Essays in Market Design

by

Oğuzhan Çelebi

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

Authored by Oğuzhan Çelebi

Department of Economics

April 10, 2023

Certified by Daron Acemoglu

Institute Professor, Thesis Supervisor

Certified by Parag Pathak

Class of 1922 Professor of Economics, Thesis Supervisor

Accepted by Abhijit Banerjee

Ford Foundation International Professor of Economics

Essays in Market Design

by

Oğuzhan Çelebi

Submitted to the Department of Economics
on April 10, 2023, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

This thesis analyzes existing allocation mechanisms and proposes new mechanisms for two-sided matching markets, with a particular focus on the role played by diversity preferences and affirmative action.

In the first chapter, "Diversity Preferences, Affirmative Action and Choice Rules", I introduce a framework to analyze diversity preferences and their effect on the affirmative action policies and choice rules adopted by institutions. I characterize the choice rules that can be rationalized by diversity preferences and demonstrate that the rule used to allocate government positions in India cannot be rationalized. I show that if institutions evaluate diversity using marginal (*i.e.,* not cross-sectional) distribution of identities, then choices induced by their preferences cannot satisfy the substitutes condition, which is crucial for the existence of competitive equilibria and stable allocations. I characterize a class of choice rules that satisfy the substitutes condition and are rationalizable by preferences that evaluate diversity and quality separately and identify the preferences that induce some widely used choice rules. The framework and results presented in this chapter provide a systematic way of evaluating the diversity preferences behind the choices made by institutions.

In the second chapter, "Adaptive Priority Mechanisms" (coauthored with Joel Flynn), we ask how authorities that care about match quality and diversity should allocate resources when they are uncertain of the market they face? Such a question appears in many contexts, including the allocation of school seats to students from various socioeconomic groups with differing exam scores. We propose a new class of *adaptive priority mechanisms* (APM) that prioritize agents as a function of both scores that reflect match quality and the number of assigned agents with the same socioeconomic characteristics. When there is a single authority and preferences over scores and diversity are separable, we derive an APM that is optimal, generates a unique outcome, and can be specified solely in terms of the preferences of the authority. By contrast, the ubiquitous priority and quota mechanisms are optimal if and only if the authority is risk-neutral or extremely risk-averse over diversity, respectively. When there are many authorities, it is dominant for each of them to use the optimal APM, and each so doing implements the unique stable matching. However, this is generally inefficient for the authorities. A centralized allocation mechanism that first uses an aggregate APM and then implements authority-specific quotas restores efficiency. Using data from Chicago Public Schools, we estimate that the gains from adopting APM are considerable.

In the third chapter, "Best Response Dynamics in Boston Mechanism", I introduce and analyze a dynamic process called Repeated Boston Mechanism (RBM), where the Boston Mechanism (BM) is used for multiple periods, and students form their application strategies by best responding to the admission cutoffs of the previous period. If students are truthful in the initial period, the allocation under RBM converges in finite time to the student optimal stable matching (SOSM), which is the Pareto-dominant equilibrium of BM and the outcome of the strategy-proof Deferred Acceptance Mechanism. If some students are sincere and do not strategize, then the allocation converges to the SOSM of a market in which sincere students lose their priorities to sophisticated ones. When students are not truthful in the first period but best reply to some initial admission cutoffs, the allocation converges to SOSM if students are initially optimistic about their admissions chances but may cycle between allocations Pareto-dominated by SOSM if they are pessimistic. These results provide a foundation for the earlier characterizations of equilibria of BM and are in line with the observations of non-equilibrium play in BM in real-world markets.

JEL Classification Codes: D47, D61

Thesis Supervisor: Daron Acemoglu
Title: Institute Professor

Thesis Supervisor: Glenn Ellison
Title: Gregory K. Palm (1970) Professor of Economics

Thesis Supervisor: Parag Pathak
Title: Class of 1922 Professor of Economics

Acknowledgments

This thesis wouldn't be possible without the support of my advisors, colleagues, friends and family. Parag's guidance was invaluable during my PhD. He made sure I was on track to be a successful researcher. Daron was always very generous with his time and wisdom. From him, I learned to ask the most important questions and not get bogged down to the details. Tayfun's desire to have a positive impact on the society was a source of inspiration and will always be a goal to strive for. Our conversations shaped my approach to research. I am grateful to Glenn Ellison for his geniune interest in my research and the wonderful comments he has provided. Elias Papaioannou was not only an amazing coauthor, but also a constant source of support during my PhD. I am also thankful for the invaluable advice I have received from Charles Angelucci, Bob Gibbons, Stephen Morris, Birger Wernerfelt and Alex Wolitzky.

I am indebted to Kamil Yılmaz for encouraging me to pursue a PhD in Economics, Alp Atakan and Levent Koçkesen for being great role models, and Özgur Yılmaz for introducing me to market design and guiding me while I was working on my first project. Without their support, I wouldn't be able to start a PhD at MIT, let alone complete this thesis.

I am thankful for the countless conversations I had with Arda Gitmez, Ömer Karaduman, Mert Demirer, Ahmet Gülek, Ali Kakhbod and Joseph Shayani. I enjoyed the table tennis games, dinners and research discussions I shared with Roberto Corrao, who made long nights in the office bearable. Ari Bronsoler was always there when I needed cheering up. Doing a PhD at MIT Economics means you get the most capable classmates in the world: Joel Flynn co-wrote Chapter 2 of this thesis, as well as the first paper I have published, and gave invaluable comments for all my other papers. But more importantly, he was an amazing roommate, office mate and friend.

I am grateful for the wonderful friendship I have shared with Berkay Balkancı and Yunus Atçalı for the last 16 years. They played a huge role in shaping how I look at the world since high school. The unconditional love of my mom and dad was with me for all my live and I am proud to call myself their son. Finally, I would like to dedicate this thesis to my person, Elif, who has blessed me with her endless love and support. She is the joy in my life and I cannot imagine going through the last 9 years without her.

Contents

Chapter 1

Diversity Preferences, Affirmative Action and Choice Rules

1.1 Introduction

Institutions in charge of allocating resources or hiring individuals make their decisions based on multiple criteria, such as the quality of the individuals they hire, the benefits individuals receive from the allocated resource or the socio-economic characteristics of individuals who are hired or allocated the resource. For example, school districts in Chicago and Boston as well as universities in Brazil prefer schools to have a diverse student body (Dur, Kominers, Pathak, and Sönmez, 2018; Dur, Pathak, and Sönmez, 2020; Aygun and Bó, 2021; Grigoryan, 2021), medical authorities prefer the allocation of scarcely available treatments to consider values such as equity and diversity while making sure medical workers themselves are able to receive the treatment (Pathak, Sönmez, Ünver, and Yenmez, 2021; Akbarpour, Budish, Dworczak, and Kominers, 2021) and Indian government uses protections for historically discriminated groups when allocating government positions (Aygün and Turhan, 2017; Sönmez and Yenmez, 2022a). In all these settings, individuals are heterogeneous in two domains. First is their identities, a student's socio-economic status, a patient's healthcare worker status or the caste of a government position applicant. Second is their score, such as exam scores in student assignment and government job allocation or index of clinical need in medical resource allocation. These scores can denote a measure of match-quality, where allocating the medical resource to a sicker individual generates more benefit, or represent individual's property rights over the resource, where candidates who performed better in the exam deserve the government jobs more than lower performing candidates.[1] Moreover, these institutions implement affirmative action programs aimed at increasing the representation of individuals from certain underrepresented groups. These programs suggest that these institutions prefer to allocate the resource not only to the individuals with highest scores, but also care about values such as diversity and equity.

The goal of this paper is to present a theory of diversity preferences. This allows us to understand the relationship between (i) how institutions evaluate diversity and consider the trade-offs between diversity and scores, and (ii) how they determine the allocation rules and affirmative action programs. To this end, I develop a model where individuals are heterogeneous in two domains, their (possibly multidimensional) identity, and their score. An institution chooses whom to hire or allocate a resource to. In most of my applications, the institution uses a mechanism that determines how the resource will be allocated, which I refer to as the choice rule. However, one can also consider settings where these choices are based on observations of the choices of the institution on different instances (for example, different sets of applicants), as in standard consumer theory models.

I start the analysis by studying when do the choice rules adopted by the institution can be rationalized by diversity preferences. I establish a connection between the standard consumer theory setting and the problem of choosing sets of individuals with different identities and qualities. This allows me to leverage existing results from consumer theory to gain insights about diversity preferences. First, I analyze a setting where the institution does not have a strict preference for allocating the resource to higher scoring agents.[2] I characterize the choice rules that can be rationalized by a preference relation (or a utility function) over the scores and the identities of the chosen individuals. Invoking Theorem 1 in Richter (1966), I give a necessary and sufficient condition under which the choice rule is rationalizable. This condition is the *congruence axiom* of Richter (1966): if a set of individuals I is indirectly revealed preferred to another set I' under the choice rule, then I' cannot be revealed preferred to I. Next, I assume that the individuals are ordered according to their scores and the institution prefers individuals with higher scores to ones with lower scores, as in many of the applications this paper considers. If the institution has preference for higher scores and does not value diversity or is not concerned with the identities of

[1] When scores represent exam scores, we can interpret the preference for higher scores as a preference for procedural fairness.

[2] Indeed, in this setting, I do not even assume set of scores is an ordered set. For example, the "score" of an individual may encode the different qualities they have or jobs they can perform, rather than the exam score of a student or clinical need of a patient that ranks individuals.

the individuals, then they must choose the individuals with the highest scores. However, various policies ranging from explicit quotas and priority bonuses in centralized allocation systems to hiring policies that mandate representation from diverse groups allow lower scoring individuals to be chosen before higher scoring individuals from other groups. Motivated by this, I characterize the class of choice rules that can be rationalized by a utility function that is increasing in the scores. These are the choice rules that a rational decision maker can adopt under diversity preferences in order to increase the representation of certain groups. This characterization is obtained by using a result from Nishimura, Ok, and Quah (2016), who give a generalization of Richter's congruence axiom that characterizes the choice rules that are rationalized by utility function that is increasing with respect to a given exogenously given preorder. To apply these theorems to test the rationality of a decision-maker, one needs to observe choice data. However, for the allocation mechanisms considered in this paper, the choice function itself (*i.e.*, the set of chosen individuals for all possible set of applications) is fully known. Thus, rationality of existing allocation mechanisms can be evaluated by checking whether they satisfy the conditions given in the characterizations. I apply this test to the main choice rule that has been used in India for between 1995 and 2020 to match individuals to government and show that it does not satisfy the conditions given in both characterizations. This indicates that the rule wouldn't be chosen by a rational decision maker, building on and complementing the analysis of Sönmez and Yenmez (2022a) who show other important shortcomings of this choice rule and illustrating the practicality of the characterization.

