Balancing Welfare and Distributional Goals in School Choice*

Justin Hadad† Irene Lo‡ Ivan-Aleksandar Mavrov§

August 23, 2026

Abstract

We study a school choice problem in which a school district jointly values student welfare and diversity across its schools, trading off the two objectives rather than treating either as a constraint. In practice, the tradeoff can be large and is a function of the preferences and priorities in the market. We take an axiomatic approach: we provide an ordered menu of matchings which are student-optimal for a given level of diversity and which are within-type fair, where no student envies another of her same type. We additionally identify a lattice of matchings that are both within- and across-type fair, where no student envies another student (regardless of type) who does not improve diversity at the school. Both collections of matchings provide a tradeoff between welfare and diversity: matchings later in the ordering or higher in the lattice improve diversity at the expense of welfare. To find them, we introduce a new class of mechanisms, *deferred acceptance with waiting rooms*. Finally, we show that existing approaches to diversity, such as hard and soft reserve policies, frequently correspond to extreme points of the welfare–diversity tradeoff, and characterize the intermediate allocations. Taken together, our results provide a framework for navigating the diversity-welfare tradeoff in practice.

1 Introduction

Public school assignment systems are designed to serve multiple objectives. While the canonical goal is student welfare, districts frequently articulate distributional goals—for example, that schools reflect the demographic composition of the applicant pool, that historically underrepresented groups are not concentrated in a small set of schools, or that each school enrolls a balanced share of students across salient categories. These distributional goals are commonly described as "diversity" objectives.

In practice, diversity is rarely pursued in isolation, and policy rules which impose a fixed diversity requirement while remaining agnostic to its welfare cost can be too blunt of an instrument.

*For their comments, we would like to thank Itai Ashlagi, Aytek Erdil, Yannai Gonczarowski, Daniel Quigley, Augustus Smith, and Alex Teytelboym, as well as seminar participants at the 8th International Workshop on Matching Under Preferences (Match-UP '26) and the Twenty-Seventh ACM Conference on Economics and Computation (EC '26).

†University of Oxford.

‡Stanford University.

§Stanford University.

A planner may instead want a principled way to select an assignment that trades off diversity and welfare to a reasonable extent, while still satisfying normative constraints that make the outcome defensible and explainable to families.

We study the problem of balancing welfare and distributional goals from the perspective of a school district designing a centralized school choice system. In our model, students with one-dimensional types submit preference orderings over schools, each with a finite capacity; there are weakly more students than school seats. The district cares about both (i) student welfare, captured by preference satisfaction, and (ii) diversity, captured by distance from a target profile of types at each school.

An important goal in school choice, especially in the context of diversity, is for allocations to be explainable to families. For example, in their seminal paper on school choice, Abdulkadiroglu and Sonmez (2003) cites the following from Henig (1994):

> *'The first step that districts must take to ensure fair implementation of choice within and among schools is to make the criteria for accepting and rejecting transfer requests clear and public. Vaguely worded references to "maintaining racial balance," "avoiding overcrowding," and meeting children's "individualized needs" invite selective interpretation unless they are accompanied by practical definitions.'*

We approach explainability of the allocation from an axiomatic angle: we state the district's normative commitments to priorities, student welfare, and diversity as a collection of transparent, verifiable properties, and characterize and compare the matchings satisfying these properties.

Our first set of contributions is a set of axioms and relations between matchings that allow for tradeoffs between diversity and welfare while respecting school priorities. We formalize respecting priorities using two notions of fairness. *Within-type fairness* is an existing notion of fairness that compares the priorities of students of the same type. We introduce *across-type fairness (with respect to reserves)*, which compares the priorities of students across types for students beyond the target profile, which we call the *reserves*. We capture a minimal level of diversity and efficiency by defining *reserve-consistency*: no student should prefer a school that has not met its target for that student's type. We compare matchings by their diversity, as measured by distance from the target type profile, and welfare, as measured by Pareto dominance.

Our first main result is to provide a *menu* of matchings which maximally satisfy preferences and priorities for a given level of diversity. Formally, these matchings are within-type fair, reserve-consistent, and Pareto efficient among within-type fair matchings with weakly more diverse allocations of student types to schools. The menu is ordered in that matchings later in the ordering strictly decrease welfare and strictly improve diversity. We thus allow a district to choose a matching in the menu for which the gain in diversity is worth the cost in welfare.

Our second main result is to provide a menu of matchings which are maximally fair for a given level of diversity. These matchings have no blocking pairs across all types, thus satisfying both within- and across-type fairness, and are reserve-consistent. The menu is again ordered inversely

by welfare and diversity, but the diversity guarantee is weaker: the number of total students in excess of the reserve target at each school weakly decreases along the menu.

Our results are constructive: we provide algorithms that, for each desired set of properties, can identify a corresponding menu of matchings in polynomial time. These algorithms come from a novel class of mechanisms: *deferred acceptance with waiting rooms* (DAW). DAW constructs feasible matchings in stages that temporarily assign students beyond a school's target type profile to 'waiting rooms' at that school. Procedurally, DAW is similar to standard deferred acceptance in that students propose to schools and schools reject students. However, instead of specifying an ex-ante choice rule that determines how schools select students from a set, we define path-dependent *rejection rules*: the type of the bottom-priority student who is rejected by a school depends on the type of student who most recently applied to that school. With a specific path-dependent rejection rule, the district can both control for diversity between stages on the path, and ensure the matching at each stage on the path satisfies the target axioms. To our knowledge, our use of rejection rules in this setting is novel.

Our results shed normative light on the properties of policies typically used in practice to achieve given diversity targets. One such policy is deferred acceptance with *hard reserves*, where any student in surplus of the target type profile is rejected. With weakly more students than school seats and appropriately set type targets, the matching generated by hard reserves exactly matches the target profile, but the welfare cost to achieve this matching can be large. We show the hard reserves matching is within-type fair and efficient, and is the diversity-optimal endpoint of both menus of matchings. Another oft-considered policy is deferred acceptance with *soft reserves*, where any student in surplus of the target type profile is rejected only once the school's total capacity is reached. We show that the soft reserves matching might not be efficient, and thus can be costly for both welfare and diversity. However, the matching is welfare-optimal amongst those that are within-type fair, across-type fair, and reserve-consistent.

To illustrate the costs that hard and soft reserves can have on welfare and diversity respectively, consider the market in Example 1, also illustrated in Figure 1.

Example 1 (Reserves policies achieve extreme outcomes)**.** There are two schools, A and B, with reserve targets of 50 type a and 50 type b students each. Each of the 100 type a students has preferences $A \succ_a B$, and each of the 100 type b students has preferences $B \succ_b A$. The matching from hard reserves matches half the students to their least-preferred school, meanwhile the matching from soft reserves achieves no diversity: $\mu(a) = A$ for all a and $\mu(b) = B$ for all b. □

In this market, standard reserve-based policies achieve extreme outcomes that either maximize welfare at the expense of any diversity, or maximize diversity at the expense of welfare for half the students. Our DAW mechanisms identify 50 interim matchings between the hard- and soft-reserve matchings that trade off between diversity and welfare. Our results reflect our belief that such interim matchings should exist in more general markets, and a district should be able to choose a matching which optimally trades off her objectives. For example, a district with a nonlinear

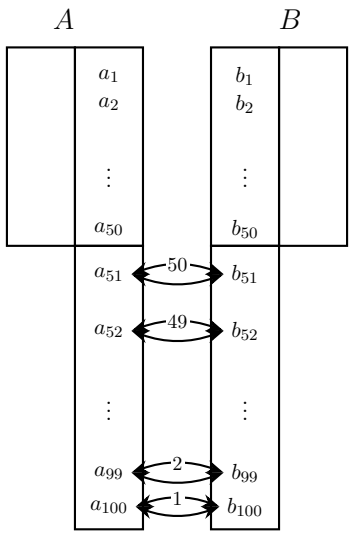


Figure 1: The economy in Example 1 has 50 interim matchings, corresponding to swapping the allocation of the students in cycles 1 through i for each $i \leq 50$.

objective over diversity and welfare might prefer a matching between the hard-reserves and soft-reserves outcomes. In addition, the optimal matching for a given school-choice problem can be highly allocation-specific.

Our results carry two normative lessons for the literature on diversity in matching. First, they shift attention from identifying a single reserve rule to designing a transparent design procedure by which a district can navigate an unavoidable tradeoff. When a district cares about both preference satisfaction and the distribution of types across schools, applying a fixed rule mechanically can be highly consequential for welfare. Different points on the welfare-diversity frontier correspond to substantively different value judgments, and our ordered menus and lattices make those judgments explicit. Our use of path-dependent rejection rules motivates restricting a designer's search through all possible type allocations to only those that appear on the frontier. Second, our results caution against treating ostensibly-milder reserve policies as optimal policy. Soft reserves are motivated by the intuition that they achieve targets while not harming welfare too much; in contrast, we show that the soft reserves matching is always weakly dominated in terms of both welfare *and* diversity.

1.1 Related Literature

Starting with the seminal paper on school choice of Abdulkadiroglu and Sonmez (2003), a large literature has studied distributional goals in matching through the lens of 'controlled school choice' (Abdulkadiroğlu, 2005; Kominers and Sönmez, 2013; Echenique and Yenmez, 2015). This literature introduced hard reserves (quotas) and soft reserves as tools for achieving desired diversity targets. It also demonstrated that increasing diversity requirements can reduce overall welfare: Kojima (2012) establishes impossibility results showing that setting maximum quotas for majority students can make minority students worse off, Hafalir et al. (2013) introduces soft reserves for minority students

as an alternative to rigid quotas, and finds welfare improvements, and Ehlers et al. (2014) compare hard and soft reserves and shows that soft reserves improve welfare.

Many of the papers in this well-established literature encode diversity as a constraint: the desired level of diversity is implemented through target hard reserves on student types at schools, and relaxed by using soft reserves. An active and growing literature refines the definition of appropriate diversity targets, particularly in settings with overlapping types (Kurata et al., 2017; Sönmez and Yenmez, 2019). Another line of work develops generalized reserve mechanisms that efficiently assign students eligible for multiple reserve categories (Aziz et al., 2020; Abdulkadiroğlu and Grigoryan, 2021). Beyond designing reserves, a complementary line of work studies designing a single school's choice rule to satisfy nuanced definitions of diversity as a constraint. Sun and Yokoo (2025) characterize single-school choice functions that achieve balanced representation across types, including overlapping categories.

Our paper is closest to an emerging line of research that treats diversity as an objective, rather than a constraint. We aim to provide a menu of matchings that trades off welfare and diversity. This tradeoff-driven approach is similar to that taken by Imamura (2025), who proposes and axiomatically studies reserve-based choice rules that trade off between merit and diversity, Çelebi and Flynn (2025), who develop adaptive priority mechanisms that adaptively adjust priorities to balance merit and diversity, and Hafalir et al. (2022), who use matroid structure to characterize the Pareto frontier of diversity-merit tradeoffs. While these works focus on merit and diversity at the school level, our paper is distinguished by trading off between welfare and diversity at the district level.

Our diversity objective is defined as the distance to a target allocation of student types across schools. A similar distributional goal has previously been articulated by Erdil and Kumano (2019), and a natural choice of target allocation is the proportional allocation, which has been studied by Nguyen and Vohra (2019). Also related is the diversity notion derived from majorization together with its associated choice rule formalized for a single school in Echenique et al. (2025). These papers treat diversity as the primary optimization objective, and provide algorithmic methods for finding matchings that are Pareto efficient among those satisfying stability or proportionality constraints. To our knowledge, we are the first to combine an axiomatic foundation with algorithmic approaches for trading off between diversity and other goals at a district level.

2 Model

We consider a school choice problem with a finite set of students S and a finite set of schools C. Each school $c \in C$ has capacity $|c|$. Students are partitioned into a finite set of types T: for each $t \in T$, let $S_t \subseteq S$ denote the set of type-t students, with $S = \bigcup_{t \in T} S_t$ and $S_t \cap S_{t'} = \varnothing$ for all $t \neq t'$.

For each student $s \in S$, let $\tau(s) \in T$ denote her type. Each school c has type-specific reserves $(r_{c,t})_{t \in T}$ satisfying $\sum_{t \in T} r_{c,t} = |c|$; we interpret $r_{c,t}$ as the number of seats at c reserved for type t, and write $r = (r_{c,t})_{c \in C, t \in T}$ when clear from the context. Each school c has a priority order over

students represented by a priority score $p_c : S \to \mathbb{R}_+$, where $p_c(s) > p_c(s')$ means that c ranks s ahead of s', and each student $s \in S$ has a preference ordering $\succ_s$ over the set of schools C.

An *economy* is a tuple

$$\mathcal{E} \;=\; \big(S, C, T, \tau, (|c|)_{c\in C}, (\succ_s)_{s\in S}, (p_c)_{c\in C}\big).$$

An *economy with reserves* is the pair $(\mathcal{E}, r)$. Given $(\mathcal{E}, r)$, the induced allocation problem is to select a matching $\mu : S \cup C \to C \cup S \cup \{\varnothing\}$ with the standard interpretation and feasibility conditions. For each student $s \in S$, $\mu(s) \in C \cup \{\varnothing\}$ is either the school to which s is assigned or $\varnothing$ (unassigned). For each school $c \in C$, $\mu(c) \subseteq S$ is the set of students assigned to c, and it must satisfy the capacity constraint $|\mu(c)| \le |c|$. Assignments are consistent in the sense that $\mu(s) = c$ if and only if $s \in \mu(c)$, and $\mu(s) = \varnothing$ if and only if $s \notin \mu(c)$ for all $c \in C$. We write $\mu'(s) \succsim \mu(s)$ to denote that student s weakly prefers her match in μ' to her match in μ. Additionally, define the following functions on the matching μ, which capture the *number* of students of type t assigned to school c under μ, and the *distance* from μ to the target allocation r:

$$n_{c,t}(\mu) := |\mu(c) \cap S_t| \qquad \text{and} \qquad \Delta_{c,t}(\mu, r) := n_{c,t}(\mu) - r_{c,t}.$$

We say that type t is in strict deficit at c if $\Delta_{c,t}(\mu, r) < 0$, and in strict surplus if $\Delta_{c,t}(\mu, r) > 0$. Write the least-preferred (bottom) student of type t at c under μ as

$$b_{c,t}(\mu) \in \arg\min_{s\in\mu(c)\cap S_t} p_c(s).$$

And define $p_c(\varnothing) := -\infty$ for all c. When $n_{c,t}(\mu) = 0$ we set $b_{c,t}(\mu) = \varnothing$. We also write $b_c(\mu)$ for the lowest-priority student assigned to c overall, with the same convention when $\mu(c) = \varnothing$.

Define the *type-t waiting room at c* as

$$w_{c,t}(\mu) := \Big\{\, s \in \mu(c) \cap S_t \;:\; \big|\{s' \in \mu(c) \cap S_t :\; p_c(s') > p_c(s)\}\big| \ge r_{c,t} \Big\}.$$

The students in the waiting room are the least-priority students who fall below the top $r_{c,t}$, according to c's priority; the waiting room is empty if $n_{c,t}(\mu) \le r_{c,t}$.

Given a matching μ, define the *type-allocation vector* induced by μ as

$$\mathrm{alloc}(\mu) \;=\; \big(n_{c,t}(\mu)\big)_{(c,t)\in C\times T}.$$

We assume that all preference and priority orderings are strict and complete.

Assumption 1 (Complete orders)**.** *Each student $s \in S$ has a complete and strict preference order $\succ_s$ over schools C and prefers any school to being unmatched. Each school $c \in C$ has a complete and strict priority order over students, represented by a score function $p_c : S \to \mathbb{R}_+$ such that $p_c(s) \neq p_c(s')$ for all distinct students $s \neq s'$.*

We then ensure that the district's type targets are feasible in the aggregate and that there is no aggregate excess of seats.

Assumption 2 (Feasible type targets)**.** *For each type $t \in T$, the number of type-t students is weakly greater than the total number of seats reserved for type t across schools:*

$$|S_t| \geq \sum_{c \in C} r_{c,t}.$$

Assumption 3 (No aggregate excess of seats)**.** *The total number of students is weakly greater than the total number of seats:*

$$|S| \geq \sum_{c \in C} |c|.$$

Our assumptions are compatible with a district who faces weak surplus in demand and who can observe the demographics of its applicants. Taken together, the assumptions ensure that priorities and preferences are well-defined and that the reserve targets are not infeasible in the aggregate; consequently, any shortfall from a reserve at one school must reflect a reallocation of students across schools rather than a shortage of that type in the market. Under Assumption 3, distance from the target reserves can be understood as misallocation of types across schools; without this assumption, this distance could be a consequence of empty seats, so distance from the target reserves becomes a less effective measure of diversity.

We note that our results (with the exception of Theorem 4) continue to hold without Assumptions 2 and 3. However, in the absence of these assumptions, some of our relations between matchings take on different economic meaning; we discuss these differences in the following section.

2.1 Relations Between Matchings

Our analysis compares matchings along two dimensions. The first is the standard welfare criterion: whether one matching makes every student weakly better off (and some strictly better off) relative to another. The second is a notion of *distributional* improvement: whether one matching moves the realized type profile at each school weakly closer to the reserve targets. We formalize these comparisons below.

