reserves. We give a reduction from the problem "minimum maximal matching" (MMM) of deciding whether a bipartite graph G admits a maximal matching of size at most k. This problem was shown to be NP-complete by Yannakakis and Gavril [1980]. Consider an instance of MMM, that is, a bipartite graph $G = (U, V, E)$ and a positive integer k. We generate an instance $I(G)$ of the selection problem under maximal reserves as follows. We set $\mathcal{P} = V$ and $\mathcal{I} = \mathcal{I}^1 \cup \mathcal{I}^2$, where $\mathcal{I}^1 = U$ and $\mathcal{I}^2$ consists of $|V|$ individuals who are not connected to any position. Let the priority order $\succ$ place all individuals in $\mathcal{I}^2$ ahead of all individuals in $\mathcal{I}^1$. Then G admits a maximal matching of size at most k if and only if it there is a selection in $\mathcal{F}^{maximal}(G)$ containing the first $|V| - k$ individuals from $\succ$.

We now prove the positive result for maximum reserves. Consider the compatibility graph $G = G(\mathcal{I}, \mathcal{P}, E)$. We make use of a result of Dulmage and Mendelsohn [1958], who established a canonical decomposition for bipartite graphs [6]: The set of applicants $\mathcal{I}$ can be partitioned into three parts $\mathcal{I}_o, \mathcal{I}_e$ and $\mathcal{I}_u$and similarly, the positions can be partitioned into three parts $\mathcal{P}_u, \mathcal{P}_e$, and $\mathcal{P}_o$ such that *every* maximum cardinality matching in G has the following structure:

- There is a perfect matching of the agents in $\mathcal{I}_e$ with the positions in $\mathcal{P}_e$.
- All the positions in $\mathcal{P}_o$ are assigned to agents in $\mathcal{I}_u$, and it is possible to find a matching that omits any specified agent in $\mathcal{I}_u$.
- All the agents in $\mathcal{I}_o$ are assigned to some position in $\mathcal{P}_u$, and it is possible to find a matching that omits any specified position in $\mathcal{P}_u$.

In particular, the agents in $\mathcal{I}_e$ and $\mathcal{I}_o$ are matched in *all* maximum cardinality matchings, and the agents do not consume any position in $\mathcal{P}_o$; and all positions in $\mathcal{P}_o$ are allocated to some subset of agents in $\mathcal{I}_u$.

Moreover, this decomposition can be found efficiently using, for example, an algorithm to find a maximum cardinality matching.

Given this result, a priority dominant selection under maximum reserves can be computed as follows: match all agents in $\mathcal{I}_e$ and $\mathcal{I}_o$; find a maximum-cardinality matching of the positions in $\mathcal{P}_o$ to the agents in $\mathcal{I}_u$; finally, pick, in decreasing order of priority, as many of the unassigned agents from $\mathcal{I}_u$ as the number of unassigned positions in $\mathcal{P}_u$. □

B.2 Quotas

Proof of Proposition 4. This follows from Lemma 4 and/or Lemma 5, each of which proves the result for a special case of the general quotas problem. □

Lemma 4. *Suppose there is a trait $0 \in T$, such that $\mathcal{I}_0 = \mathcal{I}$ and $\ell_0 = q$. Additionally, for every other trait $t \in T \setminus \{0\}$, $\ell_t = 0$. If the collection of feasible selections is defined as in* (5)*, then*

- *determining whether there exists a feasible selection is NP-complete, and*
- *the selection that respects priorities might not priority dominate all other feasible selections.*

Proof of Lemma 4. We first show that determining whether there exists a feasible selection is NP-complete, via a reduction from the Independent Set Problem. An instance of the Independent Set Problem is given by a graph $G = (V, E)$ and an integer k. An *independent set* of graph G is a set S of vertices such that no two vertices in S are adjacent. The goal is to determine whether there

[6] This was subsequently generalized, independently, by Gallai and Edmonds; see [Lovász and Plummer, 2009, Section 3.2].

exists an independent set containing at least k elements. Given an instance of the Independent Set Problem, we construct an instance of the Selection Problem with upper quotas as follows: create one individual for each vertex, and one trait (with only two individuals) for each edge. Let each trait have an upper quota of 1, and let the overall lower quota be k. Then for any selection, the upper quotas are satisfied if and only if the set of vertices associated with the selected individuals is independent. Therefore, deciding whether there exists a feasible selection is equivalent to determining whether there is an independent set of size at least k.