Even though the characterizations of rationality and monotonicity allow us to evaluate mechanisms, they place few restrictions on the choice rule and therefore the rationalizing preferences. Next, I study an important property of choice rules, the (gross) substitutes condition. Substitutes is an important theoretical property of choice rules and has proved to be crucial for existence of competitive equilibria and stable matching (Kelso and Crawford, 1982; Roth, 1984; Gul and Stacchetti, 1999; Hatfield and Milgrom, 2005). A preference for diversity may induce complementarities across individuals with different types and the choice rules that incorporate different constraints and are compatible with the substitutes conditions have been studied extensively. However, these studies do not consider the multidimensional and overlapping nature of identities, which has been introduced to market design recently (Kurata, Hamada, Iwasaki, and Yokoo, 2017; Aygun and Bó, 2021; Sönmez and Yenmez, 2022a) and are important in many settings.

In my model, each individual belongs to a group in each dimension. For example, consider a setting with three dimensions constituting gender, race and income, where an individual's identity is given by the groups they belong in each of these dimensions. When identities are multidimensional, institutions can evaluate diversity in multiple ways. For example, a company who wants to increase diversity in gender and race can look at the number of female workers and the number of workers from underrepresented groups to determine the *diversity* of their workforce. Such an approach focuses on the marginal distribution of identities, but not the cross-sectional distribution of identities as it does not take into account the number of woman from underrepresented groups. Institutions and companies often attach a great deal of value to diversity, incorporate it explicitly into their hiring practices and even publish reports evaluating the diversity of their workforce. However, many of these reports (such as yearly diversity reports of Microsoft, Apple and MIT) only include the marginal distribution of their workforce, that is, they report the percentage of workers who belong to underrepresented groups (*e.g.*, black, asian or hispanic) and gender separately, but do not report, for example, the percentage of woman from underrepresented groups.[3] Similarly, many affirmative action programs aimed at widening representation in legislatures have quotas for both woman and minorities, but these have been typically evolved separately and work independently (Hughes, 2018). Indeed, the large literature on *intersectionality* study how different identities combine to create different modes of discrimination and privilege, with a particular focus on the experience of black woman in the United States (see, *e.g.* Crenshaw (2013)). However, until recently, such considerations have been neglected in the market design literature. Motivated by this, I study the relationship between the way multidimensional/overlapping identities are evaluated by an institution and how its allocation/hiring decisions (*i.e.,* choice rules) satisfy the crucial substitutes conditions.

In my model, an institution does not consider intersectionality if they are indifferent between two sets of individuals with the same marginal distributions of identities. I first analyze a setting where individuals have identical scores and show that if the preferences of an institution do not consider intersectionality and satisfy a mild nontriviality assumption called interior optimum, then the choice rule induced by their preferences does not satisfy the substitutes condition of Roth (1984) (Proposition 3).[4]

[3] Apple and Microsoft report the fraction of employees that belong to different races and genders, while MIT does the same for its student body. An exception is Google, where the percentage of each gender is reported separately within each race in the intersectional hiring section of their annual report. See Figures A-1 and A-2 in Appendix A.2 for examples of different diversity reporting practices.

[4] A special case of preferences that do not consider intersectionality is when an institution that does not care about the diversity of the individuals at all and is indifferent between all allocations.

Next, I allow agents to have different scores that denote their quality. When the institution prefers higher quality individuals to lower quality ones (*monotonicity*), the scores can be viewed as (inverse) salaries, and the model is closely related to the classical job matching model of Kelso and Crawford (1982). In Proposition 4, I show that when the preferences of an institution does not take intersectionality into account, satisfy monotonicity and interior optimum, then the choice rule induced by their preferences does not satisfy the gross substitutes condition of Kelso and Crawford (1982), extending the first result. Both results show that when firms and institutions choose workers or allocate resources and value diversity, intersectionality emerges as an important consideration for selection or allocation procedures to satisfy (gross) substitutes.

Finally, I present a framework to study the choice rules that satisfy gross substitutes and treat diversity and score domains separately. That is, the choice between two individuals can depend on their scores and the representation of their groups, but not the scores of other individuals or representation of other groups. I show that a choice rule is rationalizable by a utility function that is additively separable in score and diversity domains (where the utility is increasing in quality and concave in representation of each group) if and only if it satisfies the gross substitutes condition and an adaptation of the acyclicity condition of Tversky (1964). I then map existing choice rules such as quotas and reserves to this framework. I show how they can be rationalized (in the case of quotas and certain types of reserves) and how they can fail the conditions and therefore cannot be rationlized by preferences that treat scores and diversity domains separately.

Related Literature This paper is contributes to the large literature on matching with affirmative action and diversity concerns initiated by Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu (2005). Most papers in this literature consider models where agents do not have multidimensional and overlapping identities. For example, Kojima (2012) studies quota policies and shows how affirmative action policies that place an upper bound on the enrollment of non-minority students may harm all minority students, Hafalir, Yenmez, and Yildirim (2013) introduce the alternative and more efficient minority reserve policies and Ehlers, Hafalir, Yenmez, and Yildirim (2014) generalize reserves to accommodate policies that have floors and ceilings for minority admission. Dur, Kominers, Pathak, and Sönmez (2018) and Dur, Pathak,

To rule out such cases, I require that the institution has a non-trivial preference for diversity by assuming the most preferred distributions are not the ones that do not include any individuals from a group. This assumption makes sure that diversity plays at least a minimal role in hiring decisions by indicating a preference for representing each group with at least one individual.

and Sönmez (2020) study reserves in Boston and Chicago public schools and show that the precedence order (the order in which reserve and non-reserve portions of the positions are processed) is important for the allocation. Kamada and Kojima (2017), Kamada and Kojima (2018) and Goto, Kojima, Kurata, Tamura, and Yokoo (2017) study stability and efficiency in more general matching-with-constraints models.

Another strand of literature that this paper builds upon is the study of (gross) substitutes conditions of Kelso and Crawford (1982) and Roth (1984). These conditions and their generalizations play an important role in characterizing the existence of competitive equilibria Gul and Stacchetti (1999); Milgrom (2000) and existence and structure of stable allocations Hatfield and Milgrom (2005); Hatfield and Kojima (2010, 2008); Aygün and Sönmez (2013). A number of papers have studied the choice rules that satisfy substitutability. Echenique and Yenmez (2015) characterize the choice rules that regard students as substitutes under some additional axioms. Kojima, Sun, and Yu (2020a) characterize all feasibility constraints that preserve the substitutes condition, which they call the generalized interval constraints. Kojima, Sun, and Yu (2020b) complements the analysis of their previous paper by characterizing when softer pecuniary transfer policies (rather than compulsory constraints) preserve substitutes conditions.[5] My results complement these earlier works by explicitly considering multidimensional and overlapping structure of types of individuals, adopting the alternative approach of formally studying the preferences and characterizing choice rules that satisfy gross substitutes and can be rationalized by additively separable utility functions.

This paper builds on the literature that studies affirmative action with multidimensional and overlapping identities. Kurata, Hamada, Iwasaki, and Yokoo (2017) propose a mechanism called Deferred Acceptance for Overlapping Types that is strategy-proof and obtains the student-optimal matching within all stable matchings when students have strict preferences over reserve and non-reserve positions. Sönmez and Yenmez (2022a) give a detailed account of and analyze the affirmative action policies and the mechanisms that assign government positions to individuals as a function of their merit and socio-economic status in India, where protections for different and overlapping domains such as gender (woman) and underrepresented socio-economic groups (scheduled castes) play an important role. Sönmez and Yenmez (2022a) formalize the main mechanism that has been used in India to match workers to government jobs between 1995 and 2020 (the SCI-AKG choice rule) and

[5] Both papers build on and contribute to the theory of discrete convex analysis, see Murota (1998) and Murota et al. (2016).

show that a lack of consideration for the overlapping nature of the protected identities has led to violations of important properties such as no justified envy and incentive compatibility. These issues has resulted in many court cases across the country and led to the subsequent termination of the mechanism. Sönmez and Yenmez (2022b) allow allocation of multiple resources and generalizes their earlier paper. I show that the SCI-AKG mechanism cannot be rationalized, adding to the deficiencies this rule exhibits. Aygun and Bó (2021) study the affirmative action policies where students can qualify for affirmative action in two different dimensions, in and show that the way overlapping identities are treated in university admissions in Brazil can cause unfairness and incentive compatibility issues (e.g., minority students becoming worse off by declaring they are minorities). Using the admissions data for the year 2013, they document that this is the case for more than 54 percent of the programs and propose an incentive-compatible and fair mechanism that solves these issues.

This paper is also related to my prior work. Çelebi and Flynn (2022a) and Çelebi and Flynn (2021) consider a mechanism designer with an additively separable utility function in quality and diversity domains and study the optimal design of affirmative action policies and coarsenings that map some underlying score to priorities, respectively. The analysis in Section 1.5 complements those papers, by analyzing when do choice rules adopted by institutions are rationalizable by a utility function that is additively separable in quality and diversity domains.[6]

Finally, in a recent paper, Carvalho, Pradelski, and Williams (2022) study the representativeness of the affirmative action policies that do not consider intersectionality in a model with continuum of applicants with heterogeneous qualities and overlapping identities. They show that for generic distributions policies that do not consider intersectionality cannot yield a representative outcome. Even though the models are quite different, the results in this paper complement theirs by focusing on the properties, and not the outcomes, of non-intersectional policies.

Outline Section 1.2 introduces the main model and notation. Section 1.3 characterizes choice rules that can be rationalized by a utility function and when does that utility function is increasing in the scores of individuals. Section 1.4 formally defines intersectionality and (gross) substitutes and shows that preferences must consider intersectionality for the induced choice rule to satisfy the (gross) substitutes condition. Section 1.5 characterizes the choice rules that are rationalizable by an utility function which is additively separable in diversity and quality domains. Section 1.6 concludes.

[6] Chan and Eyster (2003) and Ellison and Pathak (2021) also model the preferences of a school district with a utility function separable across diversity and quality domains.

The proofs of all results are in Appendix A.1.

1.2 Model

There are N dimensions that represent identities of individuals. For each $l \in \{1, \ldots, N\}$, Θ_l denotes the finite set of possible groups to which an individual can belong in dimension l, $\Theta = \Theta_1 \times \ldots \times \Theta_N$ and $|\Theta_l| > 1$ for all l.[7] Each individual has a score $s \in \mathcal{S}$, where $\mathcal{S}$ is a finite set of possible scores.[8] $T = \Theta \times \mathcal{S}$ denotes all possible *types* of individuals.