Definition 1 (Pareto dominance)**.** *A matching μ Pareto dominates μ' if $\mu(s) \succsim_s \mu'(s)$ for all $s \in S$ and $\mu(s) \succ_s \mu'(s)$ for some $s \in S$.*

When we study matchings which may be identical, we will say that they weakly Pareto dominate the other.

Our notion of diversity improvement compares two matchings by asking whether the vector of distances to the target allocation moves weakly toward zero at every school–type coordinate.

Definition 2 (Diversity dominance)**.** *A matching μ diversity-dominates μ' if for every school $c \in C$ and type $t \in T$, either*

$$0 \leq \Delta_{c,t}(\mu, r) \leq \Delta_{c,t}(\mu', r) \qquad \text{or} \qquad 0 \geq \Delta_{c,t}(\mu, r) \geq \Delta_{c,t}(\mu', r).$$

Intuitively, one matching diversity-dominates another if the type allocation at *every* school is weakly closer to the target. Under Assumption 3, this pointwise improvement also results in improved diversity under any measure of distance of types or type proportions to the target allocation. However, in the absence of Assumption 3, an allocation could be pointwise closer to the target allocation, but farther from the desired allocation proportions.[1]

We are interested in assessing the change in diversity along sequences of matchings; we call a sequence of matchings obtained in order a *path*. If each matching along a path diversity-dominates its predecessor, we say that the path increases in diversity *pointwise*.

Definition 3 (Increase in diversity pointwise)**.** *A path $(\mu^0, \mu^1, \ldots, \mu^K)$ is increasing in diversity pointwise if, for every $k = 0, 1, \ldots, K-1$, μ^{k+1} diversity-dominates μ^k.*

Another measure of diversity along a path is the ℓ_1 distance from the reserve targets, which aggregates deviations across all school–type pairs. The ℓ_1 distance measure is a relaxation of pointwise diversity: a larger surplus at one school–type pair can be offset by a larger deficit reduction elsewhere.

Definition 4 (Increase in diversity in the ℓ_1 norm)**.** *Define the ℓ_1 diversity distance from the reserve targets r by*

$$D_1(\mu, r) := \sum_{c \in C} \sum_{t \in T} \left|\Delta_{c,t}(\mu, r)\right|.$$

A matching μ is (weakly) more diverse than μ' in the ℓ_1 norm if

$$D_1(\mu, r) \leq D_1(\mu', r).$$

We say a path $(\mu^0, \mu^1, \ldots, \mu^K)$ is increasing in diversity in the ℓ_1 norm if, for every $k = 0, 1, \ldots, K-1$,

$$D_1(\mu^{k+1}, r) \leq D_1(\mu^k, r).$$

2.2 Axioms

We evaluate matchings using a collection of axioms that capture three normative ideas that arise in school assignment systems: (i) the matching should be "fair" in the sense that priority is respected; (ii) seats should not be left empty when some student strictly prefers the empty seat to her assignment; and (iii) conditional on a target type-composition, the matching should not leave

[1] For example, for a single school c and two types, consider $r_c = (50, 50)$ and allocation vectors μ, μ' satisfying $n_c(\mu) = (30, 30)$ and $n_c(\mu') = (40, 30)$. Then μ' is pointwise closer to the target allocation but does not achieve the desired proportions, while μ is farther but achieves the desired proportions.

obvious Pareto improvements on the table.

In the classical matching model without reserves, fairness rules out any student–school blocking pair: a student who prefers a school should not be rejected in favor of a lower-priority student at that school. Distributional concerns complicate this comparison by allowing for priority reversals: a school may admit a lower-priority student of one type while rejecting a higher-priority student of another type in order to meet its type-specific targets. The notion of within-type fairness which we use is common in the literature.

Definition 5 (Within-type fairness)**.** *A matching μ is* within-type fair *if there do not exist a student $s \in S$ and a school $c \in C$ such that $c \succ_s \mu(s)$ and there exists a student $s' \in \mu(c)$ with $\tau(s') = \tau(s)$ and $p_c(s) > p_c(s')$.*

We then introduce a new notion of *across-type fairness with respect to reserves*, which treats reserves as the only justification for priority reversals across type. Intuitively, if a school is admitting a low-priority student of some type beyond that type's reserve target (a surplus), then a higher-priority student who prefers the school should not be excluded in favor of that surplus student.

Definition 6 (Across-type fairness with respect to reserves)**.** *A matching μ is across-type fair with respect to reserve vector r if there does not exist a student $s \in S$ such that, for some school $c \in C$ where $c \succ_s \mu(s)$, there is a student $s' \in \mu(c)$ where $p_c(s) > p_c(s')$ and where $\Delta_{c,\tau(s')}(\mu, r) > 0$.*

We say a matching is *fair everywhere* if it is within-type fair and across-type fair with respect to reserves.

We then introduce two axioms which specify that no seat or reserve should remain unused. *Non-wastefulness* rules out capacity slack, where seats are empty at a school that some student prefers to her assignment.

Definition 7 (Non-wastefulness)**.** *A matching μ is* non-wasteful *if there do not exist a student $s \in S$ and a school $c \in C$ such that $c \succ_s \mu(s)$ and $|\mu(c)| < |c|$.*

We then define *reserve-consistency*: if a school is below its reserve target for type t, then no type-t student is assigned elsewhere while preferring that school. Equivalently, whenever a type-t student prefers school c, school c is already meeting (at least) its type-t target.

Definition 8 (Reserve-consistency)**.** *A matching μ is* reserve-consistent *if for any student $s \in S$ with type $t = \tau(s)$ and any school $c \in C$,*

$$c \succ_s \mu(s) \quad \Longrightarrow \quad \Delta_{c,t}(\mu, r) \geq 0.$$

Toward describing a welfare-diversity frontier, we will often compare matchings in their level of efficiency to those with weakly more diverse allocation vectors. The next axiom holds if the district

cannot find a weakly more diverse within-type fair matching which achieves a Pareto improvement. Intuitively, *allocation-constrained efficiency* is our notion of student optimality where comparisons can only be made between matchings where diversity is not exchanged for a welfare improvement.

Definition 9 (Allocation-constrained efficiency)**.** *A matching μ is allocation-constrained efficient if there does not exist a within-type fair matching μ' which both diversity-dominates and Pareto dominates μ, i.e. $\mu'(s) \succsim_s \mu(s)$, and $0 \leq \Delta_{c,t}(\mu', r) \leq \Delta_{c,t}(\mu, r)$ or $0 \geq \Delta_{c,t}(\mu', r) \geq \Delta_{c,t}(\mu, r), \forall c \in C, \forall t \in T$.*

We additionally define a notion of limited allocation-constrained efficiency, where comparisons between matchings are only made between allocations of the same type.

Definition 10 (Limited allocation-constrained efficiency)**.** *A matching μ is limited allocation-constrained efficient if there does not exist a within-type fair matching μ' which Pareto dominates μ and where* $\text{alloc}(\mu) = \text{alloc}(\mu')$.

3 Hard and Soft Reserves

To illustrate and motivate our axioms, consider the *hard reserves* and *soft reserves* policies, which have an intuitive structure and are frequently considered in the literature. We relegate the formal descriptions and proofs to Appendix A.

Under *hard reserves*, each school tentatively accepts only up to the top-priority $r_{c,t}$ students within each type t among the students currently under consideration. Every student who is not among the top-priority $r_{c,t}$ of their type is rejected. The second is to use *soft reserves*, which treats reserves as soft cutoffs. Each school first tentatively holds the top-priority $r_{c,t}$ applicants within each type t, exactly as under hard reserves. Then, if the school still has remaining capacity, it fills the leftover seats by tentatively admitting the highest-priority remaining applicants regardless of type.

Under Assumptions 1-3, the type allocation of the hard-reserves matching is exactly the desired type profile: every student will eventually apply to a school where the type of that student is in deficit. A run of the hard-reserves mechanism is equivalent to running $|T|$ separate deferred acceptance mechanisms where $r_{c,t}$ is treated as the capacity of each school. It follows, then, that the resulting allocation is efficient and within-type fair.[2]

Proposition 1. *The hard-reserves matching is within-type fair, across-type fair with respect to reserves, reserve-consistent, and allocation-constrained efficient. Under Assumptions 2-3, the matching is also non-wasteful.*

The hard-reserves matching can impose large welfare losses for a small gain in diversity, as Example 2 shows.

[2]While satisfied, the definition of across-type fairness is vacuous here; the type profile is met exactly. Still, this property is important when we establish the lattice where the hard- and soft-reserve matchings are its endpoints.

Example 2. Consider an economy with 100 students of type x, labeled $x_0, \ldots, x_{99}$, and 100 students of type y labeled $y_0, \ldots, y_{99}$. There are 100 schools labeled $c_1, c_2, \ldots, c_{100}$, each reserving a single seat for each type. Each school prioritizes students of each type with lower index. Preferences of each student are as follows:

- $x_0 : c_1 \succ_{x_0} c_2 \succ_{x_0} \cdots$
- $x_i : c_i \succ_{x_i} c_{i+1} \succ_{x_i} \cdots$ for $i \geq 1$
- $y_0 : c_{100} \succ_{y_0} c_{99} \succ_{y_0} \cdots$
- $y_i : c_{101-i} \succ_{y_i} c_{100-i} \succ_{y_i} \cdots$ for $i \geq 1$

Observe that the top-choice matching, where each student matches to her most-preferred school, is feasible: $\mu(c_1) = \{x_0, x_1\}$, $\mu(c_{100}) = \{y_0, y_1\}$, and $\mu(c_i) = \{x_i, y_{101-i}\}$ for $2 \leq i \leq 99$. The hard-reserves matching $\mu^{HR}(c_i) = \{x_{i-1}, y_{100-i}\}$ improves diversity at only two of the 100 schools, and only by one student at each of those schools, but moves 198 out of 200 students to their second-choice school.[3] □

The soft-reserves matching, meanwhile, does not have a type allocation that is guaranteed to be the same as the target type profile. Intuitively, fewer agents are rejected in pursuit of the target, and so the resulting allocation weakly Pareto dominates the hard-reserves matching.

Proposition 2. *The soft-reserves matching is within-type fair, across-type fair with respect to reserves, reserve-consistent, non-wasteful and weakly Pareto dominates the hard-reserves matching.*

The efficiency loss in the soft-reserves matching can be large. The soft-reserves choice rule corresponds to always rejecting the least-preferred surplus students—across *all* types—when the school's capacity is reached. Allowing comparisons beyond students of the same type causes more rejections to occur. We highlight this in Example 3. Note that the possible improvement cycles which appear in the soft-reserves matchings are precisely the cycles highlighted by Kesten (2010), which the *efficiency-adjusted deferred acceptance mechanism* resolves.

Example 3 (Efficiency loss of soft reserves)**.** Consider the following balanced market. There are three types of students, *prime*, *double prime*, and *triple prime*, and three schools, A, B, C. The reserve vector across schools, for *prime*, *double prime*, and *triple prime*, are as follows: $r_A = (1, 1, 49)$; $r_B = (1, 1, 49)$, and $r_C = (98, 1, 2)$.

Students and schools have the following preferences and priorities; we omit the preference and priority orders of student c and school C respectively as they are irrelevant to the example. (The following notation will be useful. Student j'_1 is a student of type *prime* who most prefers school J. The index of the student differentiates among all type *prime* students who most prefer school J. In the same way, student k''_2 is a student of type *double prime* who most prefers school K.)

- $a : A \succ_a B \succ_a C$
- $b : B \succ_b A \succ_b C$

[3] In Appendix A, we provide an adaptation of this example where the 198 students are moved to either their second-least-preferred or least-preferred school, while again only improving diversity at two of the 100 schools.

- $A : b'_1 > b'_2 > \ldots > b'_{50} > b''_1 > a''_1 > b''_2 > a'_1 > a'_2 > \ldots > a'_{50} > c'''_i$
- $B : a'_1 > a'_2 > \ldots > a'_{50} > b''_1 > a''_1 > b'_1 > b'_2 > \ldots > b'_{50} > b''_2 > c'''_i$.

Note that c'''_i is ranked last by A, B for all indices i.

Each student begins by applying to her top-choice school. The (infeasible) top-choice matching which these preferences and priorities generate is in Matching 1.

Matching 1: (Infeasible) matching where each student gets her top-choice school in Example 3.

A			**B**			**C**		
1	1	49	1	1	49	98	1	2
a'_1	a''_1		b'_1	b''_1				c'''_1
a'_2			b'_2	b''_2				c'''_2
$\vdots$			$\vdots$					$\vdots$
a'_{50}			b'_{50}					c'''_{100}

There is a clear tension when moving away from the top-choice matching toward the soft-reserves matching. School B is in surplus, and must reject its least-preferred surplus student, which is b''_2. Student b''_2 then applies to her next most-preferred school, A, which prefers to her its least-preferred surplus student, a'_{50}. This rejection causes a cascade of rejections: each *prime* agent then swaps places. Only then does b''_2 get rejected again, this time by A, and apply to her least-preferred school, C. As all schools now meet capacity, the resulting matching corresponds to the soft-reserves matching, shown in Matching 2.

Matching 2: Soft-reserves matching from economy in Example 3.

A			**B**			**C**		
1	1	49	1	1	49	98	1	2
b'_1	a''_1		a'_1	b''_1			b''_2	c'''_1
b'_2			a'_2					c'''_2
$\vdots$			$\vdots$					$\vdots$
b'_{50}			a'_{50}					c'''_{100}

Matching 3: A matching that improves on the soft-reserves matching in Example 3.

A			**B**			**C**		
1	1	49	1	1	49	98	1	2
a'_1	a''_1		b'_1	b''_1			b''_2	c'''_1
a'_2			b'_2					c'''_2
$\vdots$			$\vdots$					$\vdots$
a'_{50}			b'_{50}					c'''_{100}

The soft-reserves matching satisfies the properties in Proposition 2. To verify that the matching is across-type fair with respect to reserves, note there is no student who prefers a school where she has higher priority than its least-preferred surplus student. But the cost of across-type fairness here is large: each student of type *prime* is at her second-choice school, and each has a suitable partner with whom she would trade her match. Indeed, if b''_2 did not apply to A — equivalently, waived her right to apply to A, as in Kesten (2010) — then the rejection cascade would not have formed and this would result in the matching shown in Matching 3. This stark efficiency loss is thus a consequence of the initial *across-type comparison* between b''_2 and the students of type *prime*.

□

We now consider a third matching: the hard-reserves matching with a reserve vector equal to the type allocation that emerges from the soft-reserves matching. This new matching maintains the properties from Proposition 1 with respect to this new vector, and thus Pareto dominates the soft-reserves matching.

Proposition 3. *The hard-reserves matching with reserve vector equal to the type allocation of soft reserves is within-type fair, reserve-consistent, and allocation-constrained efficient, and it Pareto dominates the soft reserves matching. Under Assumptions 2-3, the matching is also non-wasteful.*

By Proposition 3, the soft-reserves matching is weakly dominated in terms of both welfare and diversity: the hard-reserves matching with reserve vector equal to the type allocation of soft reserves generates the same diversity level but weakly improves welfare. In Example 3, Matching 3 is the hard-reserves matching with reserve given by the soft reserves allocation, and achieves the same type allocation while Pareto dominating the soft reserves matching.

4 Tradeoffs

Reserve policies achieve extreme outcomes: for a given preference profile, hard reserves can substantially reduce welfare, while soft reserves might not achieve any diversity. Example 1 highlights that interim matchings may be optimal for a district with a nonlinear objective. Thus we recommend that the principal choose a matching from a set of feasible matchings, each satisfying axioms in Section 2.2, and select the one that best achieves her objective in the given school choice problem.

However, as Example 3 shows, there is a clash between allocation-constrained efficiency and across-type fairness with respect to reserves. By allowing across-type comparisons in order to preserve across-type fairness, additional rejection cycles may decrease welfare. Hence we create two separate menus, an *efficient* and a *fair-everywhere menu*. These menus contain paths which are ordered: the efficient menu contains paths which increase in diversity pointwise, and the fair-everywhere menu contains paths which increase in diversity in the ℓ_1 norm.

4.1 An Efficient Menu

Our first main result constructs a menu of efficient matchings that strictly trade off between student welfare and diversity. The menu we construct is ordered: every matching later in the ordering is guaranteed to have more diversity pointwise and less welfare.

One insight from Propositions 2 and 3 is that a soft-reserves matching is efficient only if it can be generated by a hard-reserves procedure. Hence to guarantee efficiency along a path, our mechanism must consider a starting point which differs from the soft-reserves matching.

The mechanism used to achieve this menu is *deferred acceptance with multiple waiting rooms* (DAW(M)).[4] Intuitively, the mechanism takes as its starting point the (likely infeasible) matching

[4]We write M as a parameter because it specifies the rejection rule which is used in DAW.

where each student gets her top-choice school, and identifies *interim matchings* via a process of schools rejecting students and students proposing to schools. "Waiting rooms" allow the mechanism to tentatively assign surplus students to a school in these interim matchings.[5]

Each stage of DAW(M) proceeds as follows. If the interim matching μ_i at stage i is diversity-optimal (i.e. $\Delta_{c,t}(\mu_i, r) = 0$ for all c, t), the mechanism ends. Otherwise, it selects some type t in surplus at some school c and rejects $s := b_{c,t}(\mu_i)$ from c. The rejected student s then applies to her next most-preferred school c', leading to one of three cases:

(i) If student s contributes to the surplus of type t at c', then c' rejects its least-priority student of type t.