We conclude by noting that Example 3 gives an instance satisfying the conditions of Lemma 4 in which the selection that respects priorities does not priority dominate all other feasible selections. □

Lemma 5. *Suppose there is a trait* $0 \in T$*, such that* $\mathcal{I}_0 = \mathcal{I}$ *and* $u_0 = q$*. Additionally, for every other trait* $t \in T \setminus \{0\}$*,* $u_t = |\mathcal{I}|$*. If the collection of feasible selections is defined as in* (5)*, then*

- *determining whether there exists a feasible selection is NP-complete, and*
- *the selection that respects priorities might not priority dominate all other feasible selections.*

Proof of Lemma 5. First, we show that determining whether there exists a feasible selection is NP-complete, using a reduction from the Set Cover Problem. An instance of the Set Cover Problem is given by a ground set $\mathcal{U}$, a collection $\mathcal{S}$ of subsets of $\mathcal{U}$, and an integer k. A *cover* for $\mathcal{U}$ is a subcollection $C \subseteq \mathcal{S}$ of sets whose union is $\mathcal{U}$. The objective is to determine whether there exists a set cover $C \subseteq \mathcal{S}$ for $\mathcal{U}$ containing at most k elements. Given an instance of the Set Cover Problem, we construct an instance of the Selection Problem with lower quotas as follows: create one trait for each element in $\mathcal{U}$, and one individual for each set in $\mathcal{S}$ (having traits identified by the set). There is a lower quota of 1 for each trait, and an overall upper quota equal to k. Then a selection of individuals satisfies all minimum quotas if and only if the corresponding subsets cover $\mathcal{U}$. Therefore, determining whether there exists a feasible selection is equivalent to determining whether there is a set cover with at most k elements.

Secondly, we give an instance where the selection that respects priorities does not priority dominate all other feasible selections. Consider an instance with four individuals $1, 2, 3, 4$, and three traits $T = \{0, A, B\}$. Let $\mathcal{I}_A = \{2, 4\}$, $\mathcal{I}_B = \{3, 4\}$, $\ell_A = 1 = \ell_B$ and $u_0 = 2$. Suppose that $1 \succ 2 \succ 3 \succ 4$. Then the feasible selection that respects priorities is $\{1, 4\}$. However, this selection doesn't priority dominate selection $\{2, 3\}$. □

Proof of Proposition 5. The sufficiency (first) part of the result is a restatement of Lemma 3 in Yokoi [2017]. To show neccessity, suppose that $\mathcal{I}_T$ does not form a hierarchy. Then there exist traits t_1 and t_2 and individuals i_1, i_2, i_3 such that $i_1 \in \mathcal{I}_{t_1} \setminus \mathcal{I}_{t_2}$, $i_2 \in \mathcal{I}_{t_2} \setminus \mathcal{I}_{t_1}$, and $i_3 \in \mathcal{I}_{t_1} \cap \mathcal{I}_{t_2}$.

Set all lower quotas to be zero, set an upper quota of 1 for each of traits t_1 and t_2, and a (vacuous) upper quota of $|\mathcal{I}|$ for all other traits. Let $3 \succ 2 \succ 1$ be the three highest-ranked individuals according to $\succ$. Then the outcome based rule selects individual 3, along with everyone who does not have traits t_1 or t_2. However, it is also feasible to select individuals 1 and 2 along with everyone who does not have traits t_1 or t_2. This selection is not priority dominated by the selection of the outcome based rule. By Proposition 1, it follows that there is no feasible selection that priority dominates all others.

□

C PROOFS FROM SECTION 5

Proof of Theorem 2. In what follows, it will be convenient to relax the gender disparity requirement and allow for the gap between the number of male and female electees to be at most

K for some $K \geq 1$. (The rule in the chilean election requires K to be exactly 1.) Note that the specification of the rule in Chile applies just as well to any K.

Let $\mathcal{F}_K$ be the collection of feasible sets for the Chilean election problem under the condition that the gender gap is required to be at most K. That is, every set $S \in \mathcal{F}_K$ satisfies the following two properties: (i) the number of candidates in S from party i is exactly E_i; and (ii) the difference between the number of men in S and the number of women in S is at most K.

Suppose there are N candidates in all. Clearly,

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \ldots \subseteq \mathcal{F}_{N-1} \subseteq \mathcal{F}_N.$$

We assume that $\mathcal{F}_1$ is nonempty so that there is always a feasible selection regardless of the gender disparity requirement.

We claim that the outcome of this procedure is a selection that respects priorities. Thus, the algorithm used in Chile can be seen as the outcome based rule applied to $\mathcal{F}_K$.

We do this by showing that Algorithm 2 maintains the following invariant throughout its execution: suppose candidate j belonging to party i is *not* selected:

(a) if j is a minority gender candidate, then E_i higher priority candidates from party i are selected
(b) if j is a majority gender candidate, either E_i higher priority candidates from party i are selected, or the following conditions are both true:
  (i) every lower priority candidate belonging to party i who is selected is a minority gender candidate; and
  (ii) if a lower priority candidate of the *majority* gender is selected, every subsequent (i.e. lower priority) minority gender candidate from the same party is also selected.