For each individual i, $\theta_l(i)$ denotes the group of i in dimension l, $\theta(i) = (\theta_1(i), \ldots, \theta_N(i))$ denotes the groups to which i belongs in all dimensions which I refer as the *identity* of i. $s(i)$ denotes the score of individual i and $t(i) = (\theta(i), s(i))$ denotes the *type* of i.

For each set I of individuals, $M_l(I)$ returns the number of individuals that belong to each group in dimension l and $M(I) = (M_1(I), \ldots, M_N(I))$.[9] $M(I)$ is the *marginal distribution* of I, as it returns the number of individuals that belong to each group in each dimension, but does not have any information about the cross-sectional distribution of groups. $N_\theta(I)$ denotes the number of individuals with identity θ in I and $s(I)$ denotes the vector of scores of individuals in I.

There is an institution that chooses q individuals from $\mathcal{I}$, which denotes the set of all individuals. For each I with $|I| \geq q$, 2^I_q denotes all q element subsets of I. A choice rule is a correspondence $C : 2^{\mathcal{I}} \to 2^{2^{\mathcal{I}}}$ such that if $I \in C(I')$, then (i) $I \subseteq I'$ (if I is a chosen set from I', then it must be included in I), (ii) $I \in 2^{\mathcal{I}}_q$ whenever $|I| \geq q$ (the institution fills its capacity whenever there are enough individuals) and (iii) $I = I'$ whenever $|I| < q$ (the institution chooses all individuals if there are not enough individuals to fill the capacity.).[10] Throughout the paper, I assume that the choices of the institution only depends on the scores and groups of the individuals. Therefore, if $t(i) = t(j)$ and $I^* \cup \{i\} \in C(I)$ with $j \notin I^*$, then $I^* \cup \{j\} \in C(I)$.

The relation $\succeq$ denotes the preferences of the institution over $\mathcal{I}$. A choice function C is induced by $\succeq$ if it returns the set of $\succeq$-maximal elements in 2^I_q. The choice function induced by $\succeq$ is denoted by $C_\succeq$. Formally,

[7] For example, Θ_1 can denote gender where $\Theta_1 = \{\text{men, woman}\}$ and Θ_2 can denote income where $\Theta_2 = \{\text{rich, middle class, poor}\}$.

[8] Even though $\mathcal{S}$ is not necessarily an ordered set, I use the terminology of the score as the order will be assumed for the most of the paper.

[9] Continuing the example, if $I = \{i, j\}$ where $\theta_1(i) = \theta_1(j) = \textit{men}$, $\theta_2(i) = \textit{rich}$ and $\theta_2(j) = \textit{poor}$, then $M_1(I) = (2, 0)$, $M_2(I) = (1, 0, 1)$ and $M(I) = \{(2, 0), (1, 0, 1)\}$.

[10] These are referred as capacity filling choice rules in the literature.

$$C_{\succeq}(I) = \begin{cases} \{I' \in 2^I_q : I' \succeq I'' \text{ for all } I'' \in 2^I_q\} & \text{if } |I > q| \\ I & \text{if } |I \leq q| \end{cases}$$

1.3 Rationality and Monotonicity

I begin the analysis by studying the rationality of a choice rule. Let $\prod_1^q T = \mathcal{T}$ denote all possible type distributions q individuals can have. For $|I| = q$, $\tau(I) \in \mathcal{T}$ denotes the types of individuals in I.[11] Note that $\tau \in \mathcal{T}$ is enough to determine the choices, as if $\hat{I} \in C(I)$ and $\tau(\hat{I}) = \tau(\tilde{I})$, then $I' \in C(I)$. I now define a *choice cycle*, which is the main axiom that characterizes the rationality of a choice rule.

Definition 1. *$I_1, \ldots, I_n, I_1$ is a choice cycle if*

- *for each $i < n$, there exists an $\hat{I}_i$ such that $I_i \in C(\hat{I}_i)$ and $I_{i+1} \subset \hat{I}_i$*
- *there exists $\hat{I}_n$ such that $I_n \in C(\hat{I}_n)$, $I_1 \subset \hat{I}_n$ and $I_1 \notin C(\hat{I}_n)$.*

A choice cycle is a violation of *congruence* axiom of Richter (1966). If there is a choice cycle under C, then the institution has chosen I_1 when I_2 was available, I_2 when I_3 is available, ..., I_{n-1} when I_n is available. Therefore, I_1 is indirectly (weakly) revealed preferred to I_n. The fact that I_n being chosen when I_1 is available and is not chosen means that I_n is directly (strictly) revealed preferred to I_1, which contradicts the rationality of the choice rule as I_1 was indirectly revealed preferred to I_n.

A function $U : \mathcal{T} \to \mathbb{R}$ rationalizes C if $C(I) = \{I' : U(\tau(I')) = \max_{\hat{I} \in 2^I_q} U(\tau(\hat{I}))\}$. In other words, the choice rule is rationalized by U if it always chooses the U-maximal sets of individuals. The following proposition shows that the preferences of the institution can be rationalized by some U if and only if C does not admit a choice cycle.

Proposition 1. *C does not admit a choice cycle if and only if there exists a function U (and a preference relation $\succeq$) that rationalizes C.*

Proposition 1 characterizes the rationalizable choice rules, but does not put any restrictions on the utility function. In particular, $\mathcal{S}$ may not even be an ordered set. This will be the case if scores are multidimensional or they can represent the tasks the individual can perform so that it is not always possible to rank two different individuals. However, in most of the applications I consider, the scores are from an ordered set. Therefore, for the rest of the paper, I will assume that $\mathcal{S} \subset [0, 1]$. In this

[11] Formally, $\tau(I) = (t(i_1), \ldots, t(i_q))$.

setting, it is reasonable to expect U to be increasing in $\mathcal{S}$, which I refer as *monotonic* utility functions. We say that $I \trianglerighteq I'$ ($I \triangleright I'$) if I is obtained from I' by (strictly) increasing the scores of some individuals. Formally, $I \trianglerighteq I'$ if there exists a bijection $h : I \to I'$ such that $\theta(i) = \theta(h(i))$, $s(i) \geq s(h(i))$ for all $i \in I$ and $I \triangleright I'$ if at least one of the inequalities is strict. U is increasing in $\mathcal{S}$ if $I \triangleright I'$ implies $U(I) > U(I')$. It turns out that the following generalization of choice cycles characterize the choices that can be rationalized by a monotonic utility function.

Definition 2. *$I_1, \ldots, I_n, I_1$ is a score-choice cycle if*

- *for each $i < n$, either (i) there exists an $\hat{I}_i$ such that $I_i \in C(\hat{I}_i)$ and $I_{i+1} \subset \hat{I}_i$ or (ii) $I_i \trianglerighteq I_{i+1}$.*
- *either (i) there exists $\hat{I}_n$ such that $I_n \in C(\hat{I}_n)$, $I_1 \subset \hat{I}_n$ and $I_1 \notin C(\hat{I}_n)$ or (ii) $I_n \triangleright I_1$.*

A score-choice cycle has additional requirement that the choice rule must prefer higher scoring individuals to lower scoring ones. It is a generalization of congruence axiom and is introduced by Nishimura, Ok, and Quah (2016) to generalize the rationalizability result of Richter (1966) to settings where an exogenous order is present. Applying their result, we can characterize the choice rules that are rationalizable by a *monotonic* utility function.

Proposition 2. *C does not admit a score-choice cycle if and only if there exists a function U that rationalizes C and is increasing in $\mathcal{S}$.*

One might think that lack of choice and score-choice cycles are fairly weak requirements and would be satisfied in all circumstances, which would make the analysis here trivial. However, the choice rule mandated by the Supreme Court of India and have been used in India for 25 years fails these properties. I now define this supreme court mandated choice rule (C_S) when there are two groups in two dimensions. $\Theta_1 = \{g, r\}$ where g denotes the general population and r denotes reserve eligible population (scheduled castes in their setting). $\Theta_2 = \{m, w\}$ denotes the gender of the individuals. I^g denotes the set of general category individuals who are not eligible for reserves. In this simpler setting, the choice rule C_S is characterized by 4 numbers, (q, r, r_w, o_w). q denotes the total number of positions the institution is looking to fill. $r \leq q$ is the number of reserve positions that are only open to reserve eligible individuals, while the remaining $q - r$ positions are open to all individuals. r_w is the number of reserve positions that are protected for woman, where $r_w \leq r$. Similarly, o_w is the number of open positions protected for woman, where $o_w \leq q - r$. In this simpler setting, C_S proceeds as follows.[12]

Supreme Court Mandated Choice Rule C_S

Step 1: Define I^m as the *meritorious reserve candidates*. I^m is the set of reserve eligible candidates who are among the $q - r$ highest scoring individuals in the population.

Step 2: Allocate o_w positions to the highest scoring woman in $I^g \cup I^m$.

Step 3: Allocate remaining $q - r - o_w$ positions to highest scoring individuals who are in the set $I^g \cup I^m$ and are not already allocated to a position.

Step 4: Allocate r positions to r_w highest scoring woman in I^r who are not already allocated to a position.[13]

Step 5: Allocate $r - o_r$ positions to highest scoring individuals in I^r who are not already allocated to a position.

As detailed in Sönmez and Yenmez (2022a), C_S has many shortcomings stemming from the concept of meritorious reserve candidates. If a reserve eligible candidate is not amongst the meritorious reserve candidates, then she is not considered for the open positions that are protected for woman. As a result, C_S does not satisfy *no justified envy*, which means that a reserve-eligible individual can score higher than a reserve-ineligible individual and fail receive a position while the reserve-ineligible individual receives one. As protections are adopted to help the reserve eligible individuals, this goes against the philosophy of affirmative action. Moreover, if there is a woman who does not have a high enough score to be a meritorious reserve candidate, but has a high enough score to receive an open category position protected for woman, then she can receive a position by not disclosing their reserve eligibility, violating *incentive compatibility*. Although these deficiencies are extremely important and led to the subsequent termination of the mechanism, they do not imply that the decision-maker is irrational. Next example shows that C_S admits a score-choice cycle and therefore, cannot be rationalized by a monotonic utility function. Moreover, similar to other two problems, these cycles exist because of the way meritorious reserve candidates are processed and gives yet another reason why the rule is rescinded.[14]

[12] See Sönmez and Yenmez (2022a) for the definition of the choice rule in more general settings.