(ii) If student s contributes to the surplus of students at c' (but there is no surplus of type t), then there is some other type $t' \neq t$ in surplus at school c', and c' rejects its least-priority student of type t'.

(iii) Otherwise, the stage ends in the next interim matching: μ_{i+1}.

Observe that, in a given stage, the choices of school c and surplus type t to start the stage, and of surplus type t' to continue the stage, are non-unique. Hence the described procedure can lead to different paths. Nonetheless, each path leads to a menu of interim matchings that satisfies the axioms and ordering, which we present in Theorem 1.

Theorem 1. DAW(M) *finds an ordered menu of matchings that are allocation-constrained efficient, within-type fair, and reserve-consistent. The menu is ordered: matchings earlier in the ordering Pareto dominate and are diversity dominated by matchings later in the order. The diversity-maximal endpoint of this menu is the hard-reserves matching.*

Theorem 1 shows that there is a menu of efficient matchings that provides a clear tradeoff between student welfare and diversity. Given any pair of matchings in this menu, one will Pareto dominate the other, but be strictly less diverse. Indeed, the order of diversity dominance implies that, for any strictly monotone function of the pointwise distances to the target reserves, every school becomes weakly more diverse along a path. In addition, such a menu can be constructed in polynomial time by running DAW(M).

The proof of Theorem 1 shows that each matching in the menu corresponds to running hard reserves for some other reserve vector r'. Hence the properties of efficiency and within-type fairness in Proposition 1 are preserved along paths in the menu. We derive the reserve-consistency of interim matchings by the allocation monotonicity that the DAW(M) provides. If, at some point in the execution of DAW(M), a given school does not meet its reserve for some type, then there must have been no other students of that type that preferred that school to their current match at any previous point of the procedure.

[5]Students of type t assigned to school c occupy either one of the $r_{c,t}$ reserved seats for their type at school c, or are in the waiting room $w_{c,t}$ for their type at school c. Higher-priority students of each type are assigned to the reserved seats.

A more general version of our DAW(M) mechanism achieves a complete characterization of matchings that are allocation-constrained efficient, within-type fair, and reserve-consistent. Recall that DAW(M) begins by assigning each student to their top choice school. Because the menu produced by DAW(M) is guaranteed to be monotone in diversity pointwise, the allocation of any interim matching in the menu must be weakly more diverse than the matching which assigns every student to their top school. This can be too restrictive; there may be matchings that are allocation-constrained efficient, within-type fair, and reserve-consistent which are thus not achievable.

We thus create a *deferred acceptance with multiple waiting rooms and activations* (DAW(M, a)) mechanism which characterizes the complete set of axioms from Theorem 1. The mechanism relies on an *activation vector* a: the school rejects a student of a given type only once the allocation of that type at that school exceeds the activation. With this approach, any matchings satisfying the axioms, including the matchings that are *less* diverse pointwise than the top-choice matching, can be achieved with an appropriately set activation vector.

Theorem 2. *A matching is allocation-constrained efficient, within-type fair, and reserve-consistent if and only if it lies in a menu that can be achieved by* DAW(M, a) *for a given activation vector.*

4.2 A Fair-Everywhere Menu

We now present an ordered menu where each matching is across-type fair with respect to reserves: no student envies another student of any type who is also in surplus at their respective school. Each matching later in the menu has less welfare and is more diverse in the ℓ_1 norm; i.e., there is weakly less total number of surplus students across all schools.

The mechanism used to achieve the fair-everywhere menu is *deferred acceptance with single waiting rooms* (DAW(S)). Similar to DAW(M), the mechanism takes as its starting point the (likely infeasible) matching where each student gets her top-choice school, and identifies *interim matchings* via a process of schools rejecting students and students proposing to schools. The 'waiting rooms', as in DAW(M), allow the mechanism to tentatively assign surplus students to schools in these interim matchings. Unlike DAW(M), however, the rejection rule in this mechanism does not consider the type of the student that applied last.

Each stage of DAW(S) proceeds as follows. If the interim matching μ_i at stage i is diversity-optimal (i.e. $\Delta_{c,t}(\mu_i, r) = 0$ for all c, t), the mechanism ends. Otherwise, it selects the surplus student s with lowest priority matched to c under the current matching. Student s then applies to her next most-preferred school c', leading to one of three cases:

(i) If student s contributes to the surplus of students at c', then the surplus student matched to c' with the least priority is rejected, and the stage continues.

(ii) If there is no surplus of students at c', but the current matching violates across-type fairness, then the surplus student matched to c' with the least priority is rejected, and the stage continues.

(iii) Otherwise the stage ends and we add the current matching to the menu.

Because under rule (i) the least-priority surplus student is rejected regardless of her type, the only possible student who may envy a surplus student is the one that is rejected first in the current stage. Rule (ii) verifies that this student does not envy any of the students at her previous school. However, note that diversity then only increases in the ℓ_1 norm; there is no guarantee that matchings on the path diversity dominate each other at every school–type coordinate, and the only guarantee is that the number of surplus students does not increase at a given school.

Theorem 3. DAW(S) *finds an ordered menu of matchings which are within-type fair, across-type fair with respect to reserves, and reserve-consistent. The menu is ordered: matchings earlier in the ordering Pareto dominate and have weakly less diversity in the ℓ_1 norm than matchings later in the order. The diversity-optimal endpoint of this menu is the hard-reserves matching.*

We then adapt DAW(S) to achieve the complete set of feasible matchings that are characterized by the within-type fair, across-type fair with respect to reserves, and reserve-consistent axioms. We call this adapted procedure DAW(SF): *deferred acceptance with single and flexible waiting rooms*. The key is to enlarge the set of students that are considered possible starting points of a stage; we include not just surplus students, but the bottom (least-prioritized) student of types which have exactly met their reserve target at a school. By defining starting points in this way, DAW(SF) adds interim matchings to its menu which are "skipped" in the DAW(S) procedure. Further, each path on the lattice can be achieved by DAW(SF) by some order of stages, and for any order of stages the menu produced by DAW(SF) forms a path on the lattice.

Theorem 4. *The set of feasible and non-wasteful matchings that are within-type fair, across-type fair with respect to reserves, and reserve-consistent forms a lattice. Each lattice path can be achieved by* DAW(SF) *by some order of stages and for any order of stages the menu produced by* DAW(SF) *forms a path on the lattice. One endpoint of the lattice is the soft-reserves matching, and the other exactly meets the target reserves.*

Theorem 4 is the only result which relies on Assumptions 2 and 3; if the reserves were not appropriately set for the student population, then some schools can reject surplus students without receiving applications from deficit students later in the algorithm. Indeed, the hard-reserves matching may then be wasteful and thus not be on the lattice of matchings that are stable with respect to the soft-reserves choice rules.

Both Theorems 3 and 4 rely on axiomatic characterizations of the matchings that are stable (as in Hatfield and Milgrom (2005)) with respect to the soft-reserves choice rules. To show that DAW(SF) can traverse any path on that stable lattice, we prove that *every* matching satisfying the relevant properties can be achieved by a sequence of rejecting least-preferred weak-surplus students, starting from any other stable matching that Pareto dominates it.

5 Theorem 1: Algorithm and Proof

In this section, we prove Theorem 1, which shows that deferred acceptance with multiple waiting rooms gives an ordered menu of matchings that trade off between diversity and welfare, with earlier matchings diversity dominating later ones, and later matchings Pareto dominating earlier ones. The proof also illustrates the definitions and techniques we use for the other results in the paper.

5.1 Deferred Acceptance with Multiple Waiting Rooms

We begin with a formal definition of deferred acceptance with multiple waiting rooms (DAW(M)). The definition formalizes and provides notation for some choices left implicit in the description of the DAW(M) procedure in Section 4.1. This notation will be useful in the proofs of our main results.

Recall that the description of the DAW(M) procedure in Section 4.1 leaves some room to choose the school and surplus type to reject. We will use a sequence $\Phi : \mathbb{N} \to C \times T$ to uniquely determine the school and surplus type to reject when there is a choice. We will define whether a sequence Φ is *valid* for DAW(M), i.e., provides a school and type in each step that satisfies the requirements of DAW(M). Let the *rejection rule* denote the student who is rejected from school c under the tentative assignment μ when a student of type t must be rejected and the target type allocation is r':

$$\rho_{c,t,r'}(\mu) = \begin{cases} b_{c,t}(\mu) & \text{if } |\mu(c) \cap S_t| > r'_{c,t}, \\ \varnothing & \text{otherwise.} \end{cases}$$

We may drop the reserve vector r from the set of parameters and write $\rho_{c,t}(\mu)$ if it is clear from context.

Definition 11 (Deferred acceptance with multiple waiting rooms (DAW(M, r, Φ)))**.** DAW(M, r, Φ) *proceeds in steps k and stages i. In the initial stage, $\mu_0(s)$ is set to the top-choice school for each student s, and the step is set to $k \leftarrow 1$. Each stage i starts with an interim matching μ_i and step counter k and proceeds as follows.*[6]

1. *If μ_i is diversity-optimal (i.e. $\Delta_{c,t} = 0$ for all t, c), the process ends and outputs μ_i. Otherwise, set $\mu \leftarrow \mu_i$.*

2. *Set $(c, t) \leftarrow \Phi_k$, i.e., the school c and type t given by Φ_k, where k is the current step. Set $s \leftarrow \rho_{c,t}(\mu)$, i.e., select the next student s to reject given the school c and type t.*

3. *Set $\mu(s) \leftarrow c'$, where c' is the next highest school in the preferences of s; i.e., tentatively assign s to her next most-preferred school c'.*

[6] For expositional convenience we call all stage-starting objects μ_i, but only feasible μ_i are included in the menu of matchings.

4. *Select the next student to reject as follows:*
   Set $s' \leftarrow \rho_{c',t}(\mu)$, i.e., select the next student s' to reject given the school c' and type t. This leads to one of three cases:

   (i) If $s' = b_{c',t}(\mu) \neq \varnothing$, set $s \leftarrow s' = b_{c',t}(\mu)$, i.e., reject the bottom student of the same type t. Increment the step counter $k \leftarrow k+1$, and return to (3).

   (ii) If $s' = \varnothing$ and $\mu(s)$ is over its capacity, then $\mu(s)$ must have deficit of type t students and strict surplus of some other type. Let t' be the type given by $\Phi(k+1)$. Set $s \leftarrow \rho_{c',t'}(\mu) = b_{c',t'}(\mu)$, update the type $t \leftarrow t'$, increment the step counter $k \leftarrow k+1$, and return to (3).

   (iii) Otherwise, $s' = \varnothing$ and $\mu(s)$ is not over its total capacity, and the stage terminates. We set $\mu_{i+1} \leftarrow \mu$ and add this matching, if feasible, to our menu.[7] Finally, we increment the step $k \leftarrow k+1$, stage $i \leftarrow i+1$, and return to (1).

The three cases in (4) for the next student to reject correspond to the three cases in the description of DAW(M) in Section 4.1, if the sequence Φ is *valid* for the economy:

Definition 12 (Valid sequence for DAW(M, r, Φ))**.** *A sequence Φ is* valid *for* DAW(M) *if it satisfies the following conditions:*

- *In (2), Φ gives a school c and type t that is in surplus at c.*
- *In (4i), Φ gives a school c' and type t.*
- *In (4ii), Φ gives a school c' and a type t' that is in surplus at c'.*

We let DAW$_k(M, r, \Phi)$ be the matching of students after the kth step in Definition 11.

Throughout our proofs in the remainder of this section, we assume the existence of an additional school which each student ranks last, and which has capacity and reserve targets such that the inequalities in Assumption 2 and 3 hold at equality. This is without loss; each of these students applies to this dummy school last and thus any student matched to this school otherwise remains unmatched.

5.2 Proof of Theorem 1

The driving observation behind our proof is that each matching along the steps, and thus along any path of the menu, is more diverse pointwise. We then use this to directly establish the reserve-consistency, within-type fairness, and allocation-constrained efficiency properties of any matching μ_k in a menu. Finally, we show that, because the final matching in the menu has no school with a strict surplus of students, the matching is stable with respect to the hard-reserves choice rules (with respect to the target allocation). Because all matchings in a menu must be allocation-constrained efficient, the final matching must be identical to the hard-reserves matching as a corollary of the following lemma. In addition, the most diverse matching at the end of each path is the hard-reserves matching.

[7] The first $\frac{1}{2}\sum_{c \in C} \big||\mu_0(c)| - |c|\big|$ matchings will be infeasible, as each of their corresponding stages begins at a school with strict capacity surplus of students and ends at a school with strict capacity deficit.

Lemma 1 (Stability of DA(HR))**.** *Running student-proposing deferred acceptance with hard-reserves school choice rules results in a matching that is stable with respect to these choice rules. Furthermore, this matching is weakly preferred to any other matching that is stable with respect to the hard-reserves choice rules.*

We prove Lemma 1 in Appendix A.

Then we show that each matching μ in the output of DAW(M, r, Φ) is exactly the hard-reserves matching for some other reserve vector $r' = \text{alloc}(\mu)$. Each matching μ then inherits the hard-reserves guarantees from Proposition 1: μ is within-type fair and allocation-constrained efficient.

Lemma 2 (Monotone sequence in DAW(M, r, Φ))**.** *For a fixed valid sequence Φ, if $\mu_0, \mu_1, \ldots, \mu_n$ is the menu outputted by DAW(M, r, Φ), then for all $c \in C$ and for all types $t \in T$, the sequence $|\mu_0(c) \cap S_t|, |\mu_1(c) \cap S_t|, \ldots, |\mu_k(c) \cap S_t|, r_{c,t}$ is monotone.*

Proof. Consider the execution of DAW(M, r, Φ) as a sequence of discrete steps evaluated at individual schools. Let a *step* at school c be defined as the moment immediately following a pair of an application and rejection at c. There are three cases which change the number of type t students, $|\mu(c) \cap S_t|$:

1. A stage begins by selecting a school c with a strict surplus $|\mu(c) \cap S_t| > r_{c,t}$ of type t students and rejecting a student of type t. This rejection brings the type t allocation at c monotonically closer to the target allocation.

2. A student of type t applies to school c. If c had already met its target for type t (i.e. $|\mu(c) \cap S_t| \geq r_{c,t}$), it immediately rejects a student of the same type t, which preserves monotonicity. If c is over capacity but has a strict deficit of type t ($|\mu(c) \cap S_t| < r_{c,t}$), it accepts the type t student and rejects a student of a different type, which brings the type t allocation at c monotonically closer to the target allocation.

3. A student of type t applies to school c, which is not over capacity and has a deficit of type t students. In this case no student is rejected, and the stage terminates, which moves the allocation monotonically toward the target allocation.

Because for any school c and type t every possible step either rejects a strict surplus student, admits a strict deficit student, or results in a zero net allocation change, the sequence $|\mu_k(c) \cap S_t|$ approaches $r_{c,t}$ monotonically for all $c \in C$ and $t \in T$. □

Lemma 3 (Reserve consistency of intermediate matchings)**.** *Every intermediate matching μ_k in the menu outputted by DAW(M, r, Φ) is reserve-consistent.*

Proof. Fix an intermediate matching μ_k and suppose for the sake of contradiction that it is not reserve-consistent. Then there must exist a student s of type $t = \tau(s)$ and a school c with strict deficit of type t under μ_k (i.e. $|\mu_k(c) \cap S_t| < r_{c,t}$) such that $c \succ_s \mu_k(s)$.

The DAW(M, r, Φ) procedure begins by assigning every student to their top-choice school. Because s strictly prefers c to their current match $\mu_k(s)$, student s must have applied to c and subsequently been rejected at some step prior to reaching μ_k. Let us denote the tentative matching at that step by μ'.

Under DAW(M, r, Φ), a school c will only reject a student of type t if its allocation of type t students meets or exceeds the target reserve. Therefore, at the step when student s was rejected, it must have been true that $|\mu'(c) \cap S_t| \geq r_{c,t}$. By Lemma 2, the number of type t students at school c moves monotonically toward the target reserve $r_{c,t}$ throughout the execution of the procedure. Because the allocation at μ' was weakly greater than the reserve, the monotonicity property guarantees that the allocation can never strictly fall below $r_{c,t}$ at any subsequent step.

Thus, we must have $|\mu_k(c) \cap S_t| \geq r_{c,t}$, which contradicts our assumption that school c is in strict deficit of type t at step k. Therefore, μ_k must be reserve-consistent. □

Lemma 4 (Within-type fairness of intermediate matchings)**.** *Every intermediate matching μ_k in the menu outputted by DAW(M, r, Φ) is within-type fair.*

Proof. Suppose for the sake of contradiction that an intermediate matching μ_k is not within-type fair. Then there must exist a student s and a school c such that s strictly prefers c to their current match ($c \succ_s \mu_k(s)$), and there exists another student $s' \in \mu_k(c)$ of the same type $\tau(s) = \tau(s') = t$ who has strictly lower priority at c than s (i.e. $p_c(s) > p_c(s')$).