We first argue that any selection satisfying the stated invariant respects priorities. Consider any candidate j who is not selected, and suppose j belongs to party i.

If E_i candidates of higher priority from party i are selected, then clearly j cannot be included in a feasible selection while respecting priorities.

Suppose the selection has fewer than E_i higher priority candidates from party i. By the invariant, this implies that j is a majority gender candidate, and at least one lower-priority candidate from party i is selected. Thus, j must have been replaced by a candidate ℓ of the minority gender in the swap step of the algorithm to satisfy the gender disparity constraint.

Consider any selection that includes every selected candidate with higher priority than j as well as candidate j. To satisfy the party quota constraint, at least one lower-priority candidate from party i must be *unselected*, but every such candidate is a minority candidate, so the gender parity constraint will still be violated. Thus, the only way to reduce the gender disparity is to unselect a (at least one) lower-priority majority-gender candidate of some party—necessarily other than party i—and add a minority-gender candidate of the same party. By, property b(ii) no such choice exists.

We next prove that Algorithm 2 maintains the invariant. This is established by induction on K: For K large enough, this is clear: The top E_i candidates from each party i are chosen, and each candidate that is not chosen satisfies (a) and (b) in the claim above: in fact (a) and (b) are true for *all* unselected candidates.

Suppose the claim is true for $K \geq k$. We want to show that the claim remains true for $K = k - 1$. Let S_k be the selection found by this algorithm when the gender gap is required to be at most k. If $S_k \in \mathcal{F}_{k-1}$, then the claim is clearly true (as $\mathcal{F}_{k-1} \subseteq \mathcal{F}_k$.)

The only remaining case to consider is when $S_k \notin \mathcal{F}_{k-1}$: that is the gender gap in S_k is exactly k. Suppose S_k includes more men than women. This implies that at least one of the men selected in S_k must not be selected and at least one woman not selected in S_k must be selected. In this case, suppose the algorithm in the swap step replaces man m with woman w, both belonging to party

i. Note that in this step, the status of candidates from parties other than party i is unaffected; a majority candidate is *unselected*; and a minority candidate is *selected*. We observe also that this step of the algorithm cannot affect b(i) or b(ii) for *any* candidate: these properties can be affected only if a minority gender candidate is *unselected* or if a majority gender candidate is *selected*; the algorithm however *unselects* a majority gender candidate and *selects* a minority gender candidate. Finally, (**) by the choice of m, any man m' with lower-priority than m who is selected in S_{k-1} (and hence also S_k) should belong to a party different from i; moreover, every woman w' belonging to the same party as m' and having a lower-priority than m' is also selected in S_{k-1} (and also S_k).

We now argue that the choice made in the swap step of the algorithm is such that the invariant is maintained. Pick any j who is not selected by the algorithm.

- **Case 1: Suppose j belongs to party $i' \neq i$.** In this case, j was not chosen in S_k as well; by the induction hypothesis, j satisfied the invariant before the replacement of m with w. Moreover, replacing m with w does not impact (a) and (b) of the invariant as the only changes effected by this step of the algorithm are in the composition of the candidates chosen from party i; as noted earlier, properties b(i) and b(ii) also remain unaffected by the replacement of m with w. Thus, the invariant is satisfied by j.
- **Case 2: Suppose j belongs to party i.**
  - If j has higher priority than m or lower priority than w, then conditions (a) and (b) are unaffected; moreover b(i) and b(ii) are also unaffected by our earlier observation. By the induction hypothesis, j satisfied the invariant before the current step and so the invariant is maintained.
  - Suppose j has (strictly) lower priority than m but higher priority than w. Note that j must be a majority gender candidate (as otherwise j would have been chosen instead of w), and was not selected in S_k as well. If fewer than E_i candidates with higher priority than j were chosen in S_k, then j must have satisfied (b)(i) and (b)(ii) in S_k, and so will continue to satisfy them in this step (by the induction hypothesis). Suppose exactly E_i candidates with higher priority than j were chosen in S_k. We know that S_k does not contain any lower priority candidate belonging to party i. Moreover, by (**) earlier, any lower priority majority gender candidate included in S_k *was not chosen* in the current step (to find S_{k-1}) because no minority gender candidate belonging to that party could be included instead. Thus, j satisfies b(i) and b(ii) in S_{k-1}.
  - Finally, suppose $j = m$. By the choice made by the algorithm, m is the lowest priority majority-gender candidate in S_k, and w is the highest-priority minority-gender candidate not chosen in S_k. Thus, b(i) is satisifed for j in S_{k-1}. Moreover, by (**) earlier, any lower priority majority gender candidate included in S_k *was not chosen* in the current step (to find S_{k-1}) because no minority gender candidate belonging to that party could be included instead. Thus, j satisfies b(i) and b(ii) in S_{k-1}.

□