[13] If at any point there fewer than o_w or r_w woman, then these positions are allocated to highest scoring men who are considered in that stage.

[14] One important detail about C_S it is not capacity filling, it can leave some positions empty if there are not enough individuals who belong to the groups those positions are reserved for. This

Example 1. *$I = \{m_1^g, w_1^r, m_1^r, w_1^g\}$, where m (w) are men (woman) and superscripts denote the groups individuals belong. Capacity and reserves are $q = 3$, $r = 1$, $o_w = 1$ and $r_w = 1$. The scores of individuals are given by*

$$s(m_1^g) > s(w_1^r) > s(m_1^r) > s(w_1^g)$$

As there are two open positions and w_1^r is the second highest scoring individual, $I_m = \{w_1^r\}$. In the first stage, m_1^g and w_1^r are chosen for the open positions. In the second stage, the only remaining reserve eligible individual, m_1^r, is chosen as there is no reserve eligible woman candidate. Therefore, $\{m_1^g, w_1^r, m_1^r\}$ is chosen when $\{m_1^g, w_1^r, w_1^g\}$ is available.

Now consider a setting where individual m_1^r is replaced by $\tilde{m}_1^r$ and $s(\tilde{m}_1^r) \in (s(m_1^g), s(w_1^r))$. Thus, $\{m_1^g, w_1^r, \tilde{m}_1^r\} \triangleright \{m_1^g, w_1^r, m_1^r\}$. In this case, w_1^r becomes the third highest scoring individual and $I_m = \{\tilde{m}_1^r\}$. Since one of the open positions is reserved for woman, in the first stage m_1^g and w_1^g (who is the only woman eligible at this stage) are chosen for the open positions. In the second stage, the only remaining reserve eligible woman, w_1^r is chosen. Thus, $\{m_1^g, w_1^g, w_1^r\}$ is chosen when $\{m_1^g, w_1^r, \tilde{m}_1^r\}$ is available, creating the following score-choice cycle:

$$\{m_1^g, w_1^g, w_1^r\}, \{m_1^g, w_1^r, \tilde{m}_1^r\}, \{m_1^g, w_1^r, m_1^r\}, \{m_1^g, w_1^g, w_1^r\}$$

This example shows that C_S cannot be rationalized by preferences that are increasing in scores. Indeed, m_1^r was chosen at first, but was not chosen after his score has increased. This also shows that C_S does not *respect improvements*, which requires a chosen agent to be chosen after an increase of her score. Respecting improvements and stability are the two main conditions that characterize the Deferred Acceptance Mechanism (Balinski and Sönmez, 1999).The next example shows that the shortcomings of this choice rule are even greater and it cannot be rationalized at all.

Example 2. *$I = \{m_1^g, m_2^g, w_1^r, w_2^r, w_1^g\}$. Capacity and reserves are $q = 3$, $r = 1$, $o_w = 1$ and $r_w = 0$. The scores of individuals are given by*

$$s(m_1^g) > s(m_2^g) > s(w_1^r) > s(w_2^r) > s(w_1^g)$$

As there are two open positions and m_1^g and m_2^g are the two highest scoring individ-

creates some further problems when there are not enough individuals from each group. However, if there are enough individuals from each group, then C_S always allocates all positions. Therefore, my results show that even if availability of individuals from each group is not an issue and C_S allocates all the positions, it is still not rationalizable.

uals, $I^m = \emptyset$ *and only* I^g *are eligible for the open slots. Thus in the first stage,* m_1^g *and* w_1^g *are chosen. In the second stage, highest scoring reserve eligible individual,* w_1^r *is chosen. This means that the set* $I_1 = \{m_1^g, w_1^r, w_1^g\}$ *is chosen when* $I_2 = \{m_1^g, w_1^r, w_2^r\}$ *is available.*

Now consider $\tilde{I} = \{m_1^g, w_1^r, w_2^r, w_1^g\}$*, which removes* m_2^g *from the set of applicants. Then* $I_m = \{w_1^r\}$ *and in the first stage,* m_1^g *and* w_1^r *are chosen. In the second stage, the only remaining reserve eligible individual,* w_2^r *is chosen. Therefore,* I_2 *is chosen when* I_1 *is available and* I_1, I_2, I_1 *is a choice cycle.*

This example shows that the C_S rule not only violate monotonicity, but it also cannot be result of preferences of a rational decision-maker. Examples 1 and 2 provide evidence of shortcomings of the choice rule that has been used in allocating government jobs in India and show how my results can be used to determine the rationality of the choice rules and inform policymakers, assuming they have well defined preferences over the allocations.

As the examples demonstrate, the main problem with this mechanism is the fact that only high scoring reserve eligible individuals are considered for the open positions. In this simple setting with two groups in two dimensions, modifying the mechanism to consider all individuals for the open positions, would solve this issue.[15]

1.4 Intersectionality and (Gross) Substitutes

This section analyzes how institutions evaluate diversity and how this affects the properties of the choice rules they implement. If an institution evaluates diversity of a set of individuals by the marginal, but not the cross-sectional distributions of groups, then it does not consider intersectionality. I show that when the preferences do not consider intersectionality, then the choice rules induced by those preferences fail to satisfy the (gross) substitutes condition, an important property that is required for existence of a competitive equilibrium and a stable matching.

1.4.1 Homogeneous Individuals and the Substitutes Condition

I start the analysis with individuals who are homogeneous in terms of quality, but may belong to different socio-economic groups, assuming $|\mathcal{S}| = 1$ and suppressing any dependence to scores. Since $M(I)$ denotes the marginal, and not cross-sectional distribution of groups, if the preferences of the institution are based on $M(I)$, then

[15] Sönmez and Yenmez (2022a) propose the 2SMG mechanism that remedies the flaws of the previous mechanism in their model general model and was subsequently endorsed by the Supreme Court.

they cannot incorporate intersectionality.

Definition 3. $\succeq$ *does not consider intersectionality if for all* I *and* I' *with* $M(I) = M(I')$, $I \sim I'$.

Definition 3 indicates that the institution considers only marginal, and not cross-sectional, distribution of groups when evaluating the diversity of a set of individuals. As described in the introduction, policies that reserve positions for different groups in different dimensions separately, or diversity statistics that only report the marginal distribution of individuals are examples of such preferences. Next, I state the substitutes condition.

Definition 4. *Let* $\tilde{I} \subseteq I' \subset I$. C *satisfies the substitutes condition if for all* $\tilde{I}$ *with* $\tilde{I} \subseteq \hat{I} \in C(I)$, *there exists* $\bar{I}$ *such that* $\bar{I} \in C(I')$ *and* $\tilde{I} \subseteq \bar{I}$.

This condition is the generalization of the substitutes condition of Roth (1984) to choice correspondences.[16] Substitutes condition states that whenever a set of individuals are chosen from a set I, then the same set of individuals are also chosen from any set $I' \subset I$. The following example illustrates the relationship between intersectionality and the substitutes condition.

Example 3. *Assume* $\succeq$ *does not consider intersectionality. Let* $q = 4$, $\Theta_1 = \{1, 2\}$ *and* $\Theta_2 = \{1, 2\}$, $\mathcal{I} = \{i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8\}$. *Moreover, assume that the institution strictly prefers to have exactly* 2 *individuals from all* 4 *groups (the "most" diverse outcome when the institution form their preferences by the marginal distribution of groups) to any other distribution. The types of individuals are*

	Θ_1	Θ_2
i_1, i_5	1	1
i_2, i_6	1	2
i_3, i_7	2	1
i_4, i_8	2	2

The subset of agents $\hat{I} = \{i_1, i_5, i_4, i_8\}$ *has two individuals from each group and therefore* $\hat{I} \in C_{\succeq}(\mathcal{I})$. *However, if the set of available agents are* $\mathcal{I} \setminus \{i_1, i_5\}$, *then the unique subset of agents that has two individuals from each group is* $\tilde{I} = \{i_2, i_6, i_3, i_7\}$, *and thus* $C_{\succeq}(\mathcal{I} \setminus \{i_1, i_5\}) = \tilde{I}$. *As* $\hat{I} \in C(\mathcal{I})$ *but* $\hat{I} \notin C(\mathcal{I} \setminus \{i_1, i_5\})$, $C_{\succeq}$ *does not satisfy the substitutes condition.*

[16] Note that when C is a choice function (*i.e.*, $C(I)$ is singleton for all I), this condition is equivalent to the following: If $i \in C(I)$ and $I' \subseteq I$, then $i \in C(I')$. A similar generalization is employed by Kojima, Sun, and Yu (2020b) for a model with salaries.

As Example 3 illustrates, when preferences do not consider intersectionality and value diversity, then the substitutes condition might fail. As the institution evaluates diversity by marginal distributions, individuals $\{i_1, i_5\}$ and $\{i_4, i_8\}$ become *complements*: when $\{i_4, i_8\}$ is chosen, the institution already has two individuals from group 2 in both dimensions, and prefers individuals who belong to group 1 in these dimensions. Therefore, when $\{i_1, i_5\}$ is not available, choosing $\{i_4, i_8\}$ cannot be optimal as no two individuals from $\{i_2, i_3, i_6, i_7\}$ can complement $\{i_4, i_8\}$ to attain the preferred distribution.

It is possible that the preferences of the institution put no weight on diversity. This would, for example, be the case if the institution is indifferent between all possible group distributions and therefore indifferent between any two sets of individuals (with q elements) and substitutes condition is trivially satisfied. Therefore, I make a minimal assumption that makes diversity preferences non-trivial and assume that the most preferred diversity level of the institution does not completely exclude a group from being chosen. Formally, $M(I)$ is *at boundary* if I has no individual with some trait, *i.e.*, there exists $\hat{\theta} \in \Theta_l$ such that $\theta_l(i) \neq \hat{\theta}$ for all $i \in I$. Conversely, $M(I)$ is *interior* if it is not at boundary, *i.e.*, if I has at least one individual from each group. The following assumption states that when comparing groups of individuals, compositions that have no individuals from some group are not the most preferred ones.

Definition 5. *$\succeq$ satisfies* ***interior optimum*** *if for all I where $M(I)$ is at boundary, there exists I' where $M(I')$ is interior and $I' \succ I$.*

This is a reasonable assumption for diversity preferences: it requires that the institution values diversity and prefers to choose at least one individual from each group, but puts no other restrictions on how it values different compositions of individuals.