Recall that whenever a student is rejected in the DAW(M, r, Φ) procedure, they immediately apply to their next favorite school. Because s strictly prefers c to $\mu_k(s)$, student s must have applied to c and subsequently been rejected at some step j of the DAW(M, r, Φ) procedure prior to the completion of stage k.

According to the rejection rules of the algorithm, a school c only rejects a student of type t if c has met or exceeded its target reserve $r_{c,t}$ for that type. Therefore, at step j when s was rejected, c must have met its target allocation for type t. Furthermore, s must have been the lowest-priority student of type t at c at step j (i.e. $s = b_{c,t}(\mu_j)$).

By the monotonicity established in Lemma 2, once c meets its target reserve $r_{c,t}$ for type t, it continues to weakly meet that reserve for the rest of the procedure. Because c rejected student s at step j, any subsequent application to c by a student of type t will induce the mechanism to reject another type t student from c to preserve the type t monotonicity at c.

Since s was the lowest-priority student at step j, any remaining student of type t must have a priority strictly greater than $p_c(s)$. Furthermore, with every subsequent type t application at school c, and its corresponding rejection removing the new lowest-priority student, the priority of the bottom student of type t at school c (i.e. $p_c(b_{c,t}(\cdot))$) must be monotonically increasing across all steps following step j.

At the end of stage k, student s' is matched to c. Therefore by definition, the priority $p_c(b_{c,t}(\mu_k))$ of the bottom student of type t at matching μ_k must be weakly smaller than the priority of s'. Thus, we have:

$$p_c(s') \geq p_c(b_{c,t}(\mu_k)) \geq p_c(b_{c,t}(\mu_j)) = p_c(s)$$

This yields $p_c(s') \geq p_c(s)$, which constitutes a contradiction with our initial assumption that $p_c(s) > p_c(s')$. Therefore, the intermediate matching μ_k must be within-type fair. □

Lemma 5 (Allocation-constrained efficiency of intermediate matchings)**.** *Every intermediate matching μ_k in the menu outputted by DAW(M, r, Φ) is allocation-constrained efficient.*

Proof. Suppose for the sake of contradiction that an intermediate matching μ_k is not allocation-constrained efficient. Then there exists a within-type fair matching μ that Pareto dominates and diversity-dominates μ_k. We will show that if both dominance conditions hold, it must be that μ is not within-type fair.

The procedure begins with the assignment μ_0 in which every student is assigned to her top-choice school. Hence every student weakly prefers her assignment under μ_0 to her assignment under μ. Since μ Pareto dominates μ_k, there must be a first point in the execution of DAW(M, r, Φ) at which this property fails. Let μ^* denote the last tentative assignment before this point such that

$$\mu^*(i) \succeq_i \mu(i) \quad \text{for all } i \in S.$$

The assignment μ^* occurs strictly before reaching μ_k.

Let s be the student rejected immediately after μ^*, and let c be the school from which she is rejected. Let $t = \tau(s)$. Since this is the first rejection after which some student falls below her assignment in μ, and since students move one school at a time down their preference lists, we must have $c = \mu^*(s) = \mu(s)$.

Because s is rejected from c, school c has strict surplus of type t immediately before the rejection. Immediately after the rejection, c still weakly meets its type-t reserve. By Lemma 2, the number of type-t students at c then moves monotonically toward $r_{c,t}$ and therefore remains weakly above the reserve through μ_k. Moreover, since the rejection of s strictly decreases the number of type-t students at c, we have $|\mu^*(c) \cap S_t| \geq |\mu_k(c) \cap S_t| + 1$. Since μ diversity-dominates μ_k and μ_k weakly meets the type-t reserve at c, μ cannot assign more type-t students to c than μ_k does. Hence

$$|\mu^*(c) \cap S_t| \geq |\mu_k(c) \cap S_t| + 1 \geq |\mu(c) \cap S_t| + 1.$$

Thus μ^* contains strictly more type-t students at c than μ does. Therefore there exists a student $s' \in \mu^*(c) \cap S_t$ with $s' \notin \mu(c)$. By the definition of μ^*, we have $\mu^*(i) \succeq_i \mu(i)$ for all $i \in S$. Since $\mu^*(s') = c$ and $s' \notin \mu(c)$, it follows that $c \succ_{s'} \mu(s')$.

Finally, the rejection rule in DAW(M, r, Φ) rejects the lowest-priority student of the selected type. Since s is the type-t student rejected from c immediately after μ^*, and since $s' \in \mu^*(c) \cap S_t$, we have $p_c(s') > p_c(s)$. But $s \in \mu(c)$, s' has the same type as s, s' strictly prefers c to $\mu(s')$, and $p_c(s') > p_c(s)$. This contradicts the within-type fairness of μ.

Therefore no such matching μ exists, and μ_k is allocation-constrained efficient. □

Proposition 4 (Termination at hard-reserves)**.** *If for a fixed valid sequence Φ, the menu outputted by* DAW(M, r, Φ) *is $\{\mu_0, \mu_1, \cdots \mu_n\}$, then $\mu_n \equiv$* DA(HR(r)).

Proof. First, we demonstrate that the final matching μ_n is stable with respect to the hard-reserves choice rules. Suppose for the sake of contradiction that it is not. Then there must exist a student s and a school c such that $c \succ_s \mu_n(s)$, and s would be chosen by c under its hard-reserves choice rule, meaning $s \in Ch_c^{\text{HR}}(\mu_n(c) \cup \{s\})$.

Because μ_n is reserve-consistent, the fact that s strictly prefers c implies that school c must have already weakly met its target allocation for type $t = \tau(s)$. That is, $|\mu_n(c) \cap S_t| \geq r_{c,t}$. Furthermore, the DAW$(M, r, \Phi)$ procedure terminates only when there are no strict surplus students of any type at any school. Because μ_n is the final matching outputted by the procedure, we must have that $|\mu_n(c) \cap S_t| \leq r_{c,t}$. Combined with the reserve consistency inequality, it must be that exactly $r_{c,t}$ students of type t are matched to c under μ_n.

Because μ_n is within-type fair, no student of type t currently matched to c can have lower priority than s. Under the hard-reserves choice rule Ch_c^{HR}, the school selects at most $r_{c,t}$ students of type t with the highest priority. Since there are already $r_{c,t}$ applicants with strictly higher priority than s, the hard-reserves choice rule will reject s. This yields $Ch_c^{\text{HR}}(\mu_n(c) \cup \{s\}) = \mu_n(c)$, which contradicts our assumption that s forms a blocking pair with school c.

Therefore, μ_n must be stable with respect to the hard-reserves choice rules.

Let $\mu^{\text{HR}} := \text{DA}(\text{HR}(r))$. By Lemma 1, μ^{HR} is the student-optimal stable matching with respect to the hard-reserves choice rules. Since μ_n is also stable with respect to these choice rules, we have that μ^{HR} Pareto dominates μ_n. It remains to rule out strict Pareto dominance. By Proposition 1, μ^{HR} is within-type fair. Moreover, μ^{HR} attains the target reserve allocation r, while at termination μ_n has no strict surplus of any type at any school. Thus μ^{HR} diversity-dominates μ_n. If μ^{HR} strictly Pareto dominated μ_n, then μ^{HR} would be a within-type fair matching that both Pareto dominates and diversity-dominates μ_n, contradicting Lemma 5. Therefore μ^{HR} cannot strictly Pareto dominate μ_n.

Since μ^{HR} weakly Pareto dominates μ_n but does not strictly Pareto dominate it, the two matchings are welfare-equivalent, and so every student is matched with the same school across the two matchings. Hence the matchings are equivalent. □

6 Multi-Dimensional Types

Several policies that are intended to implement diversity goals are deliberately *intersectional*: their designs address multiple social and economic attributes simultaneously.[8] While generally understood as necessary for achieving diversity in practice, intersectionality is notoriously challenging to implement; as a result, many policies focus on achieving diversity one dimension at a time.[9] A natural extension of our model accommodates a notion of intersectionality. In this extension, each student has a type identified by a series of values, one for each attribute. As in Section 2, the designer specifies a target number of each type at each school, which in turn induces a target

[8] For more on the foundations and agenda of intersectionality, see Crenshaw (2017).

[9] This difficulty has been well documented, for instance, in Carvalho et al. (2025): "The obvious reason for the reductive approach to assessing/addressing underrepresentation is simplicity."

number of students with each attribute.

We first discuss why the model presented in Section 2 does not accommodate intersectionality in an intuitive manner, even when each possible cross-section of attributes is defined as a single type. In practice, it is natural to expect types with similar attributes to be similar in the diversity which their admission achieves; for example, a black woman is closer in her attributes to a black man than she is to a white man. Our original model does not accommodate this proximity. Further, DAW(M), which moves types pointwise toward the target along the paths it generates, does not guarantee that the number of students with a particular attribute will move closer to the induced target number.

We adapt our model to allow for natural notions of proximity across types, even while each type is treated as a disjoint set. We then show that this adaptation invites a modified version of DAW(M) which treats each attribute separately across the algorithm, and which admits an additional monotonicity axiom: along a path of matchings in a menu, the number of students with each attribute moves monotonically toward the target. We formalize the adapted model and the additional mechanism and axiom below.

Let $\mathcal{D}$ be a finite set of type dimensions. For each dimension $d \in \mathcal{D}$, let T^d be a finite set of possible values in that dimension. The multi-dimensional type space is

$$T := \prod_{d \in \mathcal{D}} T^d.$$

Each student $s \in S$ has a type vector

$$\tau(s) = \left(\tau^d(s)\right)_{d \in \mathcal{D}} \in T,$$

where $\tau^d(s) \in T^d$ denotes student s's value in dimension d. Where i is the value in dimension d, we write t_i^d. Each dimension d can be ordered and then partitioned; Figure 2 presents an example with three type dimensions and two possible values each.

Population	All students S							
(τ^1)	$(1,\cdot,\cdot)$				$(2,\cdot,\cdot)$			
(τ^1,τ^2)	$(1,1,\cdot)$		$(1,2,\cdot)$		$(2,1,\cdot)$		$(2,2,\cdot)$	
(τ^1,τ^2,τ^3)	$(1,1,1)$	$(1,1,2)$	$(1,2,1)$	$(1,2,2)$	$(2,1,1)$	$(2,1,2)$	$(2,2,1)$	$(2,2,2)$

Each bottom bin represents one full type. Here, (i,j,ℓ) denotes $\left(t_i^1, t_j^2, t_\ell^3\right)$.

Figure 2: A nested representation of a multi-dimensional type space; the figure illustrates the case of three dimensions, each with two possible values. The full population of students is first partitioned by the first type dimension, then each block is partitioned by the second, and so on.

We present an adapted version of our main mechanism, called *hierarchical* DAW(M), to preserve

monotonicity in the relevant attributes along a path, as well as in the allocation of types. Hierarchical DAW(M) requires the designer to specify an order of the type dimensions, which determines the sequence of dimensions with respect to which surplus and deficit are evaluated. Among the nested sequence of categories, the coarser categories appear "higher" in the hierarchy, and more detailed type vectors appear "lower" in the hierarchy.

The mechanism proceeds like DAW(M), except that the rejection rule is hierarchical in the following sense. If a student applies to a school at which its type is in surplus, then the school rejects the least-priority student with that same type vector. If the type is in deficit, the mechanism chooses the type from which to reject by first moving "upward" in the hierarchy until it finds the first coarser category that is in surplus. Then, the mechanism moves back "downward" within that category to find a surplus type vector; from that type the least-priority student is rejected. A visualization of this procedure is in Figure 3.

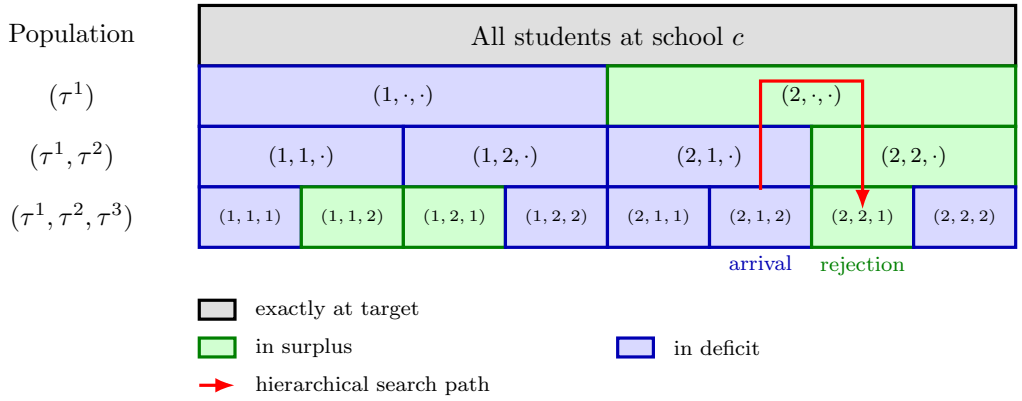


Figure 3: Illustration of the hierarchical rejection rule at a school c. As in Figure 2, (i,j,ℓ) denotes $\left(t_i^1, t_j^2, t_\ell^3\right)$. The bottom row shows the full type vectors at school c with five deficit types (blue) and three surplus types (green). Suppose a student of type $\left(t_2^1, t_1^2, t_2^3\right)$ arrives at school c; this is the blue bin labeled "arrival." Because that type is in deficit, the mechanism does not reject from that bin. The mechanism first moves upward to the parent block $\left(t_2^1, t_1^2, \cdot\right)$ in the second row; since that block is not in surplus, the mechanism continues upward to the block $\left(t_2^1, \cdot, \cdot\right)$ in the first row, which is the first block above the arriving type that is in surplus. It then moves downward within that block and identifies the finer block $\left(t_2^1, t_2^2, \cdot\right)$ in the second row, which is also in surplus. Finally, it moves down again to a surplus type contained in that block, here $\left(t_2^1, t_2^2, t_1^3\right)$, and rejects the least-priority student from that bottom-row type. The red arrows trace this search path.

Each rejection is chosen so that the assignment moves toward the target not only for the full type vectors, but also for the coarser parent categories. We call this *category monotonicity*: along the path, no school moves away from its induced target for any category generated by the ordered type dimensions. To formalize the definition, fix an ordering of the type dimensions and relabel them so that $\mathcal{D} = \{1, \ldots, D\}$. For any full type $t = \left(t^1, \ldots, t^D\right) \in T$ and any $m \in \{1, \ldots, D\}$, write

$$t^{\leq m} := \left(t^1, \ldots, t^m\right)$$

for the category of depth m containing t. For each $m \in \{0, \ldots, D\}$, define

$$T_{t^{\leq m}} := \left\{ \tilde{t} \in T : \tilde{t}^d = t^d \text{ for every } d \leq m \right\}.$$

When $m = 0$, we say $T_{t^{\leq 0}} = T$. Thus $T_{t^{\leq m}}$ is the set of full type vectors contained in the category $t^{\leq m}$. Define the deviation of this category from its target at school c by

$$\Delta_{c,t^{\leq m}}(\mu, r) := \sum_{\tilde{t} \in T_{t^{\leq m}}} \Delta_{c,\tilde{t}}(\mu, r).$$

When $m = D$, this definition reduces to the original type-level deviation $\Delta_{c,t}(\mu, r)$; when $m = 0$, it gives $\Delta_{c,t^{\leq 0}}(\mu, r) = |\mu(c)| - |c|$.

Definition 13 (Category monotonicity)**.** *Fix an ordering of the type dimensions. A path* $\left(\mu^0, \mu^1, \ldots, \mu^K\right)$ *satisfies* category monotonicity *if, for every* $k = 0, 1, \ldots, K-1$*, every school* $c \in C$*, every full type* $t \in T$*, and every* $m \in \{1, \ldots, D\}$*, either*

$$0 \leq \Delta_{c,t^{\leq m}}\left(\mu^{k+1}, r\right) \leq \Delta_{c,t^{\leq m}}\left(\mu^k, r\right)$$

or

$$0 \geq \Delta_{c,t^{\leq m}}\left(\mu^{k+1}, r\right) \geq \Delta_{c,t^{\leq m}}\left(\mu^k, r\right).$$

Each stage of Hierarchical DAW(M) proceeds as follows. If the interim matching μ_i at stage i is diversity-optimal, then the mechanism ends. Otherwise, it selects some type t in surplus at some school c such that all parent types are also in surplus, and rejects the least-preferred student of this type at c. The rejected student s then applies to her next most-preferred school c'. Let $\tilde{\mu}$ denote the tentative matching after s is added to c', and let $t = \tau(s)$. There are three cases.

(i) If student s contributes to the surplus of her full type at c', then c' rejects its least-priority student of type t.

(ii) If student s does not contribute to the surplus of her full type at c', but contributes to the surplus of a coarser category at c', let

$$m^* := \max\left\{ m \in \{0, \ldots, D-1\} : \Delta_{c',t^{\leq m}}\left(\tilde{\mu}, r\right) > 0 \right\}.$$

The school then moves downward within the category $t^{\leq m^*}$, choosing at each lower depth a subcategory that is in surplus, until it reaches a full type $t' \in T$ such that

$$t' \in T_{t^{\leq m^*}} \qquad \text{and} \qquad \Delta_{c',t'}\left(\tilde{\mu}, r\right) > 0.$$

School c' rejects its least-priority student of type t'.