Proposition 3. *If $\succeq$ does not consider intersectionality and satisfies interior optimum then $C_{\succeq}$ does not satisfy the substitutes condition.*

Proposition 3 shows that the logic of Example 3 is indeed much more general: whenever the preferences of the institutions satisfies the mild interior optimum assumption, considering intersectionality is necessary to satisfy the substitutes condition. The proof of this result starts with an arbitrary $\succeq$ that satisfies the two assumptions of the proposition. It then proceeds to determine a particular distribution of groups, similar to the "most" diverse outcome in Example 3 and two sets of agents that complement each other the way $\{i_1, i_5\}$ and $\{i_4, i_8\}$ complements each other in Example 3.

1.4.2 Heterogeneous Qualities and Gross Substitutes

This section extends the analysis to the setting where, in addition to their socio-economic groups, each individual also has a score $s \in \mathcal{S} \subseteq [\underline{s}, \bar{s}]$ with $|\mathcal{S}| > 1$. I assume that holding their identities constant, the institution prefers higher quality agents to lower quality ones.

Definition 6. *$\succeq$ satisfies **monotonicity** if for all $I = \{i_1, \ldots, i_q\}$ and $I' = \{i'_1, \ldots, i'_q\}$ with $\theta(i_k) = \theta(i'_k)$ and $s(i) \geq s(i'_k)$ with at least one element strict, $I \succ I'$.*

Under monotonicity assumption, the scores in this model are analogous to (inverse) salaries in Kelso and Crawford (1982), where a higher salary is worse for the institution. Therefore, I adopt the following gross substitutes definition given in Kelso and Crawford (1982).

Definition 7. *Let $\tilde{I} \subseteq \hat{I} \in C(I)$. Define I' by (weakly) decreasing the scores of all $I \backslash \tilde{I}$. If C satisfies gross substitutes, then there exists $\bar{I}$ such that $\tilde{I} \subset \bar{I}$ and $\bar{I} \in C(I')$.*

Gross substitutes condition requires that if a set of individuals are chosen, and the scores of other individuals decrease, then that set of individuals must still be chosen. I also extend the definition of preferences that do not consider intersectionality to settings with heterogeneous qualities.

Definition 8. *$\succeq$ does not consider intersectionality if $\{s(I), \theta(I)\} = \{s(I'), (\tilde{I}')\}$ implies $I \sim I'$.*

With heterogeneous qualities, an institution does not consider intersectionality is indifferent between two sets of individuals whenever they have the same cross-sectional distribution of groups and the same scores. The following proposition shows that the the relationship between intersectionality and the substitutes condition generalizes to this setting. These results show that how diversity is evaluated is important for choice rule to satisfy certain properties, and in this particular case, seemingly unrelated criterion of intersectionality emerges as a critical consideration for the choice rules to satisfy the substitutes conditions.

Proposition 4. *If $\succ$ does not consider intersectionality, satisfies monotonicity and interior optimum, then $C_\succ$ does not satisfy gross substitutes.*

Proposition 4 is proved by making the appropriate adjustments to the proof of Proposition 3, where decreasing the scores of a set of individuals mirrors the effect of removing those individuals. It shows that whenever the institutions values higher scoring individuals and has a non-trivial preference for diversity, not considering intersectionality when evaluating diversity will cause failure of the gross substitutes condition.

1.5 (Gross) Substitutes and Separability

The analysis in the previous sections has focused on the rationality of the institution and how it evaluates diversity, but it was silent on the possible trade-offs between quality and diversity domains. This section characterizes choice rules that satisfy gross substitutes and are induced by a class of preferences that treat diversity and quality domains separately. Separability is a reasonable property to study; in various settings, the contribution of an individual to an institution is independent from their identity and the institution prefers having a diverse body due to equity concerns, not because it would boost productivity. Moreover, I will show that many (but not all) of the choice rules adopted by different institutions satisfy these conditions and can be rationalized by the preferences I characterize.

For this section, I assume that if $I' \in C(I)$ and $I'' \in C(I)$, then I' is equivalent to I'', in other words, $\tau(I) = \tau(I')$. This means that although C is a choice correspondence as there can be many individuals with same type, it is actually a choice function if we restrict attention to equivalence classes $\mathcal{T}$. A choice rule C satisfies *gross substitutes** if C satisfies both the substitutes condition and the gross substitutes condition.

Let $R(I)$ denotes all agents who are not chosen in some $\hat{I} \in C(I)$. Formally, $i \in R(I)$ if there exists some $\hat{I} \in C(I)$ and $i \notin \hat{I}$. Given C, construct the following binary relation $>_C$, for $\theta(j) \neq \theta(k)$ and $|I| = q - 1$,

$$k \in R(I \cup \{j,k\}) \implies (s(j), \theta(j), N_{\theta(j)}(I \cup \{j\})) >_C (s(k), \theta(k), N_{\theta(k)}(I \cup \{k\}))$$

$>_C$ represents the revealed preference induced by C. For $\theta \neq \theta'$, $(s, \theta, n) >_C (s', \theta', n')$ states that an individual with identity θ and score s is chosen together with $n-1$ agents that shares her identity in favor of an individual with score s', identity θ' with n' other individuals who shares her identity. To economize on notation, let $Q = \{1, \ldots, q\}$ and $D = \Theta \times Q$ denote the set of all (θ, n) with generic element $d \in D$.

Definition 9. *A collection*

$$\begin{aligned}(s_1, d_1) &> (s'_1, d'_1)\\ (s_2, d_2) &> (s'_2, d'_2)\\ &\vdots\\ (s_m, d_m) &> (s'_m, d'_m)\end{aligned}$$

is a cycle if $(s'_1, \ldots, s'_m)$ *is a permutation of* $(s_1, \ldots, s_m)$ *and* $(d'_1, \ldots, d'_m)$ *is a permutation of* $(d_1, \ldots, d_m)$

This definition is due to Tversky (1964) (see also Scott (1964); Adams (1965)) and is used it to characterize preferences that admit an additively separable utility representation. It implies transitivity, but is more restrictive. Existence of a cycle under $>_C$ means that the evaluation of diversity and quality domains are connected, since both $\{(s_i, d_i)\}_{i\leq m}$ and $\{(s'_i, d'_i)\}_{i\leq m}$ are formed from exactly same scores and diversity levels, but $\{(s_i, d_i)\}_{i\leq m}$ are revealed strictly preferred to $\{(s'_i, d'_i)\}_{i\leq m}$ for all i.

Definition 10. *C satisfies **Acyclicity** if there exists no cycles under $>_C$.*

Acyclicity of $>_C$ rules out any connection between the diversity and quality domains. We are now ready to characterize a general class of choice rules that can be induced by a utility function that is separable in diversity and quality domains.

Proposition 5. *If C satisfies gross substitutes* and acyclity, then there exists increasing u and concave $\{h_\theta\}_{\theta\in\Theta}$ such that*

$$U(I) = \sum_{i\in I} u(s(i)) + \sum_{\theta\in\Theta} h_\theta(N_\theta(I)) \tag{1}$$

where U rationalizes C.

For each increasing u and concave $\{h_\theta\}_{\theta\in\Theta}$, C_U satisfies gross substitutes and acyclity.

Proposition 5 shows that as long as the choice rules satisfy gross substitutes* and acyclicity, they can be induced by a utility function given in Equation 1. For a given set of individuals, this utility function evaluates the quality and diversity dimensions separately. The utility from quality dimension is given by $\sum_{i\in I} u(s(i))$, where u is the benefit of choosing an individual with score $s(i)$, which is increasing in $s(i)$. The utility from diversity of the chosen set is given by a set of functions $\{h_\theta\}_{\theta\in\Theta}$, each of which denotes the benefit of choosing a given number of individuals with each identity θ. Moreover, these functions are concave, which means that the marginal benefit of choosing an individual with a given identity is decreasing in the number of such individuals, representing a preference against choosing many individuals with the same identity. Conversely, if the preferences of the institution can be represented by a utility function given in Equation 1, then the choice rules induced by those preferences will satisfy gross substitutes.

A short sketch of the proof is useful for illustrating how these two conditions yield the structure in Proposition 5. First, the choice rule induces an incomplete binary relation $>_C$ over (s, θ, n) tuples. Under acyclicity, this incomplete binary relation can be extended to a complete preference relation which can be represented by an additively separable utility function $u(s) + h(\theta, n)$ over (s, θ, n) tuples. This is an application of the results in Tversky (1964). Next, $h_\theta(N_\theta(I))$ are constructed for each θ using h. At this point, u and $\{h_\theta\}_{\theta\in\Theta}$ represent C for decisions between any two individuals, but not necessarily for decisions over sets of individuals. I then show that under gross substitutes*, the information contained in these binary decisions is actually enough to pin down choices over sets of individuals, yielding the representation.

Many of the choice rules used in different contexts and studied in the literature can be mapped to this framework. For example, a quota policy that restricts admission of individuals with each type θ by some $k_\theta \geq 0$ (Kojima, 2012) can be rationalized by any strictly increasing u and $\{h_\theta\}_{\theta\in\Theta}$ given by

$$h_{\boldsymbol{\theta}}(\tilde{N}_\theta(I)) = \begin{cases} 0 & \text{if } N_\theta(I) \leq k_\theta \\ -q\hat{u} & \text{if } N_\theta(I) > k_\theta \end{cases}$$

where $\hat{u} = u(\overline{s}) - u(\underline{s})$. This indicates that the utility loss in the diversity dimension from going above the quota is larger than any utility gain from the quality dimension. Another example is a reserve policy (see Hafalir, Yenmez, and Yildirim (2013) and Dur, Pathak, and Sönmez (2020)), that reserves r_θ of positions for individuals with each identity. These reserves and remaining open positions are then processed according to some order (which is called a precedence order in the literature) where in each step, highest scoring reserve eligible individuals are chosen. When open slots are processed after all reserve slots, then reserve policies can be rationalized by by any strictly increasing u and $\{h_\theta\}_{\theta\in\Theta}$ given by

$$h_\theta(\tilde{N}_\theta(I)) = \begin{cases} N_\theta(I)\hat{u} & \text{if } N_\theta(I) \leq k_\theta \\ k_\theta \hat{u} & \text{if } N_\theta(I) > k_\theta \end{cases}$$

where $\hat{u} = u(\overline{s}) - u(\underline{s})$. This indicates that the diversity utility is increasing and more important than any gains in the score dimension until the reserve is met and is constant after the reserve requirements are satisfied. However, as the following example shows, if open slots are processed before reserves, the choice rule may fail acyclicity and cannot be represented by an additively separable utility function.