(iii) Otherwise, the stage ends in the next interim matching: $\mu_{i+1} := \tilde{\mu}$.

Thus when an arriving student would make her own full type in surplus, the rule rejects within that same full type. When her full type remains weakly below target, the rule searches upward for the nearest parent category that is in surplus, and then searches downward inside that parent category for a surplus full type from which to reject.

Theorem 5. *Hierarchical* DAW(M) *finds an ordered menu of matchings that are allocation-constrained efficient, within-type fair, and reserve-consistent. The menu is ordered: matchings earlier in the ordering Pareto dominate and are diversity dominated by matchings later in the order; additionally, matchings along a path satisfy category monotonicity. The diversity-maximal endpoint of this menu is the hard-reserves matching.*

Hierarchical DAW(M) is a specific implementation of DAW(M); it constrains the set of types from which students can be rejected. Hence allocation-constrained efficiency, within-type fairness, and reserve consistency axioms, as well as the hard-reserves endpoint, follow from Theorem 1. Category monotonicity follows directly from the construction of the procedure.

7 Discussion

Our primary contributions are *menus* of matchings which satisfy normative axioms. Previous literature has primarily focused on the limitations of quota- and reserve-based policies, or welfare subject to a fixed distributional target. In practice, the preferred matching in a given school choice problem can be highly allocation-specific. Our results reflect our belief that the designer, for any objective she has over diversity and welfare, should be able to choose the matching that optimally trades off her objectives, while also guaranteeing appealing, explainable, and verifiable axioms. We achieve these menus using *rejection rules*, which we believe to be novel in this setting. Our approach provides schools with the freedom to reject students without knowledge of the full set of students whom they will eventually reject.

We now speak to the practical implementation of our mechanisms. Each DAW mechanism which we have described proceeds in stages. The stages correspond to either "chains," which start at a surplus type in one school and end at (weak) deficit type in another school, or "cycles", where a student replaces the student which initiated the series of rejections. Both the DAW(M) and DAW(S) mechanisms (and their adaptations) only run chains until the interim matching is feasible, and then they only run cycles. When the mechanisms switch from chains to cycles, the district then can move between feasible matchings in a way that decreases welfare at the benefit of diversity: either pointwise (DAW(M)) or in the ℓ_1 norm (DAW(S)).

These cycles correspond to the stages described in Section 4. In effect, the designer is then choosing cycles containing the least-preferred students at schools and swapping them along the edges. Thus the construction presented here is parallel to *top trading cycles*, but instead of students pointing at a more-preferred school to which they would like to be matched, they instead point to a *less-preferred* school (their next most-preferred school) at which they could feasibly be admitted.

Hence a natural way to implement our DAW mechanisms is for the district to select a set of cycles, which might be called *bottom trading cycles*, in a way that maximizes her objective.

The practical tool of bottom trading cycles gives a helpful visual for how our procedure works: schools swap students to improve diversity while minimizing fairness violations and/or welfare loss. We thus believe that, while bottom trading cycles are isomorphic to the procedure described in this paper, they present a fruitful opportunity for implementing distributional goals across a school district.

We have taken seriously the objective of describing matchings that lie on the welfare-diversity frontier for a given set of axioms. At the same time, we have been agnostic to how the designer chooses the matching from the frontier. Indeed, how this choice is made will affect whether the DAW mechanism is strategyproof. Further, the existence of cycles relies on complete preferences — a common assumption in school choice and which can be imputed on students given their characteristics. In light of this, the completions of preferences must be "fair", in the sense that how preferences are made complete is independent of student types. Future work should consider under what commitments of the principal our mechanisms are incentive-compatible.

References

Abdulkadiroğlu, A. (2005). College admissions with affirmative action. *International Journal of Game Theory*, 33(4):535–549.

Abdulkadiroğlu, A. and Grigoryan, A. (2021). Priority-based assignment with reserves and quotas. Technical report, National Bureau of Economic Research.

Abdulkadiroglu, A. and Sonmez, T. (2003). School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3):729–747.

Aziz, H., Gaspers, S., and Sun, Z. (2020). Mechanism design for school choice with soft diversity constraints. In *AAMAS*, pages 1756–1758.

Carvalho, J.-P., Pradelski, B. S. R., and Williams, C. (2025). Intersectionality: Affirmative action with multidimensional identities. *Management Science*, 71(5):4495–4509.

Crenshaw, K. W. (2017). *On Intersectionality: Essential Writings*. The New Press, New York.

Echenique, F., Mekonnen, T., and Yenmez, M. B. (2025). Diversity in Choice as Majorization. arXiv:2407.17589 [econ].

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

Ehlers, L., Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. (2014). School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153:648–683.

Erdil, A. and Kumano, T. (2019). Efficiency and stability under substitutable priorities with ties. *Journal of Economic Theory*, 184.

Gale, D. and Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9.

Hafalir, I. E., Kojima, F., Yenmez, M. B., and Yokote, K. (2022). Design on matroids: Diversity vs. meritocracy.

Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. (2013). Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/TE1135.

Hatfield, J. W. and Milgrom, P. R. (2005). Matching with Contracts. *American Economic Review*, 95(4):913–935.

Henig, J. R. (1994). *Rethinking School Choice: Limits of the Market Metaphor*. Princeton University Press.

Imamura, K. (2025). Meritocracy versus diversity. *Journal of Economic Theory*, page 106047.

Kesten, O. (2010). School Choice with Consent. *The Quarterly Journal of Economics*, 125(3):1297–1348.

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

Kominers, S. D. and Sönmez, T. (2013). Designing for diversity in matching. In *EC*, pages 603–604.

Kurata, R., Hamada, N., Iwasaki, A., and Yokoo, M. (2017). Controlled school choice with soft bounds and overlapping types. *Journal of Artificial Intelligence Research*, 58:153–184.

Nguyen, T. and Vohra, R. (2019). Stable matching with proportionality constraints. *Operations Research*, 67(6):1503–1519.

Sönmez, T. and Yenmez, M. B. (2019). *Affirmative action with overlapping reserves*. Boston College.

Sun, Z. and Yokoo, M. (2025). Achieving balanced representation in school choice with diversity goals. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 14129–14138.

Çelebi, O. and Flynn, J. P. (2025). Adaptive Priority Mechanisms.

A Hard and Soft Reserves

We begin with an example that adapts Example 2. The soft-reserves matching suffers from inefficiency that arises from possible improvement cycles that exist both across and within types. The hard-reserves matching only bears the possibility for within-type improvement cycles; such cycles are inevitable when requiring within-type fairness. The improvement cycles are all of the type mentioned in Kesten (2010).

Example 4. Consider an economy with 100 students of type x, labeled $x_0, \ldots, x_{99}$, and 100 students of type y, labeled $y_0, \ldots, y_{99}$. There are 100 schools labeled $c_1, c_2, \ldots, c_{100}$, each reserving a single seat for each type.

Throughout this example, x_0 and y_0 are special students. All other student indices are interpreted cyclically over $\{1, \ldots, 99\}$; for any integer m, let

$$\langle m \rangle := 1 + ((m-1) \bmod 99).$$

Preferences of each student are as follows:

- $x_0 : c_1 \succ_{x_0} c_2 \succ_{x_0} \cdots \succ_{x_0} c_{100}$.
- $x_i : c_i \succ_{x_i} c_{\langle i+1 \rangle} \succ_{x_i} \cdots \succ_{x_i} c_{\langle i+98 \rangle} \succ_{x_i} c_{100}$ for $i \in \{1, \ldots, 99\}$; that is, their preferences are cyclical over $c_1, \ldots, c_{99}$, with c_{100} ranked last.
- $y_0 : c_{100} \succ_{y_0} c_{99} \succ_{y_0} \cdots \succ_{y_0} c_1$.
- $y_i : c_{101-\langle i \rangle} \succ_{y_i} c_{101-\langle i+1 \rangle} \succ_{y_i} \cdots \succ_{y_i} c_{101-\langle i+98 \rangle} \succ_{y_i} c_1$ for $i \in \{1, \ldots, 99\}$; that is, their preferences are cyclical over $c_{100}, c_{99}, \ldots, c_2$, with c_1 ranked last.

School priorities are given within each type as follows (comparisons across types can be completed arbitrarily).

- $p_{c_1}(y_2) > p_{c_1}(y_3) > \cdots > p_{c_1}(y_{99}) > p_{c_1}(y_1) > p_{c_1}(y_0)$.
- $p_{c_{100}}(x_2) > p_{c_{100}}(x_3) > \cdots > p_{c_{100}}(x_{99}) > p_{c_{100}}(x_1) > p_{c_{100}}(x_0)$.
- $p_{c_i}(x_0) > p_{c_i}(x_{\langle i+1 \rangle}) > p_{c_i}(x_{\langle i+2 \rangle}) > \cdots > p_{c_i}(x_{\langle i+98 \rangle}) > p_{c_i}(x_i)$ for $i \in \{1, \ldots, 99\}$.
- $p_{c_{101-i}}(y_0) > p_{c_{101-i}}(y_{\langle i+1 \rangle}) > p_{c_{101-i}}(y_{\langle i+2 \rangle}) > \cdots > p_{c_{101-i}}(y_{\langle i+98 \rangle}) > p_{c_{101-i}}(y_i)$ for $i \in \{1, \ldots, 99\}$.

Observe that the top-choice matching, where each student matches to her most-preferred school, is feasible: $\mu(c_1) = \{x_0, x_1\}$, $\mu(c_{100}) = \{y_0, y_1\}$, and $\mu(c_i) = \{x_i, y_{101-i}\}$ for $2 \leq i \leq 99$.

The hard-reserves matching is $\mu^{HR}(c_1) = \{x_0, y_2\}$, $\mu^{HR}(c_{100}) = \{x_2, y_0\}$, and $\mu^{HR}(c_i) = \{x_{\langle i+1 \rangle}, y_{\langle 102-i \rangle}\}$ for $2 \leq i \leq 99$. This matching improves diversity at just two of the 100 schools, but moves 196 students to the school that is second to last in their preferences and moves two students, x_2 and y_2, to their least-preferred school.

The structure of this example is that x_0 ranks c_1 first and has highest priority among type-x students at c_1, while every type-x student ranks school c_{100} last. Since x_0 occupies the single type-x reserve at c_1, the remaining 99 type-x students compete for the 98 type-x seats at $c_2, \ldots, c_{99}$ before any of them uses c_{100}. Thus, because schools have cyclic priorities that oppose the students'

preferences, the hard-reserves stable matching for type-x students is unique. In this example, the priority of x_0 at c_1 prevents any welfare improvement of type x students. Symmetrical structure holds for type y students. The main difference between these two improvement cycles and the cycle from Example 3 is that here, the type of the student who prevents the improvement cycle from occurring is the *same* as the type of the students in the cycle itself. □

A.1 Formalizing the Hard- and Soft-Reserves Choice Rules

Fix an economy with reserves $(\mathcal{E}, r)$. For each school $c \in C$, let

$$Ch_{c,r} : 2^S \to 2^S$$

be a *choice rule* satisfying $Ch_{c,r}(X) \subseteq X$ and $|Ch_{c,r}(X)| \leq |c|$ for every $X \subseteq S$. We interpret $Ch_{c,r}(X)$ as the set of students that c would (tentatively) hold from the applicant set X given reserve vector r.

Definition 14 (Student-proposing deferred acceptance with a round-contingent choice rule)**.** *Given $(\mathcal{E}, r)$ and school choice rules $(Ch_{c,r})_{c \in C}$, define an iterative procedure indexed by rounds $k = 0, 1, 2, \ldots$.*

Let μ^k denote the tentative matching after round k. For each student s, let $R^k(s) \subseteq C$ denote the set of schools that have rejected s by the end of round k. Set $\mu^0(s) = \varnothing$ for all $s \in S$ and $\mu^0(c) = \varnothing$ for all $c \in C$, and set $R^0(s) = \varnothing$ for all $s \in S$.

Then in round $k+1$:

(i) Each student s with $\mu^k(s) = \varnothing$ and $C \setminus R^k(s) \neq \varnothing$ applies to her most-preferred school among those that have not rejected her:

$$a^{k+1}(s) \in \arg\max_{\succ_s} \big(C \setminus R^k(s)\big).$$

For each school $c \in C$, let

$$A_c^{k+1} := \{s \in S : \ a^{k+1}(s) = c\}$$

denote the set of students who apply to c in round $k+1$.

(ii) Each school c considers the set

$$X_c^{k+1} := A_c^{k+1} \cup \mu^k(c)$$

consisting of its new applicants together with the students it was tentatively holding from the previous round, and tentatively holds

$$\mu^{k+1}(c) := Ch_{c,r}(X_c^{k+1}).$$

Define the set of students rejected by c in round $k+1$ as

$$\text{Rej}_c^{k+1} := X_c^{k+1} \setminus \mu^{k+1}(c).$$

(iii) For each student $s \in S$, update

$$R^{k+1}(s) := R^k(s) \cup \{c \in C : \ s \in \text{Rej}_c^{k+1}\}.$$

Define students' tentative assignments consistently with schools' holdings: for each $s \in S$,

$$\mu^{k+1}(s) = c \iff s \in \mu^{k+1}(c), \qquad \mu^{k+1}(s) = \varnothing \iff s \notin \bigcup_{c \in C} \mu^{k+1}(c).$$

The procedure terminates at the first round K such that $\text{Rej}_c^K = \varnothing$ for all $c \in C$. The output matching is $\mu := \mu^K$.

The deferred acceptance algorithm presented here is defined with respect to types; if it were independent of types, then we have the deferred acceptance algorithm as in Gale and Shapley (1962).

We now introduce the two choice rules from Section 3 that reflect natural ways to treat reserves: *hard reserves* and *soft reserves*. Under hard reserves, each school c chooses to tentatively admit the highest priority $r_{c,t}$-students of type t. No seat of type t is ever used to enroll a student of another type t'.

Definition 15 (Hard-reserves choice rule)**.** *Fix $(\mathcal{E}, r)$ and a school $c \in C$. For any set of students $X \subseteq S$, define $Ch_{c,r}^{\text{HR}}(X) \subseteq X$ as follows:*

(i) For each $t \in T$, define

$$H_{c,t}(X) \in \arg \max_{Y \subseteq X \cap S_t: \ |Y| \leq r_{c,t}} \sum_{s \in Y} p_c(s).$$

(ii) Let

$$H_c(X) := \bigcup_{t \in T} H_{c,t}(X), \qquad Ch_{c,r}^{\text{HR}}(X) := H_c(X).$$

Another way to treat reserves is to let each school c tentatively admit the highest priority $r_{c,t}$-students of type t, but when capacity is not yet met, any type-t seat which is unfilled can be given to a student of type t'. This characterizes the soft-reserves choice rule.

Definition 16 (Soft-reserves choice rule)**.** *Fix $(\mathcal{E}, r)$ and a school $c \in C$. For any set of students $X \subseteq S$, define $Ch_{c,r}^{\text{SR}}(X) \subseteq X$ as follows:*

(i) For each $t \in T$, define

$$H_{c,t}(X) \in \arg \max_{Y \subseteq X \cap S_t: \ |Y| \leq r_{c,t}} \sum_{s \in Y} p_c(s).$$

Let

$$H_c(X) := \bigcup_{t \in T} H_{c,t}(X).$$

(ii) Define the leftover number of seats

$$\ell_c(X) := |c| - |H_c(X)| \geq 0,$$

and define

$$F_c(X) \in \arg \max_{Y \subseteq X \setminus H_c(X):\ |Y| \leq \ell_c(X)} \sum_{s \in Y} p_c(s).$$

For each type we also write $F_{c,t}(X) := F_c(X) \cap S_t$.

(iii) Let

$$Ch^{\mathrm{SR}}_{c,r}(X) := H_c(X) \cup F_c(X).$$

We use the formal descriptions of these choice rules to prove Propositions 1–3.

A.2 Proofs of Propositions 1–3

We first observe that students are substitutes for both choice rules according to the following definition from Hatfield and Milgrom (2005):

Definition 17 (Substitutes (Hatfield and Milgrom, 2005))**.** *Students in a set S are substitutes for a school c with a choice rule Ch_c, if for any subsets $X' \subseteq X'' \subseteq S$ we have $X' \setminus Ch_c(X') \subseteq X'' \setminus Ch_c(X'')$.*

Under hard reserves, the claim follows directly; under soft reserves we have to additionally account for a match occurring with a student in the waiting room.