Example 4. *There are 3 positions to be allocated and two groups in one dimension, $\Theta = \{a, b\}$. There is one reserve position for each of the groups. The institution first processes the open position, then group a reserve and then group b reserve. We will consider two applicant sets, I and I'. The following table lists the score of each individual in these cases, where letters denote their groups.*

	a_1	b_1	a_2	b_2
$s(I)$	3	2	1	1
$s(I')$	2	3	1	1

In each case, the open position goes to the individual with a score of 3 *and the other individual from that group receives the reserve position. Under I, $C(I) = \{a_1, b_1, a_2\}$, which implies $(1, a, 2) >_C (1, b, 2)$ and $C(I') = \{a_1, b_1, b_2\}$, which implies $(1, b, 2) >_C (1, a, 2)$. This violates acyclicty and shows that this choice rule cannot be rationalized by a separable utility function.*

1.6 Conclusion

This paper contributes the study of affirmative action and diversity concerns in market design. On the theoretical side, I introduce a model of diversity preferences and establish a connection between decision theory and market design. This model allows me to characterize when choice rules can be rationalized by diversity preferences, when they satisfy the crucial gross substitutes condition and when they can be rationalized by a utility function that treats diversity and quality separately. On the applied side, I show that the mechanism used to match workers to government jobs in India cannot be rationalized by diversity preferences and that considering intersectionality is crucial for the choice rule to satisfy the substitutes condition under multidimensional identities. Moreover, I identify the preferences that induce some well-known choice rules such as quotas and reserves. My results provide a systematic way of evaluating the diversity preferences behind the choices made institutions.

Chapter 2

Adaptive Priority Mechanisms

This chapter is jointly authored with Joel Flynn.

2.1 Introduction

Authorities that allocate resources such as school seats, university places, and medical supplies often face conflicting objectives. On the one hand, they want to maximize match quality or appear fair by allocating resources to the highest-scoring agents according to various criteria such as academic attainment, mortality risk, or distance. On the other hand, they want to achieve diversity across a range of socioeconomic attributes including race, religion, and gender. Resolving this conflict is complicated, especially in new markets, due to uncertainty regarding the distribution of individuals' scores, characteristics, and preferences.

To balance these trade-offs, when the use of prices is seen as infeasible or unethical, authorities have broadly used two classes of policies: *quotas*,[1] where a certain portion of the resource is set aside for given groups; and *priorities*, where individuals in given groups receive higher scores. These policies have been applied across many different markets in many different countries, for example: the Indian government reserves some government jobs for disadvantaged groups; Chicago Public Schools employs quotas for students from different socioeconomic groups at its competitive exam schools; the University of California, Davis instituted a quota system for minority students; many countries gave differential priority to healthcare workers in the receipt of Covid-19 vaccines; church-run schools in the UK give explicit priority points to students from various religious groups; and the University of Michigan and the University of Texas have used different priority scales for minority students.

[1] We use *quota* as a general term that includes the widely used reserve policies (see Definition 14).

But what mechanism *should* such an authority use? Despite its revealed practical importance, we currently possess no formal understanding of this question. Thus, we do not know if (and under what circumstances) an authority should use a priority mechanism, a quota mechanism, or something else entirely.

In this paper, we formulate and solve the optimal mechanism design problem of an authority that allocates a resource to agents who are heterogeneous in their individual scores and belong to different groups. The authority cares about individuals' *scores*, through some aggregate index, and *diversity*, through the numbers of agents from different groups who are allocated the resource.[2] Moreover, they are uncertain about the market they face and have some beliefs about the joint distribution of scores and groups in the population.

We propose a new class of *adaptive priority mechanisms* (APM) that adjust agents' scores as a function of the number of assigned agents with the same characteristics and that allocate the resource to the set of agents with the highest adjusted scores. With a single authority, we derive an APM that is optimal, implements a unique outcome, and can be specified solely in terms of the *preferences* of the authority (*i.e.,* it is optimal regardless of their beliefs). By contrast, we show that priorities and quotas are optimal if and only if risk aversion over diversity is extremely low or high, respectively. Moreover, optimally set priority and quota policies depend on both the preferences and beliefs of the authority. Thus, the optimal APM both improves outcomes and requires less information. When there are many authorities, it is dominant for each of them to implement this APM and this leads to the unique stable outcome but generates inefficiency. To remedy this, we propose a centralized allocation mechanism, an *adaptive priority mechanism with quotas* (APM-Q), that restores efficiency. Finally, we benchmark the quantitative gains from APM using data from Chicago Public Schools and find that they are substantial.

Single-Authority Model We begin our analysis by studying a setting with a single authority that has some amount of a homogeneous resource (*e.g.,* seats at a school, medical resources) that it can allocate to a continuum of agents.[3] Agents

[2]This diversity preference can be interpreted more generally as encoding a preference of the authority over the composition of assigned agents across a range of attributes, *e.g.,* when allocating medical resources, the authority may care about ensuring that frontline medical workers are treated. Moreover, when scores represent individuals' property rights over objects (*e.g.,* higher-scoring students *deserve* better schools), we can interpret the preference for higher scores as a preference for procedural fairness. Whenever a lower-scoring agent obtains the resource while a higher-scoring agent does not, the latter agent has *justified envy* towards the former. In the two-sided matching literature, justified envy is often seen as inimical to fairness (see *e.g.,* Balinski and Sönmez, 1999).

[3]In Appendix B.3, we generalize our analysis and results to a setting with discrete agents.

differ in their scores (*e.g.,* exam score, clinical need) and discrete attributes (*e.g.,* socioeconomic status, if they are a frontline health worker). The authority cares separably about some index of the score distribution (*e.g.,* the average score) of those to whom it allocates the resource and the numbers of agents from different groups. As a result, the preferences of the authority over agents depend on the joint distribution of agents' scores and groups. We assume that this distribution is potentially unknown and varies arbitrarily across states of the world. The authority's problem is to design an allocation mechanism that is optimal regardless of their beliefs, a property that we call *first-best optimality.*

Adaptive Priority Mechanisms To this end, we introduce the class of adaptive priority mechanisms (APM), which proceed in two steps. First, each agent is given an *adaptive priority* that is a function of their own score and the number of agents from the same group to whom the resource is assigned. Second, APM allocate the resource to agents in order of adaptive priorities, subject to fully allocating the available amount. This class of mechanisms allows the implicit preference for agents from different groups to depend upon the ultimate allocation. The allocation under an APM is defined as the fixed point of the above operation: an allocation is implemented by APM if the adaptive priority of all agents who are allocated the resource (evaluated at the allocation) is higher than those who are not allocated the resource. When an agent's adaptive priority is increasing in their own score and decreasing in the number of agents with the same attributes that are assigned the resource – a property we call *monotonicity* – the APM implements a unique allocation. Moreover, this allocation can be computed greedily by prioritizing agents according to their adaptive priority, evaluated at the number of higher-scoring agents in their group.

Most importantly, we derive a particular, monotone APM that is first-best optimal. Under this optimal APM, an agent's priority is equal to the contribution of their own score plus their marginal contribution to diversity utility. Intuitively, this mechanism equates the benefits and costs of allocating to the marginal agent, regardless of the ultimate joint distribution of agents' scores and groups. Moreover, this APM can be described solely as a function of the authority's preferences, without any reference to its beliefs.

(Sub)Optimality of Priorities and Quotas We next establish that priority and quota mechanisms are generally dominated by APM. We do so by characterizing the conditions on the preferences of the authority such that priorities and quotas attain first-best optimality. Concretely, we find that priorities and quotas are first-best optimal if and only if (i) the authority is risk-neutral over diversity, in which case priorities are optimal, or (ii) the authority is extremely risk-averse over diversity, in which case quotas are optimal. Hence, outside of extreme cases, APM deliver strict improvements relative to the *status quo*.

A Price-Theoretic Intuition To both illustrate and develop the intuition behind these results, we study a detailed example that allows for a closed-form comparison of priorities, quotas, and the optimal adaptive priority mechanism. We do this in the spirit of the seminal analysis of Weitzman (1974), who compares price and quantity regulation in product markets. In the example, the resource corresponds to seats at a school and there are two groups of students (minority and majority students). The authority is uncertain over the relative scores of minority and majority students, and has linear-quadratic preferences over the scores of admitted students and the number of minority students admitted to the school.

The preference of the authority between priority and quota mechanisms is governed by its risk aversion over the number of admitted minority students: there is a cutoff value such that quotas are preferred when risk aversion exceeds this threshold and priorities are otherwise preferred. On the one hand, by mandating a minimal level of minority admissions, quotas *guarantee* a level of diversity. On the other hand, as relatively more minority students receive the resource in the states in which they have relatively higher scores, priorities *positively select* minority students. Adaptive priority mechanisms optimally exploit the guarantee effects of quotas and the positive selection effects of priorities, and are always optimal.

Dominance and Stability with Multiple Authorities While the single-authority model is relevant for studying settings with a single resource, in many markets there are multiple authorities who control heterogeneous resources (*e.g.,* school seats) over which agents have heterogeneous preferences. We generalize our analysis to this setting and show that APM arise under both cooperative (stability) and non-cooperative (dominant-strategy equilibrium) solution concepts. Concretely, we show that there is a unique stable allocation and a mechanism is consistent with stability if and only if it coincides with the single-authority-optimal APM. Moreover, when authorities sequentially admit agents, each authority using its single-authority-optimal APM is a dominant strategy and implements the unique stable matching. Thus, one could advise authorities to use APMs with confidence that outcomes will be stable and that they could do no better under any alternative mechanism.

Inefficiency of APM and an Efficient Multi-Authority Mechanism However, decentralized outcomes under APMs are generally inefficient for the authorities. This is because authorities do not internalize the "pecuniary externalities" they generate by over-admitting agents that have a preference for them but that the other authorities value more. To remedy this, we propose a centralized allocation mechanism, an *Adaptive Priority Mechanism with Quotas* (APM-Q). An APM-Q first constructs a fictitious aggregate authority, decides the aggregate levels of admissions of each group according to an optimal APM, and then allocates these groups across authorities according to optimally set quotas. This mechanism fixes the pecuniary externalities by creating a pseudo-market in which each group has a "price" and authorities are assigned agents only if they would be willing to "pay" for them.

Benchmarking the Gains from APM in Chicago Exam Schools Finally, we benchmark the improvements from APM using application and admission data from 2013-2017 on the selective exam schools of Chicago Public Schools (CPS), a setting also empirically studied by Angrist, Pathak, and Zárate (2019) and Ellison and Pathak (2021). CPS uses a reserve system to increase the admissions of underrepresented groups. In this system, as we later detail, academic scores and the socioeconomic tiers of the census tracts in which students live determine the schools that students can attend. Estimating preference parameters to best rationalize the pursued reserve policy, we find that the gains from using the optimal APM are equivalent to eliminating 37.5% of the loss to CPS' payoffs from failing to admit a diverse class of students. This gain is 2.3 times larger than the estimated gain from a 2012 policy change that increased the size of all reserves. This exercise shows both that APM could be practically implemented and that the gains from so doing may be considerable.