Lemma 6 (Substitutability under hard reserves)**.** *Students are substitutes for any school c with a hard-reserves choice rule $Ch^{\mathrm{HR}}_{c,r}$.*

Proof. Consider subsets $X' \subseteq X'' \subseteq S$. Student $s \in X$ is rejected under $Ch^{\mathrm{HR}}_{c,r}$ with a set of applicants X if only if there are at least $r_{c,\tau(s)}$ students from type $\tau(s)$ with higher priority at c than s. Observe $s \in X' \setminus Ch^{\mathrm{HR}}_{c,r}(X') \implies |\{s' \in X' \mid p_c(s') > p_c(s) > 0 \text{ and } \tau(s') = \tau(s)\}| \geq r_{c,\tau(s)}$. However, $\{s' \in X' \mid p_c(s') > p_c(s) > 0 \text{ and } \tau(s') = \tau(s)\} \subseteq \{s' \in X'' \mid p_c(s') > p_c(s) > 0 \text{ and } \tau(s') = \tau(s)\} \implies s \in X'' \setminus Ch^{\mathrm{HR}}_{c,r}(X'')$. □

Lemma 7 (Substitutability under soft reserves)**.** *Students are substitutes for any school c with a soft-reserves choice rule $Ch^{\mathrm{SR}}_{c,r}$.*

Proof. Consider subsets $X' \subseteq X'' \subseteq S$. Note $s \in X' \setminus Ch^{\mathrm{SR}}_{c,r}(X') \implies s \notin H_{c,\tau(s)}(X')$ and $s \notin F_c(X')$. Then $s \notin H_{c,\tau(s)}(X') \implies |\{s' \in X' \mid p_c(s') > p_c(s) > 0 \text{ and } \tau(s') = \tau(s)\}| \geq r_{c,\tau(s)}$, which then implies that $s \notin H_{c,\tau(s)}(X'')$. This also implies that $X' \setminus H_c(X') \subseteq X'' \setminus H_c(X'')$.

It remains to show that $s \notin F_c(X'')$. As $s \in X' \setminus H_{c,\tau(s)}(X')$ and $s \notin F_c(X')$, it must be that $|F_c(X')| = \ell_c(X')$. Therefore at least $\ell_c(X')$ students in $X' \setminus H_c(X')$ have higher priority than s; by the inclusion above, those same students belong to $X'' \setminus H_c(X'')$; moreover, for every type t', $|H_{c,t'}(X')| \leq |H_{c,t'}(X'')|$ and thus $\ell_c(X'') = |c| - |H_c(X'')| \leq |c| - |H_c(X')| = \ell_c(X')$. Thus in $X'' \setminus H_c(X'')$ there are at least as many students with higher priority than s as there are seats available in $F_c(X'')$, hence $s \notin F_c(X'')$. □

We then observe that the matchings are stable with respect to their choice rules as in Theorem 3 of Hatfield and Milgrom (2005). Each follows directly and we omit proof. For shorthand, we write DA(HR) and DA(SR) as the matchings from Definition 14 with the hard- and soft-reserves choice rules from Definitions 15 and 16 respectively.

Lemma 8 (Restatement of Lemma 1)**.** *Running student-proposing deferred acceptance with hard-reserves school choice rules results in a matching that is stable with respect to these choice rules. Furthermore, this matching is weakly preferred to any other matching that is stable with respect to the hard-reserves choice rules.*

Lemma 9 (Stability of DA(SR))**.** *Running student-proposing deferred acceptance with soft-reserves school choice rules results in a matching that is stable with respect to these choice rules. Furthermore, this matching is weakly preferred to any other matching that is stable with respect to the soft-reserves choice rules.*

Additionally, under our assumptions, the seats have no waste.

Lemma 10 (Non-wastefulness of DA(HR) and DA(SR))**.** *Under Assumptions 1-3,* DA(SR) *and* DA(HR) *both assign every seat at every school.*

We provide a sketch of the proof of Lemma 10. By Lemma 1, $\mu =$ DA(HR) is stable with respect to the hard-reserves choice rules. If $|\mu(c)| < |c|$ for some school c, there must be some strict deficit type t at that school c. Given that not all seats allocated for type t students are taken under μ, under Assumption 2 there must be some student s of type t for whom $\mu(s) = \varnothing$. However s and c would block matching μ, which would be a contradiction.

In the case of the soft-reserves choice rule, if $|\mu(c)| < |c|$ for some school c, Assumption 3 implies that there must be some student s of type t, for whom $\mu(s) = \varnothing$. Then s and c would block matching μ, which would be a contradiction.

We now can prove our main results of this section. Throughout the proofs there is an implied reserve vector r; at times we explicitly include it in the choice rule, e.g. DA(HR(r)).

Proof of Proposition 1. By Lemma 1, $\mu =$ DA(HR) is stable with respect to the hard-reserves choice rules.

Within-type fairness: Assume for the sake of contradiction that there exist students s and s' from type t for which $c \succ_s \mu(s)$ and $p_c(s) > p_c(s')$. Then for $X = \mu(c)$ we have that $\sum_{s'' \in X \cap S_t} p_c(s'') < p_c(s) + \sum_{s'' \in X \cap S_t \setminus \{s'\}} p_c(s'')$ which implies that $H_{c,t}(X \cup \{s\}) \neq H_{c,t}(X)$, which contradicts.

Across-type fairness with respect to reserves: Under the hard-reserves choice rule, for no school c can there be a type t for which $\Delta_{c,t}(\mu, r) > 0$; hence across-type fairness with respect to the reserves is satisfied.

Allocation-constrained efficiency: Suppose μ is not allocation-constrained efficient. This means that there is a matching μ' that is within-type fair and both diversity-dominates and Pareto dominates μ. Note that for all types t and schools c we have $\Delta_{c,t}(\mu, r) \leq 0$. Furthermore, given that μ' diversity-dominates μ, $\Delta_{c,t}(\mu, r) = 0 \implies \Delta_{c,t}(\mu', r) = 0$, which means that no type is at surplus at any school under μ'. Under Assumption 1 all schools are acceptable and thus no student that is matched under μ can be left unmatched under μ'. Additionally, if a student of type t is left unmatched under $\mu = \text{DA(HR)}$, that student must have applied to all schools and been rejected which implies that no school c can be at deficit of type t under μ. Together, these observations imply that $\text{alloc}(\mu) = \text{alloc}(\mu')$. However, this would result in a contradiction as μ' must be stable with respect to the $Ch^{\text{HR}}_{c,r}$ choice rules, and by Lemma 1 $\mu = \text{DA(HR)}$ must Pareto dominate any other stable matching.

Reserve consistency: Suppose that for a student s with $t = \tau(s)$ there exists a school c with $c \succ_s \mu(s)$ for which $|\mu(c) \cap S_t| < r_{c,t}$. Then clearly $|(\mu(c) \cap S_t) \cup \{s\}| \leq r_{c,t}$ and $\sum_{s'' \in X \cap S_t} p_c(s'') < p_c(s) + \sum_{s'' \in X \cap S_t} p_c(s'')$ and thus $H_{c,t}(\mu(c) \cup \{s\}) \neq H_{c,t}(\mu(c))$ which contradicts stability with respect to the hard-reserves choice rule. □

Proof of Proposition 2. By Lemma 9, $\mu = \text{DA(SR)}$ is stable with respect to the schools' soft-reserves choice rules.

Within-type fairness: The argument is similar to that for hard-reserves choice rules. Assume for the sake of contradiction that there exist students s and s' from type t for which $c \succ_s \mu(s)$ and $p_c(s) > p_c(s')$. If there exists a student $s'' \in H_{c,t}(\mu(c))$ for which $p_c(s) > p_c(s'')$, then we would have that $s \in H_{c,t}(\mu(c) \cup \{s\})$. Otherwise, it must be the case that $s' \in F_{c,t}(\mu(c))$ and then we can obtain a contradiction by noting that $\sum_{q \in F_{c,t}(\mu(c))} p_c(q) < p_c(s) + \sum_{q \in F_{c,t}(\mu(c)) \setminus \{s'\}} p_c(q)$ and so $s \in F_{c,t}(\mu(c) \cup \{s\})$.

Across-type fairness with respect to reserves: Suppose for the sake of contradiction that there is a student s such that, for some school $c \succ_s \mu(s)$, there is a student $s' \in \mu(c)$ where $p_c(s) > p_c(s')$ and where $\Delta_{c,\tau(s')}(\mu, r) > 0$. Together $p_c(s) > p_c(s')$ and $\Delta_{c,\tau(s')}(\mu, r) > 0$ imply that $\exists s'' \in F_c(\mu(c)) \cap S_{\tau(s')} : p_c(s) > p_c(s'')$. If $s \notin H_{c,t}(\mu(c) \cup \{s\})$ we should have that $s \in F_c(\mu(c) \cup \{s\})$, because $\sum_{q \in F_c(\mu(c))} p_c(q) < p_c(s) + \sum_{q \in F_c(\mu(c)) \setminus \{s''\}} p_c(q)$, thus a contradiction.

Reserve consistency: The same argument as the one for hard-reserves applies here. If there exists a student s with $t = \tau(s)$ and a school $c \succ_s \mu(s)$ for which $|\mu(c) \cap S_t| < r_{c,t}$, then $H_{c,t}(\mu(c) \cup \{s\}) = \{s\} \cup H_{c,t}(\mu(c))$. This implies $s \in Ch^{\text{SR}}_{c,r}(\mu(c) \cup \{s\})$, which contradicts stability with respect to the soft-reserves choice rule.

□

Proof of Proposition 3. Let us use the following notation: $\mu' = \text{DA}(\text{SR})(r)$, $r' = \text{alloc}(\mu')$, and $\mu = \text{DA}(\text{HR})(r')$. Within-type fairness follows from Proposition 1 applied to reserves r'.

Pareto dominating soft reserves: From Proposition 2 the soft-reserves matching μ' is within-type fair. The fact that μ Pareto dominates μ' follows from the within-type fairness of μ' combined with the argument for allocation-constrained efficiency of hard-reserves in Proposition 1. To see why this is true, note that μ' is stable with respect to $Ch^{\text{HR}}_{c,r'}, \forall c \in C$. This is because the existence of a student s and school $c \succ_s \mu'(s)$ for which $Ch^{\text{HR}}_{c,r'}(\mu'(c) \cup \{s\}) \neq Ch^{\text{HR}}_{c,r'}(\mu'(c))$ is equivalent to $H_{c,\tau(s)}(\mu'(c) \cup \{s\}) \neq H_{c,\tau(s)}(\mu'(c))$. The latter implies that $\exists s' \in H_{c,\tau(s)} : p_c(s') < p_c(s)$ which would contradict the within-type fairness of μ'. Thus μ' is stable with respect to the schools' hard-reserves choice rules, and the claim follows from Lemma 1.

Reserve consistency: By Proposition 2, μ' is reserve-consistent and so $c \succ_s \mu'(s) \implies |\mu'(c) \cap S_{\tau(s)}| \geq c_{\tau(s)}$. Because μ Pareto dominates μ', we have $c \succ_s \mu(s) \implies c \succ_s \mu'(s)$, thus $c \succ_s \mu(s) \implies |\mu'(c) \cap S_{\tau(s)}| \geq c_{\tau(s)}$ hence μ is reserve-consistent.

Allocation-constrained efficiency: We will show that if there is a within-type fair matching μ'' that Pareto dominates and diversity-dominates μ, then $\text{alloc}(\mu'') = \text{alloc}(\mu)$. Note that μ'' must also Pareto dominate μ', which is Pareto dominated by μ. If for some school c and type t we have that the deficit under μ'' is strictly smaller than the deficit under μ, i.e. $r_{c,t} \geq |\mu''(c) \cap S_t| > |\mu(c) \cap S_t|$, then we would obtain a contradiction. This is because at least one student s of type t would be matched to c under μ'' but not under μ', and because μ'' Pareto dominates μ', we will get that s blocks the soft-reserves matching μ' because type t is at deficit at c under μ' (by construction $\text{alloc}(\mu') = \text{alloc}(\mu)$). Therefore we have that for any school c and type t, $|\mu''(c) \cap S_t| \leq |\mu(c) \cap S_t|$. We then note that because every student who is matched under μ' must also be matched under μ'', we obtain $\text{alloc}(\mu'') = \text{alloc}(\mu)$. From here, the same reasoning in Proposition 1 applies and we obtain that μ must be allocation-constrained efficient. □

A.3 Deferred Acceptance with Waiting Rooms and Activations

We generalize the DAW(M) procedure by adding "activation" thresholds $a_{c,t}$ to the rejection rules for each type at each school. The rejection rule for type t at school c becomes active once, at some point in the execution, school c tentatively holds at least $a_{c,t}$ students of type t; and once active, the rejection rule remains active for the rest of the algorithm. The activation thresholds we consider are weakly higher than the reserves for their corresponding type and school, i.e. $a_{c,t} \geq r_{c,t} \, \forall c \in C, \forall t \in T$.

We use the activation vector to generate a menu on which a given within-type fair, allocation-constrained efficient, and reserve-consistent matching exists. We will set the activation vector such that $a = \max(\text{alloc}(\mu), r)$, and show that a matching generated by a procedure with this activation vector is likewise achieved by the DAW(M) procedure with reserves equal to its allocation vector $\text{alloc}(\mu)$, hence the properties of DAW(M) carry over here.

As before, we will use a sequence $\Phi : \mathbb{N} \to C \times T$ to uniquely determine the school and surplus type to reject when there is a choice. We will again define whether a sequence Φ is *valid* in the

same sense as for DAW(M).

The definition of the rejection rule is the same. For any assignment μ, let

$$\rho_{c,t,r}(\mu) = \begin{cases} b_{c,t}(\mu) & \text{if } |\mu(c) \cap S_t| > r_{c,t}, \\ \varnothing & \text{otherwise,} \end{cases}$$

where $b_{c,t}(\mu)$ is the lowest-priority type-t student assigned to c under μ; as before, we sometimes suppress r when clear from the context.

We write DAW(M, a, r, Φ) for the mechanism with activation vector $a \geq r$, where the inequality is pointwise. The procedure prioritizes initiating student rejection chains that begin with a strict surplus student whose type rejection rule is active at their current school.

Definition 18 (Deferred acceptance with multiple waiting rooms and activations)**.** DAW(M, a, r, Φ) *proceeds in steps k and stages i. In the initial stage, $\mu_0(s)$ is set to the top-choice school for each student s, and the step is set to $k \leftarrow 1$. For each school c, any type t is set to be active for the entire algorithm if $|\mu_0(c) \cap S_t| \geq a_{c,t}$. Each stage i starts with an interim matching μ_i and step counter k, and proceeds as follows.*

1. *If μ_i is diversity-optimal (i.e. $\Delta_{c,t} = 0$ for all t, c), the process ends and outputs μ_i. Otherwise, set $\mu \leftarrow \mu_i$.*

2. *Set $(c, t) \leftarrow \Phi_k$, i.e., consider the school c and type t given by Φ_k, where k is the current step. If type t is active at school c, set $s \leftarrow \rho_{c,t}(\mu)$, i.e., select the next student s to reject given the school c and type t which must be active at c.*

3. *Set $\mu(s) \leftarrow c'$ where c' is the next most-preferred school of student s, i.e., tentatively assign s to her next most-preferred school c'.*
   If $|\mu(c') \cap S_t| \geq a_{c',t}$ make type t active at school c' for the rest of the algorithm.

4. *Select the next student to reject as follows:*

   (i) If type t was active prior to this step at school c' and $\rho_{c',t,r}(\mu) \neq \varnothing$, set $s \leftarrow \rho_{c',t,r}(\mu)$, i.e., if the same-type rejection rule is active and type t is currently in strict surplus at c', reject the bottom-priority type-t student at c'. Increment the step counter $k \leftarrow k + 1$ and return to (3).

   (ii) If type t was active prior to this step at school c', $\rho_{c',t,r}(\mu) = \varnothing$, and $|\mu(c')| > |c'|$, then the application did not create a strict same-type surplus, but school c' is over capacity.[10] Let $\Phi_{k+1} = (c', t')$ select an active type t' at c' with $\rho_{c',t',r}(\mu) \neq \varnothing$. Set $s \leftarrow \rho_{c',t',r}(\mu)$, update $t \leftarrow t'$, increment the step counter $k \leftarrow k + 1$, and return to (3).

   (iii) If type t is not active at school c' and $|\mu(c')| > |c'|$, then the same-type rejection rule cannot be used. Let $\Phi_{k+1} = (c', t')$ select an active type t' at c' with $\rho_{c',t',r}(\mu) \neq \varnothing$. Set $s \leftarrow \rho_{c',t',r}(\mu)$, update $t \leftarrow t'$, increment the step counter $k \leftarrow k + 1$, and return to (3).

[10] This case is only possible if type t just became active, which likewise requires $a_{c',t} = r_{c',t}$.

(iv) Otherwise, school c' is not over capacity and no rejection is required. The stage terminates. Set $\mu_{i+1} \leftarrow \mu$ and add this matching to the menu. Increment the step counter $k \leftarrow k+1$, increment the stage $i \leftarrow i+1$, and return to (1).

A sequence $\Phi : \mathbb{N} \to C \times T$ is *valid* for $\text{DAW}(M, a, r, \Phi)$ if, whenever the procedure consults Φ, the school-type pair selected by Φ is active and currently in strict surplus.

Definition 19 (Valid sequence for $\text{DAW}(M, a, r, \Phi)$)**.** *Let μ denote the current tentative assignment at the moment when Φ is consulted. The sequence Φ is valid if*

- *In (2), if $\Phi_k = (c, t)$, then type t is active at school c and $\rho_{c,t,r}(\mu) \neq \varnothing$; equivalently, $|\mu(c) \cap S_t| > r_{c,t}$.*
- *After a student applies to school c', if the same-type rejection rule does not select a student and c' remains over capacity, then Φ_{k+1} must select a pair $\Phi_{k+1} = (c', t')$ such that type t' is active at c' and $\rho_{c',t',r}(\mu) \neq \varnothing$; equivalently, $|\mu(c') \cap S_{t'}| > r_{c',t'}$.*

As before, we assume the existence of an additional school which each student ranks last, and which has capacity and reserve targets such that the inequalities in Assumption 2 and 3 hold at equality.