Related Literature The market design literature has largely studied the comparative statics and axiomatic foundations of mechanisms. In this context, our paper relates to the literature on matching with affirmative action concerns initiated by Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu (2005). For example, in the study of quotas, Kojima (2012) shows how affirmative action policies that place an upper bound on the enrollment of non-minority students may hurt all students, Hafalir, Yenmez, and Yildirim (2013) introduce the alternative and more efficient minority reserve policies, Ehlers, Hafalir, Yenmez, and Yildirim (2014) generalize reserves to accommodate policies that have floors and ceilings for minority admissions, and Doğan (2016) shows that stronger affirmative action can (weakly) harm all minority students under reserve policies and proposes a new rule that fixes this issue. The quota policies studied in this paper are a special case of the slot-specific priorities introduced in Kominers and Sönmez (2016). Further related papers study quota policies in university admissions in India (Aygün and Turhan, 2020; Sönmez and Yenmez, 2022a,b), in Germany (Westkamp, 2013) and in Brazil (Aygun and Bó, 2021). Kamada and Kojima (2017, 2018) and Goto, Kojima, Kurata, Tamura, and Yokoo (2017) study stability and efficiency in more general matching-with-constraints models. Echenique and Yenmez (2015) characterize a class of substitutable choice rules under diversity preferences and Erdil and Kumano (2019) study tie-breaking rules under substitutable priorities under stable matching mechanisms and distributional constraints. Çelebi (2022) studies when affirmative action policies, including quota policies, can be rationalized by diversity preferences.

In this paper, we instead pursue the methodological approach of mechanism design and welfare economics by analyzing optimal mechanisms from the perspective of an authority with some given preferences over allocations. Chan and Eyster (2003) share this perspective in their analysis of the costs and benefits of banning affirmative action.[4] In this vein, we have previously analyzed the narrower problem of how to optimally coarsen agents' scores into priorities (Çelebi and Flynn, 2022b) in a continuum matching market framework in the style of Abdulkadiroğlu, Che, and Yasuda (2015) and Azevedo and Leshno (2016). This analysis nevertheless restricted authorities to use a priority mechanism that does not consider agents' characteristics and implement only allocations that are stable with respect to these priorities. Thus, our focus on comparing priorities, quotas, and optimal mechanisms distinguishes our analysis from our prior work and the previous literature, which study the properties of each policy in isolation and without an explicit treatment of uncertainty.

Outline Section 2.2 exemplifies our main results. Section 2.3 studies optimal mechanisms with a single authority. Section 2.4 studies equilibrium mechanisms with many authorities. Section 2.5 analyzes efficient mechanisms with multiple authorities. Section 2.6 quantifies the gains from APM using data from Chicago Public Schools. Section 2.7 concludes.

2.2 Comparing Mechanisms: An Example

The Setting A single school has capacity q. Students are of unit total measure, have scores in $[0, 1]$, and are either minority or majority students. The authority has

[4] Other analyses of this issue include Epple, Romano, and Sieg (2008) and Temnyalov (2021).

linear-quadratic preferences $\xi : \mathbb{R}^2 \to \mathbb{R}$ over students' total scores $\bar{s}$ and the measure of admitted minority students x:

$$\xi(\bar{s}, x) = \bar{s} + \gamma \left(x - \frac{\beta}{2} x^2 \right) \tag{2}$$

where $\gamma \geq 0$ indexes their general concern for admitting minority students relative to ensuring high scores and $\beta \geq 0$ indexes the degree of risk aversion regarding the measure of admitted minority students.

The minority students are of measure κ and have scores that are uniform over $[0, 1]$. The majority students are of measure $1 - \kappa$ and all have common underlying score $\omega \in [\underline{\omega}, \overline{\omega}] \subseteq [0, 1]$ with distribution Λ. The score of the majority students, ω, parameterizes how well the majority students score relative to the minority students. Finally, we assume that the affirmative action preference is neither too small nor too large with the following: $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$, $\kappa(1 - \underline{\omega}) < \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta}$. These conditions ensure that optimal affirmative action policies will neither be so large as to award all slots to minority students in some states nor so small that there is no affirmative action in some states.

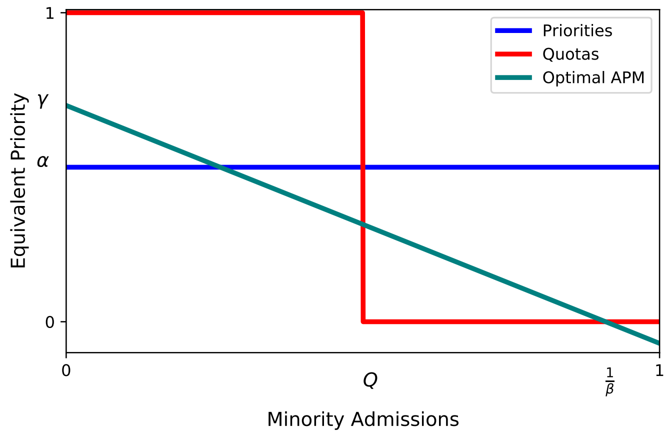
The authority can implement an APM, a priority mechanism, or a quota mechanism. An APM awards a score boost of $A(y)$ to a minority student when measure y other minority students are admitted, and allocates the seats to the highest-scoring minority students who have transformed scores higher than ω.[5] An (additive) priority mechanism $\alpha \in \mathbb{R}_+$ increases uniformly the scores of minority students: the score used in admissions becomes uniform over $[\alpha, 1 + \alpha]$. The authority then admits the highest-scoring measure q students. A quota policy $Q \in [0, \min\{\kappa, q\}]$ sets aside measure Q of the capacity for the minority students. The measure Q highest-scoring minority students are first allocated to quota slots, and all other agents are then admitted to the residual $q - Q$ places according to the underlying score.[6]

We illustrate how these three policies prioritize minority students in Figure 2-1. Priority mechanisms award a constant score boost of α. Quota mechanisms give enough points to always ensure admission until measure Q is reached and then give no advantage. APM allow any pattern of prioritization as a function of minority admissions (we plot only the optimal APM, which turns out to be linear in this context).

[5] Formally, this happens when $s(x(\omega)) + A(x(\omega)) = \omega$, where $s(x(\omega))$ denotes the score of the marginal minority student when the highest-scoring $x(\omega)$ minority students are admitted.

[6] This corresponds to a precedence order that processes quota slots first. We discuss the importance of precedence orders in Section 2.2.1 and in Appendix B.2.3.

Figure 2-1: How Priorities, Quotas, and APM Prioritize Minority Students



Notes: Illustration of the equivalent priority given to a minority student as a function of the measure of admitted minority students under: the optimal APM (see Proposition 6), a priority mechanism α, and a quota mechanism Q.

Comparing Mechanisms Let the authority's expected utility be V^* under any optimal (expected utility maximizing) mechanism, V_A under an optimal adaptive priority mechanism, V_P under an optimal priority mechanism, and V_Q under an optimal quota mechanism. The following proposition characterizes the relationships between these mechanisms:

Proposition 6. *The following statements are true:*

1. *The APM $A(y) = \gamma(1 - \beta y)$ is optimal, $V^* = V_A$*

2. *The comparative advantage of priorities over quotas is given by:*

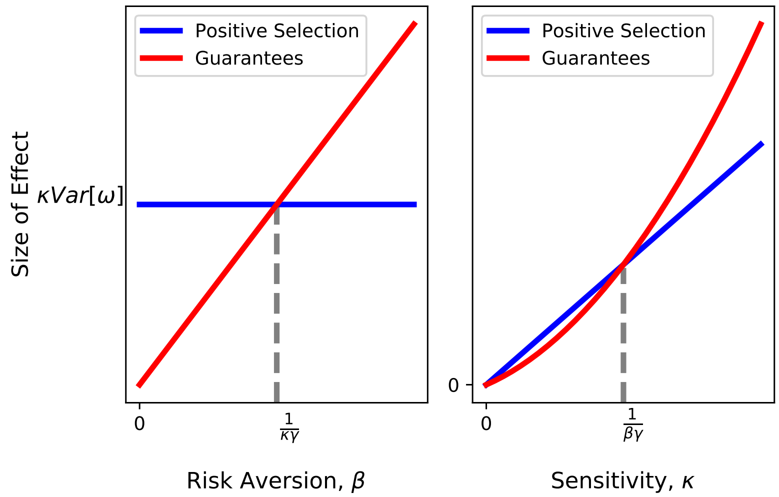
$$\Delta \equiv V_P - V_Q = \frac{\kappa}{2}\left(1 - \kappa\gamma\beta\right) \mathrm{Var}[\omega] \tag{3}$$

3. *The comparative advantage of APM over priorities and quotas is given by:*

$$\Delta^* \equiv \min\{V^* - V_P, V^* - V_Q\} = \begin{cases} \frac{1}{2}\left(\kappa\gamma\beta\right)^2 \frac{\kappa \mathrm{Var}[\omega]}{1+\kappa\gamma\beta}, & \kappa\gamma\beta \leq 1, \\ \frac{1}{2}\frac{\kappa \mathrm{Var}[\omega]}{1+\kappa\gamma\beta}, & \kappa\gamma\beta > 1. \end{cases} \tag{4}$$

The proofs of all results are in Appendix B.1. We now develop intuition for the comparative advantage of priorities over quotas. First, observe that a quota of Q admits measure Q minority students in all states of the world under our assumptions.

Figure 2-2: Comparative Statics for the Positive Selection and Guarantee Effects



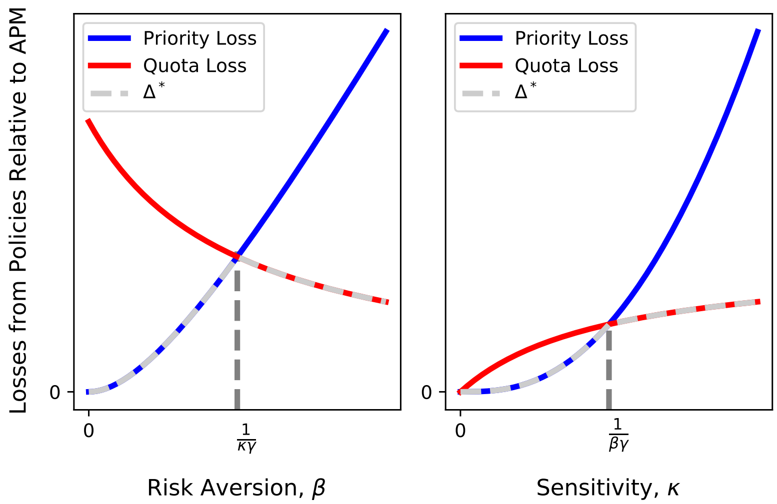
Notes: Illustration of the comparative statics for the trade-offs between priority and quota mechanisms. Positive Selection plots the positive selection effect, $\kappa\text{Var}[\omega]$, and Guarantee plots the guarantee effect, $\frac{\kappa}{2}(1+\kappa\gamma\beta)\,\text{Var}[\omega]$. As per Equation 3 in Proposition 6, priorities dominate quotas if and only if $1 \geq \kappa\gamma\beta$, where the point of indifference is denoted by the dashed grey line.