A.4 Proof of Theorem 2

We observe that reserve consistency and limited allocation-constrained efficiency imply allocation-constrained efficiency as in Definition 9.

Lemma 11 (Limited allocation-constrained efficiency and reserve consistency imply allocation–constrained efficiency)**.** *If a matching μ is limited allocation-constrained efficient and reserve-consistent, then μ is allocation-constrained efficient.*

Proof. Suppose μ is reserve-consistent but not allocation-constrained efficient, so there exists a within-type fair matching μ' that Pareto dominates and diversity-dominates μ. We will show that $\text{alloc}(\mu') = \text{alloc}(\mu)$, thus violating limited allocation-constrained efficiency. Fix a school c and type t, and write $n_{c,t} = |\mu(c) \cap S_t|$ and $n'_{c,t} = |\mu'(c) \cap S_t|$.

If $n_{c,t} > r_{c,t}$, then $n'_{c,t} \leq n_{c,t}$, and if $n_{c,t} = r_{c,t}$, then $n'_{c,t} = r_{c,t} = n_{c,t}$. Meanwhile, if $n_{c,t} < r_{c,t}$, it cannot be that $n'_{c,t} < n_{c,t}$. But if $n'_{c,t} > n_{c,t}$, then there is some type-t student s such that $s \in \mu'(c)$ but $s \notin \mu(c)$. Since μ' Pareto dominates μ, this implies $c \succ_s \mu(s)$. But $|\mu(c) \cap S_t| < r_{c,t}$, so this contradicts the reserve consistency of μ. Therefore it must be that $n'_{c,t} = n_{c,t}$ in all cases, and so $\text{alloc}(\mu') = \text{alloc}(\mu)$.

But μ' is within-type fair, Pareto dominates μ, and has the same allocation as μ. Therefore μ cannot be limited allocation-constrained efficient. Hence no such μ' exists and so μ is allocation-constrained efficient. □

Toward a proof of the main result, we present an analogue to Lemma 2. Observe that for any type t and school c the number of students of type t matched to c weakly increases up until

the moment that type t gets activated. From that moment onward until the termination of the procedure, the number of students can only decrease up until reaching the exact reserve $r_{c,t}$, as no chain of rejections of type t students can end with an application at school c under the definition of $\text{DAW}(M, a, r, \Phi)$. Furthermore, once the number of type t students becomes equal to $r_{c,t}$, no chain of rejections of type t students can start at school c by the definition of Φ being a valid sequence for a. We elide proof.

Lemma 12 (Quasi-concave sequence in $\text{DAW}(M, a, r, \Phi)$)**.** *If for a fixed valid pair of a sequence Φ, and activation vector a, the menu outputted by $\text{DAW}_k(M, a, r, \Phi)$ is $\{\mu_0, \mu_1, \cdots \mu_n\}$, then for all $c \in C$ and for all types $t \in T$ we have that $|\mu_0(c) \cap S_t|, |\mu_1(c) \cap S_t|, \cdots |\mu_k(c) \cap S_t|, r_{c,t}$ is a quasi-concave sequence.*

First, we use the quasi-concavity of $\text{DAW}(M, a, r, \Phi)$ to show that every matching which lies on its menu must be reserve-consistent.

Lemma 13 (Reserve consistency of $\text{DAW}(M, a, r, \Phi)$)**.** *Every matching μ_i on a menu generated by $\text{DAW}(M, a, r, \Phi)$ for a valid sequence Φ and activation vector a is reserve-consistent.*

Proof. Fix a feasible menu matching μ_i and suppose student s with type $t = \tau(s)$ strictly prefers school c to her assignment under μ_i, so $c \succ_s \mu_i(s)$. Since students move down their preference lists only after being rejected, student s must have applied to c and been rejected by c before μ_i was recorded in the menu.

Let μ denote the tentative assignment immediately before this rejection. Since s is a type-t student rejected from c, the rejection rule for type t must be active at c, and $\rho_{c,t,r}(\mu) \neq \varnothing$. Hence $|\mu(c) \cap S_t| > r_{c,t}$. Immediately after rejecting s, school c still weakly meets its type-t reserve. But by Lemma 12, once type t is active at c, the number of type-t students assigned to c cannot fall below $r_{c,t}$. Therefore $|\mu_i(c) \cap S_t| \geq r_{c,t}$. □

We now show that, if a matching is within-type fair and limited allocation-constrained efficient, that it is equivalent to the hard-reserves matching for *some* reserve vector.

Lemma 14 (Within-type fair characterization)**.** *A matching μ is within-type fair if and only if μ is stable with respect to the hard-reserves school choice rules $Ch^{\text{HR}}_{c,\text{alloc}(\mu)}$.*

Proof. Let us denote by $v = \text{alloc}(\mu)$. Suppose μ is not within-type fair and let $c \succ_s \mu(s)$ and $p_c(s) > p_c(s')$ for some student $s' \in \mu(c)$ where $\tau(s) = \tau(s')$. Then $Ch^{\text{HR}}_{c,v}(\mu(c) \cup \{s\}) \neq \mu(c)$ and thus μ is not stable with respect to $Ch^{\text{HR}}_{c,\text{alloc}(\mu)}$.

Suppose that μ is not stable with respect to $Ch^{\text{HR}}_{c,v}$. This means that there must exist some student s and school c such that $c \succ_s \mu(s)$ and such that $Ch^{\text{HR}}_{c,v}(\mu(c) \cup \{s\}) \neq \mu(c)$. The hard-reserves choice rule is separable across types; this separability gives us that $Ch^{\text{HR}}_{c,v}(\mu(c) \cup \{s\}) \neq \mu(c) \iff Ch^{\text{HR}}_{c,v}((\mu(c) \cap S_{\tau(s)}) \cup \{s\}) \neq (\mu(c) \cap S_{\tau(s)})$. Because $v = \text{alloc}(\mu)$, we get that $|(\mu(c) \cap S_{\tau(s)}) \cup \{s\}| = v_{c,\tau(s)} + 1$ and so there must exist a student $s' \in \mu(c)$ with priority lower than s but with the same type, hence μ cannot be within-type fair. □

Lemma 15 (Within-type fair and efficiency characterization)**.** *A matching μ is within-type fair and limited allocation-constrained efficient if and only if μ is the result of running student-proposing DA with the set of hard-reserves school choice rules $Ch^{\text{HR}}_{c,\text{alloc}(\mu)}$.*

Proof. Again let $v = \text{alloc}(\mu)$. Suppose first that μ is within-type fair and limited allocation-constrained efficient. By Lemma 14 μ is stable with respect to the hard-reserves choice rules $Ch^{\text{HR}}_{c,v}$. Since hard-reserves choice rules are substitutable, student-proposing DA produces the student-optimal stable matching, so μ^{HR} weakly Pareto dominates μ. If μ^{HR} strictly Pareto dominated μ, this would contradict limited allocation-constrained efficiency, hence $\mu = \mu^{\text{HR}}$.

Conversely, suppose $\mu = \text{DA}(\text{HR}(\text{alloc}(\mu)))$. Stability with respect to $Ch^{\text{HR}}_{c,v}$ implies within-type fairness by Lemma 14. If there were a within-type fair matching μ' with $\text{alloc}(\mu') = v$ that Pareto dominated μ, then μ' would also be stable with respect to $Ch^{\text{HR}}_{c,v}$ by Lemma 14, contradicting student-optimality of $\mu = \text{DA}(\text{HR}(v))$. Thus μ is limited allocation-constrained efficient. □

Then using Lemmas 11 and 13, we can extend the characterization to include reserve consistency and allocation-constrained efficiency.

We now show that any matching which lies on a menu produced by $\text{DAW}(M, a, r, \Phi)$ for some activation vector a is stable with respect to the hard-reserves choice rules with reserves equal to its allocation. The following proposition extends the results from Proposition 4; because of their similarity, we present a sketch of the proof.

Proposition 5 (Internal consistency and termination of $\text{DAW}(M, a, r, \Phi)$)**.** *For a valid sequence Φ and activation vector a, let the menu outputted by $\text{DAW}(M, a, r, \Phi)$ be $\{\mu_0, \mu_1, \cdots \mu_n\}$. Then $\mu_k = \text{DA}(\text{HR}(\text{alloc}(\mu_k)))$ for $k < n$ and $\mu_n \equiv \text{DA}(\text{HR}(r))$.*

Proof. We provide a sketch. Fix a matching μ_i on the menu, and let ℓ_i be the step at which μ_i is recorded; set $v^i = \text{alloc}(\mu_i)$. We will show that every rejection made before μ_i is also a rejection which would be made when running hard-reserves with reserve vector v^i.

Consider a rejection of a type-t student from school c before ℓ_i, and let μ^ℓ be the tentative assignment immediately before this rejection. So type t is active at c, and the rejected student is $b_{c,t}(\mu^\ell)$. Immediately after this rejection, the number of type-t students at c falls to $|\mu^\ell(c) \cap S_t| - 1$; by the quasi-concavity established in Lemma 12, by the time the procedure reaches the menu matching μ_i, the number of type-t students at c cannot exceed this. Hence $v^i_{c,t} \leq |\mu^\ell(c) \cap S_t| - 1$, and so $|\mu^\ell(c) \cap S_t| > v^i_{c,t}$. Therefore the same student will also always be rejected via the hard-reserves choice rule with reserve vector v^i.

The sequence Φ merely determines which legal rejection to make next: in (4i), the procedure rejects the bottom student of the arriving type; in (4ii) and (4iii), the same-type rejection is unavailable, so the valid sequence Φ selects another active type whose bottom student is rejected. As we showed above, whichever type is actually rejected is also in strict surplus relative to v^i.

When μ_i is recorded, no type at a school is in strict surplus relative to v^i by definition of $v^i = \text{alloc}(\mu_i)$, hence $\mu_i = \text{DA}\big(\text{HR}(v^i)\big)$. At the last menu element, the procedure terminates only

when the allocation is diversity-optimal relative to r. Hence $v^n = \text{alloc}(\mu_n) = r$, and therefore $\mu_n = \text{DA}(\text{HR}(r))$. □

We now complete the proof of Theorem 2. We demonstrate that the execution of $\text{DAW}(M, \text{alloc}(\mu), \Phi)$ is identical to the execution of $\text{DAW}(M, a, r, \Phi)$ up to μ for a matching μ which satisfies the relevant axioms, where we denote by $a = \max(\text{alloc}(\mu), r)$ the pointwise maximization between the vectors r and $\text{alloc}(\mu)$.

Proof. Let $v = \text{alloc}(\mu)$ and choose a valid sequence Φ such that the execution of $\text{DAW}(M, v, \Phi)$ reaches μ. First observe that if $v_{c,t} < r_{c,t}$, then no type-t student can be rejected from school c before μ is reached in the execution of $\text{DAW}(M, v, \Phi)$. If such a student s were rejected from c, then s would subsequently move only to schools that she ranks below c, but this contradicts the reserve consistency of μ.

We prove the prefix equivalence inductively. The initial assignments under $\text{DAW}(M, v, \Phi)$ and $\text{DAW}(M, a, r, \Phi)$ are identical, since both assign each student to her most-preferred school. Suppose that the two executions are identical up to some step before μ is reached, and denote the current assignment the two procedures have in common by μ_i.

First, suppose that $\text{DAW}(M, v, \Phi)$ initiates a rejection chain by selecting a type-t student from school c. Then $|\mu_i(c) \cap S_t| > v_{c,t}$. By reserve consistency, we cannot have $v_{c,t} < r_{c,t}$, so $v_{c,t} \geq r_{c,t}$ and hence $a_{c,t} = v_{c,t}$. Therefore $|\mu_i(c) \cap S_t| > a_{c,t}$ as well. Thus type t is active at school c in the execution of $\text{DAW}(M, a, r, \Phi)$, and the same bottom-priority type-t student is rejected.

Second, suppose that, during a rejection chain, a type-t student s' applies to school c in the execution of $\text{DAW}(M, v, \Phi)$, and this causes c to reject a type-t student s. Then $|(\mu_i(c) \cap S_t) \cup s'| > v_{c,t}$. Again, by the observation above, this rejection cannot occur at a coordinate with $v_{c,t} < r_{c,t}$. Hence $v_{c,t} \geq r_{c,t}$ and $a_{c,t} = v_{c,t}$. Therefore $|(\mu_i(c) \cap S_t) \cup s'| > a_{c,t}$, so the same type-$t$ rejection is triggered in the execution of $\text{DAW}(M, a, r, \Phi)$.

Third, suppose that a rejection chain in $\text{DAW}(M, v, \Phi)$ ends when a type-t student s applies to school c. By construction, this means $|(\mu_i(c) \cap S_t) \cup s| \leq v_{c,t}$. Since $v_{c,t} \leq a_{c,t}$, we also have $|(\mu_i(c) \cap S_t) \cup s| \leq a_{c,t}$. Thus in the execution of $\text{DAW}(M, a, r, \Phi)$, the application does not put type t strictly above its activation threshold at school c. Hence no same-type rejection is triggered, and the same rejection chain ends at the same school.

This completes the argument that if matching μ is within-type fair, limited allocation-constrained efficient and reserve-consistent (and thus allocation-constrained efficient by Lemma 11), then μ is in the menu produced by $\text{DAW}(M, \max(\text{alloc}(\mu), r), r, \Phi)$ for some sequence Φ.

In the other direction, suppose that μ is generated by $\text{DAW}(M, a, r, \Phi)$ for some valid sequence Φ and activation vector a. By Proposition 5, $\mu = \text{DA}(\text{HR}(\text{alloc}(\mu)))$. By Lemma 15, μ is within-type fair and limited allocation-constrained efficient, and by Lemma 13, μ is reserve-consistent, so it is also allocation-constrained efficient by Lemma 11. □

B Deferred Acceptance with Single Waiting Rooms

We now give a formal definition of our DAW(S) algorithm. The definition formalizes the description of the procedure in Section 4.2, and uses notation from Appendix 5. Now, Φ is a sequence which instead identifies schools only: $\Phi : \mathbb{N} \to C$. This is sufficient for the DAW(S) algorithm because at each step the rejected student must be the lowest-priority surplus student at a particular school, independent of the type of the last student who applied. We also adapt the definition of the rejection rule accordingly. For any assignment μ, let

$$\rho_{c,r}(\mu) = \begin{cases} \arg \min_{s \in \{b_{c,t}(\mu):\, |\mu(c) \cap S_t| > r_{c,t}\}} p_c(s) & \text{if } \exists t \in T \text{ s.t. } |\mu(c) \cap S_t| > r_{c,t}, \\ \varnothing & \text{otherwise.} \end{cases}$$

where $b_{c,t}(\mu)$ is the lowest-priority type-t student assigned to c under μ; as before, we sometimes suppress r when clear from the context. We write $\text{DAW}(S, r, \Phi)$ for the mechanism.

Also as in Appendix 5, we assume the existence of a dummy school whose capacity and reserves set the inequalities in Assumption 2 and 3 to hold at equality.

Definition 20 (Deferred acceptance with single waiting rooms ($\text{DAW}(S, r, \Phi)$))**.** $\text{DAW}(S, r, \Phi)$ *proceeds in steps k and stages i. In the initial stage, $\mu_0(s)$ is set to the top-choice school for each student s, and the step is set to $k \leftarrow 1$. Each stage i starts with an interim matching μ_i and step counter k and proceeds as follows.*

1. *If μ_i is diversity-optimal (i.e. $\Delta_{c,t} = 0$ for all t, c), the process ends and outputs μ_i. Otherwise, set $\mu \leftarrow \mu_i$.*

2. *Set $c \leftarrow \Phi_k$, i.e., consider the school c given by Φ_k, where k is the current step. Set $s \leftarrow \rho_c(\mu)$, i.e., select the next student s to reject as the surplus student with lowest priority.*

3. *Set $\mu(s) \leftarrow c'$, where c' is the next highest school in the preferences of s; i.e., tentatively assign s to her next most-preferred school c'.*

4. *Select the next student to reject as follows:*

   (i) If c' is over its capacity, set $s \leftarrow \rho_{c',r}(\mu)$ to be the lowest priority strict surplus student at c'. Increment the step counter $k \leftarrow k + 1$, and return to (3).

   (ii) If c' is not over its capacity, but μ violates across-type fairness, set $s \leftarrow \rho_{c',r}(\mu)$ to be the lowest priority strict surplus student at c'. Increment the step counter $k \leftarrow k + 1$, and return to (3).

   (iii) Otherwise, the stage terminates. We set $\mu_{i+1} \leftarrow \mu$ and if μ is feasible we add this matching to our menu. Finally, we increment the step $k \leftarrow k + 1$, stage $i \leftarrow i + 1$, and return to (1).

We again need the sequence Φ to be *valid* as in the previous sections; we elide the definition as it is analogous.

Before proving Theorem 3, we establish in Proposition 6 that a feasible matching μ is within-type fair, across-type fair with respect to reserves, reserve-consistent, and non-wasteful if and only if μ is stable with respect to the soft-reserves choice rule. We then show that the matchings which are consequence to running $\text{DAW}(S, r, \Phi)$ are stable with respect to the soft-reserves choice rules, hence the theorem holds.

Proposition 6 (Axiomatic characterization of stability with respect to soft-reserves)**.** *A feasible matching μ is within-type fair, across-type fair with respect to reserves, reserve-consistent, and non-wasteful if and only if μ is stable with respect to the soft-reserves choice rules induced by the target allocation vector r.*

Proof. In Proposition 2 we established that any matching μ that is stable with respect to the soft-reserves choice rules (induced by r), satisfies the within-type fairness, across-type fairness and reserve consistency axioms (with respect to r).

A contrapositive argument completes this proposition. If μ fails non-wastefulness, there must be a student s that is unmatched at μ, and a school c for which $|\mu(c)| < |c|$; then s would block μ. It is straightforward to check that if any of the within-type fairness, across-type fairness, or reserve-consistency axioms is violated for a student s and school c, then s would block μ because $s \in Ch^{\text{SR}}_{c,r}(\mu(c) \cup \{s\})$. □

Toward the proof of the main result, we establish that the $\text{DAW}(S, r, \Phi)$ procedure preserves the relevant axioms at every interim matching: reserve consistency, within-type fairness, and across-type fairness with respect to reserves. These properties paired with non-wastefulness imply that every feasible interim matching is stable with respect to the soft-reserves choice rules. We then identify the endpoints of the path, showing that the first feasible matching is the soft-reserves matching and that the terminal matching is the hard-reserves matching.

Lemma 16 (Reserve consistency of $\text{DAW}(S, r, \Phi)$)**.** *At every step of the* $\text{DAW}(S, r, \Phi)$ *procedure, the current assignment is reserve-consistent.*

Proof. Consider any step of the procedure and let μ be the current assignment. Suppose that there is a student s of type t and a school c such that $c \succ_s \mu(s)$. Then s must have applied to c and been rejected by c at an earlier step. By the definition of $\rho_{c,r}$, a type-t student can be rejected from c only when $|\mu(c) \cap S_t| > r_{c,t}$ immediately before the rejection. Hence immediately after the rejection at least $r_{c,t}$ type-t students remain assigned to c.

After this step, the number of type-t students assigned to c can decrease only through another rejection of a type-t student from c, and each such rejection again occurs only when there are strictly more than $r_{c,t}$ type-t students assigned to c. Thus once s has been rejected from c, every later assignment has at least $r_{c,t}$ type-t students assigned to c, hence reserve consistency holds. □

Lemma 17 (Within-type fairness of $\mathrm{DAW}(S, r, \Phi)$)**.** *Every interim matching produced by* $\mathrm{DAW}(S, r, \Phi)$ *is within-type fair.*

Proof. Let μ_i be an interim matching produced by $\mathrm{DAW}(S, r, \Phi)$. Suppose toward a contradiction that μ_i violates within-type fairness. Then there exist students s and s' of the same type t and a school c such that $s' \in \mu_i(c)$, $c \succ_s \mu_i(s)$, and $p_c(s) > p_c(s')$.

Since $c \succ_s \mu_i(s)$, student s must have applied to c and been rejected by c at an earlier step. At the moment s was rejected, s was the lowest-priority strict surplus student at c. But for the rest of the procedure, any type-t student with priority lower than s who applies to c must become the lowest-priority type-t student at c. And because type t has already met its reserve at c by Lemma 16, such a student must be in strict surplus. Therefore no type-t student with priority below s can remain assigned to c at μ_i, contradicting. □

Lemma 18 (Across-type fairness of $\mathrm{DAW}(S, r, \Phi)$)**.** *Every interim matching produced by* $\mathrm{DAW}(S, r, \Phi)$ *is across-type fair with respect to reserves.*

Proof. Let μ_i be an interim matching produced by $\mathrm{DAW}(S, r, \Phi)$. Suppose, toward a contradiction, that μ_i violates across-type fairness with respect to reserves. Then there exist students s and s' and a school c such that $s' \in \mu_i(c)$, $c \succ_s \mu_i(s)$, $p_c(s) > p_c(s')$, and $|\mu_i(c) \cap S_{\tau(s')}| > r_{c,\tau(s')}$.

Since $c \succ_s \mu_i(s)$, student s must have applied to c and been rejected by c at an earlier step. At the moment s was rejected, $\rho_{c,r}$ selected the lowest-priority student among the bottom students of all strict surplus types at c. After this rejection, no interim matching can leave at c a strict surplus student with priority below s. Indeed, if such a student appears during a stage, then the assignment violates across-type fairness at c, and the $\mathrm{DAW}(S, r, \Phi)$ procedure continues by rejecting the lowest-priority strict surplus student at c before recording an interim matching. Since s' is in strict surplus at c under μ_i and $p_c(s) > p_c(s')$, μ_i must not be an interim matching. □

By Lemmas 16, 17, and 18, matchings generated by the procedure are reserve-consistent, within-type fair, and across-type fair with respect to reserves; they are also trivially non-wasteful. Thus by Proposition 6, every feasible matching in the menu produced by $\mathrm{DAW}(S, r, \Phi)$ is stable with respect to the soft-reserves choice rules induced by r.

We now show that the endpoints of any path constructed by $\mathrm{DAW}(S, r, \Phi)$ correspond to the soft-reserves and hard-reserves matchings, which completes the proof.

Lemma 19 (Soft-reserves endpoint of $\mathrm{DAW}(S, r, \Phi)$)**.** *Let* μ_1 *be the first feasible matching in the menu produced by* $\mathrm{DAW}(S, r, \Phi)$*. Then* $\mu_1 = \mathrm{DA}(\mathrm{SR}(r))$*.*

Proof. Let $\mu^{\mathrm{SR}} := \mathrm{DA}(\mathrm{SR}(r))$. μ_1 is stable with respect to the soft-reserves choice rules induced by r and is non-wasteful. Since μ^{SR} is the student-optimal stable matching for these choice rules, μ^{SR} Pareto dominates μ_1.

Suppose that $\mu_1 \neq \mu^{\mathrm{SR}}$ and consider the first step in which a student s is rejected by the school c they are matched to under μ^{SR}; this step must come before we reach μ_1. Let us denote by μ

the matching that exists just after student s has been rejected by their school they are matched to under μ^{SR}. This implies that $|\mu(c) \cap S_t| \geq r_{c,t}$ which means that s has been the lowest priority surplus student at c right before s was rejected. We claim that there must be at least one student $s' \in \mu(c)$ who would block μ^{SR}.

First, note that every student $s' \in \mu(c)$ weakly prefers c to her matched school under μ^{SR} because s is the first student that is rejected by her match under μ^{SR}. Then note that because s has been the lowest priority surplus student at c right before s was rejected, we must have that if $\tau(s')$ is not a deficit type at c under μ, then $p_c(s') > p_c(s)$. Thus given that μ^{SR} must be stable, we have that together $s' \in \mu(c)$ and $|\mu(c) \cap S_{\tau(s')}| \geq r_{c,\tau(s')}$ imply $\mu^{\text{SR}}(s') = c$. On the other side, together $s' \in \mu(c)$ and $|\mu(c) \cap S_{\tau(s')}| < r_{c,\tau(s')}$ imply that there are at least as many students from type $\tau(s')$ matched to c under μ^{SR} as there are under μ, i.e. $|\mu(c) \cap S_{\tau(s')}| \leq |\mu^{\text{SR}}(c) \cap S_{\tau(s')}|$ because each of these students would prefer c and be eligible to enter if c is still at deficit for that type under μ^{SR}. Together these two arguments show a contradiction and thus μ_1 Pareto dominates μ^{SR}, which means that they are identical matchings. □

Lemma 20 (Hard-reserves endpoint of DAW(S, r, Φ))**.** *Let μ_n be the terminal matching in the menu produced by* DAW(S, r, Φ)*. Then* $\mu_n = \text{DA}(\text{HR}(r))$.

Proof. We give a sketch proof as the logic is similar to Lemma 19. The number of surplus students weakly decreases throughout DAW(S, r, Φ), and no student can be rejected from all schools because this would imply that all have weak surplus of student of that type. Thus the algorithm terminates when the allocation of students matches the target allocation r. At that moment we must have a matching μ that is stable with respect to the soft-reserves choice rules induced by r and has an allocation $\text{alloc}(\mu) = r$, which implies that μ is also stable with respect to the hard-reserves choice rules induced by r. The student optimality of DA(HR) under this set of school choice rules imply that DA(HR) Pareto dominates μ_n.

An analogous argument to the one for DA(SR) above demonstrates that at no step of the execution of the algorithm a student s is ever rejected by the school c she is matched to under DA(HR). Thus we must have that μ_n Pareto dominates DA(HR), and so they must be identical matchings. □

B.1 Deferred Acceptance with Single and Flexible Waiting Rooms

Deferred acceptance with single and flexible waiting rooms generalizes the DAW(S) procedure by considering a larger set of students as the starting point for a cycle in each stage, including students that are in weak surplus at a school. Algorithmically, if the rejection chain initiated by a student does not result in a cycle that satisfies our axioms, the procedure backtracks to the last intermediate matching. This approach allows DAW(SF) to yield menus of feasible matchings that constitute complete paths on the lattice of matchings that are stable with respect to the soft-reserves choice rules induced by the target allocation.

We identify this algorithm formally by DAW(SF, r, Φ), and the rejection rule is defined the same as in DAW(S, r, Φ). Crucially, this procedure derives its flexibility from how the first student is chosen: she may be rejected when she is the least-priority student that brings her type exactly to capacity, and when a different student of the same type but with higher priority may arrive back at that school, thus the matching remains reserve-consistent. We thus need the sequence to choose the bottom student of a given type at a school, $\Phi : \mathbb{N} \to C \times T$. If the cycle it initiates does not lead to the desired axioms being satisfied, we revert to the previous assignment. We assume the sequence is valid in the sense of the previous procedures.

Definition 21 (Deferred acceptance with single and flexible waiting rooms DAW(SF, r, Φ))**.** DAW(SF, r, Φ) *proceeds in steps k and stages i. In the initial stage, $\mu_0(s)$ is set to the top-choice school for each student s, and the step is set to $k \leftarrow 1$. Each stage i starts with an interim matching μ_i and step counter k and proceeds as follows.*

1. *Set $\mu \leftarrow \mu_i$. If no remaining entry of the valid sequence Φ can initiate a successful stage from μ_i, the process ends and outputs μ_i.*

2. *Let $k' \geq k$ be the first index such that, for $\Phi_{k'} = (c, t)$, we have $|\mu(c) \cap S_t| \geq r_{c,t}$. Set $k \leftarrow k'$ and $s \leftarrow b_{c,t}(\mu)$, i.e., reject the bottom type-t student at c.*

3. *If s has been rejected by all schools, set $\mu \leftarrow \mu_i$, increment the step counter $k \leftarrow k + 1$, and return to (1). Otherwise, set $\mu(s) \leftarrow c'$, where c' is the next highest school in the preferences of s; i.e., tentatively assign s to her next most-preferred school c'.*

4. *Select the next student to reject as follows:*

   (i) If c' is over its capacity, set $s \leftarrow \rho_{c',r}(\mu)$. Increment the step counter $k \leftarrow k + 1$, and return to (3).

   (ii) If c' is not over its capacity, but μ violates across-type fairness at c', set $s \leftarrow \rho_{c',r}(\mu)$. Increment the step counter $k \leftarrow k + 1$, and return to (3).

   (iii) Otherwise, the stage terminates. If μ satisfies reserve consistency, set $\mu_{i+1} \leftarrow \mu$ and if μ is feasible, add this matching to the menu. Finally, increment the step counter $k \leftarrow k + 1$, set $i \leftarrow i + 1$, and return to (1). If μ does not satisfy these axioms, set $\mu \leftarrow \mu_i$, increment the step counter $k \leftarrow k + 1$, and return to (1).

Note (4iii) is used to verify that rejecting the weak-surplus student cannot violate the axioms. To see why Assumptions 2 and 3 are necessary, consider that, if we were to assume a dummy school to contain all the residual seats in the economy, the rejection condition in (3) might render a student unmatched. This violates non-wastefulness.

B.2 Proof of Theorem 4

Recall in Proposition 6 we characterized the set of feasible and non-wasteful matchings that are within-type fair, across-type fair with respect to reserves, and reserve-consistent. By construction,

DAW(SF, r, Φ) records a feasible matching only after verifying these axioms in (4iii). Hence every feasible matching recorded by DAW(SF, r, Φ) belongs to the soft-reserves stable lattice.

We now show that every path in this lattice can be generated by a suitable valid sequence Φ. The proof has two steps. First, for any two consecutive matchings on a lattice path, we identify a student who is eligible to initiate the corresponding DAW(SF, r, Φ) stage. Second, we show that, when this student is selected, the stage produces the next matching on the path.

Lemma 21 (Existence of an eligible initiating student in DAW(SF, r, Φ))**.** *Let μ_i and μ_{i+1} be consecutive matchings in the soft-reserves stable lattice, with μ_i Pareto dominating μ_{i+1}. Let $P = \{s \in S : \mu_i(s) \neq \mu_{i+1}(s)\}$. Then P contains a student who can begin a stage in the* DAW(SF, r, Φ) *procedure.*

Proof. For each $s \in P$, let $c = \mu_i(s)$ and $c' = \mu_{i+1}(s)$. As μ_i Pareto dominates μ_{i+1}, we have $c \succ_s c'$. Observe that within-type fairness of μ_{i+1} implies that if there is a student $s' \in \mu_i(c)$ such that $\tau(s') = \tau(s)$ and $p_c(s) > p_c(s')$, then $s' \in P$. Thus the students in P must be the least-preferred students of their respective types at their schools in μ_i. Additionally, there must be at least one student in P who leaves a school at which she is weak-surplus; if not, then the net change would not balance across μ_i and μ_{i+1}.

Then there exists such a student from the bottom segment of her type at her school. Writing $c = \mu_i(s_0)$ and $t = \tau(s_0)$, we have $s_0 = b_{c,t}(\mu_i)$ and $|\mu_i(c) \cap S_t| \geq r_{c,t}$. Hence s_0 is eligible to initiate a stage. □

Lemma 22 (A flexible stage follows a lattice edge in DAW(SF, r, Φ))**.** *Let μ_i and μ_{i+1} be consecutive matchings in the soft-reserves stable lattice, with μ_i Pareto dominating μ_{i+1}. Let s_0 be an eligible initiating student from Lemma 21. If a* DAW(SF) *stage begins by rejecting s_0 from $\mu_i(s_0)$, then the next feasible matching recorded by that stage is μ_{i+1}.*

Proof. Let μ^* be the matching recorded by the stage initiated by s_0 and recall μ_i and μ_{i+1} are consecutive. Note μ_i Pareto dominates μ^*. We claim that μ^* Pareto dominates μ_{i+1} and thus the matchings are identical. Suppose not; then, during the stage where the eligible student is rejected, there is a first moment at which some student s is rejected from the school she receives under μ_{i+1}. Let $c = \mu_{i+1}(s)$, and consider the assignment immediately before this rejection. By the choice of the first such moment, every student assigned to c at that moment weakly prefers c to her assignment under μ_{i+1}.

The rejection from c is made by the single-waiting-room rule: the rejected student is the lowest-priority student among the bottom students of all strict-surplus types at c. Hence, if s is rejected from c, some student who remains assigned to c would block μ_{i+1} under the soft-reserves choice rule. This contradicts stability of μ_{i+1}. Therefore no student is rejected from her assignment under μ_{i+1} during the stage, so μ^* Pareto dominates μ_{i+1}.

By (4iii) of DAW(SF, r, Φ), every feasible matching recorded by the procedure satisfies the axioms and therefore belongs to the soft-reserves stable lattice. Thus μ_i Pareto dominates μ^* and

μ^* Pareto dominates μ_{i+1}, with all three matchings in the lattice. Since μ_i and μ_{i+1} are consecutive, there is no lattice point strictly between them. Hence $\mu^* = \mu_{i+1}$. □

Proof of Theorem 4. Choose the initial entries of Φ so that $\mathrm{DAW}(SF, r, \Phi)$ follows the same rejections as $\mathrm{DAW}(S, r, \Phi)$ until the first feasible matching is recorded, and fix a path $\mu_1, \ldots, \mu_n$. By Lemma 19, the first matching is $\mu_1 = \mathrm{DA}(\mathrm{SR}(r))$. Let μ_i and μ_{i+1} be consecutive in the lattice; by Lemma 21, there exists an eligible student who can initiate a stage from μ_i. Let Φ select this student next; by Lemma 22, the next feasible matching recorded by the procedure is μ_{i+1}, which completes the induction and proves that every path in the soft-reserves stable lattice can be generated by a valid sequence Φ.

Conversely, every feasible matching recorded by $\mathrm{DAW}(SF, r, \Phi)$ satisfies reserve consistency, within-type fairness, and across-type fairness with respect to reserves, and is non-wasteful by feasibility. Hence, by Proposition 6, every feasible recorded matching lies in the soft-reserves stable lattice. Therefore the feasible menus generated by DAW(SF) are paths in the lattice, and every path in the lattice can be generated by some valid sequence. □