However, a priority policy induces variability in the measure of admitted minority students across states of the world. We call the gain to quota policies in eliminating this variation the *guarantee effect* and find mathematically that it is equal to $\frac{\kappa}{2}(1+\kappa\gamma\beta)\,\text{Var}[\omega]$ in payoff terms.

Second, the optimal priority policy admits more minority students when minority students score relatively well and fewer when minority students score relatively poorly. To demonstrate this, we show that minority admissions in state ω under the optimal priority policy are $x(\alpha,\omega) = \bar{x}(\alpha) + \varepsilon(\omega)$ where $\bar{x}(\alpha) = \kappa(1+\alpha-\mathbb{E}[\omega])$ and $\varepsilon(\omega) = \kappa\,(\mathbb{E}[\omega]-\omega)$. Thus, in the states where minority students score relatively better ($\omega < \mathbb{E}[\omega]$), we have that $\varepsilon(\omega) > 0$ and $x(\alpha,\omega) > \bar{x}(\alpha)$. We call this effect the *positive selection* effect and find that this benefits a priority policy by $-\text{Cov}[\omega,\varepsilon(\omega)] = \kappa\text{Var}[\omega]$ in payoff terms.

The ultimate preference between priority and quota mechanisms is determined by which of the guarantee and positive selection effects dominates. This is itself determined by the extent to which the authority values diversity γ, the risk preferences of the authority β, and the measure of minority students κ. We illustrate how risk aversion and the measure of minority students affect the sizes of the positive selection

Figure 2-3: Comparative Statics for the Losses from Priorities and Quotas



Notes: Illustration of the comparative statics for the losses from optimal priority and quota policies relative to the optimal APM (as presented in Equation 4 in Proposition 6). The lower envelope of the losses, Δ^*, corresponds to the comparative advantage of the optimal APM over priorities and quotas. The point of indifference between priorities and quotas is denoted by the dashed grey line.

and guarantee effects in Figure 2-2. If the authority is close enough to risk-neutral (*i.e.*, $\frac{1}{\kappa\gamma} > \beta$), then priorities are strictly preferred as positive selection dominates guarantees. If the authority is sufficiently risk-averse (*i.e.*, $\frac{1}{\kappa\gamma} < \beta$), then quotas are strictly preferred as the guarantee effects dominate positive selection. The threshold for risk aversion scales inversely with the measure of minority students κ. Because minority students' scores are uniform, κ corresponds to the density of minority students' scores. Hence, the change in minority admissions from a small change in their priority equals κ. Thus, κ indexes the *sensitivity* of minority admissions to the state under priority policies. As a result, higher κ favors quota policies by increasing the magnitude of the guarantee effect relative to the positive selection effect. Finally, the extent of uncertainty $\text{Var}[\omega]$ may intensify an underlying preference but never determines which regime is preferred.

An APM is optimal and overcomes the limitations posed by both priorities and quotas. In this case, the optimal APM is linear in the measure of admitted minority students, with slope given by the authority's risk aversion over minority admissions, awarding each minority student a subsidy equivalent to their marginal contribution to the diversity preferences of the authority. This allows the adaptive priorities to optimally balance the positive selection and guarantee effects, and implement the first-best allocation in every state. In Figure 2-3, we illustrate how the losses from priority mechanisms and quota mechanisms vary with risk aversion and sensitivity. As risk aversion moves, the loss from priority and quota policies relative to the optimum is greatest when the authority is indifferent between the two regimes. The loss from restricting to priority or quota policies is zero when the authority is risk-neutral or there is no uncertainty regarding relative scores, and decreases as the authority becomes extremely risk-averse. As sensitivity increases, the scope for affirmative action increases and so the gains from APM also increase. Thus, we should expect there to be large gains from switching to APM precisely when authorities have intermediate levels of risk aversion and/or the scope for implementing affirmative action is significant.

Finally, optimal APM have an advantage with respect to priority and quota mechanisms that we have not yet highlighted: they depend only on the authority's preferences, γ and β, and not their beliefs about ω, Λ. This contrasts with the optimal priority and quota policies, which depend on Λ.[7] As a result, APM improve outcomes while using *less* information.

2.2.1 Discussion

Before moving to the general analysis, we discuss three additional findings that emphasize the broader economics and scope of these results.

A Price-Theoretic Intuition This comparison of *priorities vs. quotas* echoes the comparison of *prices vs. quantitities* by Weitzman (1974). We show in Appendix B.2.1 that there is a formal mapping between the two.[8] Intuitively, the positive selection effect is equivalent to the effect that price regulation gives rise to the greatest production in states where the firm's marginal cost is lowest. Moreover, the guarantee effect is equivalent to the ability of quantity regulation to stabilize the level of production. An APM corresponds in the Weitzman (1974) setting to a regulator setting neither a price nor a quantity, but completely specifying the optimal demand curve. As in Weitzman's analysis, this allows the authority to implement the optimal allocation, regardless of the firm's realized marginal costs. Thus, in this context, the comparison of mechanisms for allocating goods without prices boils down to similar

[7] The optimal quota policy is given by $Q^* = \frac{1+\gamma-\mathbb{E}[\omega]}{\frac{1}{\kappa}+\gamma\beta}$, while the optimal priority policy sets the expected measure of minorities to Q^*. The policies depend on Λ through $\mathbb{E}[\omega]$.

[8] Mapping Weitzman's curvature of production costs $C''^{-1} \mapsto \kappa$, curvature of benefits to consumers $B'' \mapsto -\gamma\beta_m$, and uncertainty over marginal costs $\mathrm{Var}[\alpha(\theta)] \mapsto \mathrm{Var}[\omega]$, we have that Weitzman's Δ coincides with our own. See Appendix B.2.1 for more details.

trade-offs between well-understood price-based mechanisms for goods allocation.

Medical Resource Allocation In Appendix B.2.2, we apply this model to understand the trade-offs between priority and quotas in the context of medical resource allocation. This topic received enormous attention during the Covid-19 pandemic (see *e.g.,* Pathak, Sönmez, Ünver, and Yenmez, 2021). Our analysis provides a formal justification for the idea that priorities may lose out relative to quotas from ignoring some groups or ethical values in the allocation of scarce resources (the guarantee effect). However, we also uncover a benefit of priorities that was not previously understood: they induce positive selection. Thus, if we care mostly about treating the neediest (β is low), priorities may yet be optimal.[9]

Optimal Precedence Orders Thus far we have modelled quotas by first allocating minority students to quota slots and then allocating all remaining students according to the underlying score. However, we could have done the opposite. The orders in which quotas are processed are called *precedence orders* in the matching literature and their importance has been the subject of a growing literature (see *e.g.,* Dur, Kominers, Pathak, and Sönmez, 2018; Dur, Pathak, and Sönmez, 2020; Pathak, Rees-Jones, and Sönmez, 2020). In Corollary 1 in Appendix B.2.3, we show that processing quotas second is equivalent to using a priority policy in this setting. Thus, processing quotas first is better than processing them second if and only if $1 \leq \kappa\gamma\beta$. We emphasize that this equivalence merely illustrates the similarity between priority policies and processing quotas second and does not hold in the more general model we study in the remainder of the paper. This notwithstanding, the main aspect of this conclusion is robust: in Theorem 2, we show that for any quota policy to be optimal in the presence of uncertainty, it must process quotas first. In the absence of uncertainty, in Appendix B.5, we show that every quota mechanism is equivalent to a priority mechanism (and *vice versa*) and use this fact to quantify the impact of changes in precedence orders for US H1-B visa allocation.

2.3 Optimal Mechanisms with a Single Authority

We begin our general analysis by studying the resource allocation problem of a single authority. In this context, we define APM and derive an optimal APM that attains the

[9] Inspired by the uncertainty over the number of medical professionals and the general population who were expected to need scarce medical resources at the onset of the Covid-19 pandemic, we consider an extension where there is also uncertainty over the number of medical professionals who get sick, κ. We show that an increase in the uncertainty regarding the need of frontline workers $\mathrm{Var}[\kappa]$ leads to a greater preference for quotas. This can be seen diagrammatically in Figure 2-2 as the guarantee effect is convex in κ.

first-best. We moreover provide necessary and sufficient conditions for the optimality of the ubiquitous priority and quota mechanisms and find that they are extremely restrictive, implying that there are likely gains from switching to APM.

2.3.1 Model

An authority allocates a single resource of measure $q \in (0,1)$ to a unit measure of agents. Agents differ in their type $\theta \in \Theta = [0,1] \times \mathcal{M}$ comprising their scores $s \in [0,1]$ and the group to which they belong, $m \in \mathcal{M}$, where their score denotes their suitability for the resource and $\mathcal{M}$ is a finite set comprising potential attributes such as race, gender, or socioeconomic status. We denote the score and group of any type θ by $s(\theta)$ and $m(\theta)$, respectively. The true distribution of types is unknown to the authority. The authority's uncertainty is parameterized by $\omega \in \Omega$, where Ω is the set of all distributions over Θ that admit a density. The authority believes that ω has distribution $\Lambda \in \Delta(\Omega)$. In state of the world ω, we denote the type distribution by F_ω with density f_ω.[10] In Appendix B.3, we generalize our analysis and results to the discrete context.

An allocation $\mu : \Theta \to \{0,1\}$ specifies for any type $\theta \in \Theta$ whether they are assigned to the resource.[11] Two allocations μ and μ' are *essentially the same* if they coincide up to a measure zero set. The set of possible allocations is $\mathcal{U}$. An allocation is feasible if it allocates no more than measure q of the resource. A mechanism is a function $\phi : \Omega \to \mathcal{U}$ that returns a feasible allocation for any possible distribution of types. A mechanism ϕ implements an *essentially unique* allocation if all allocations implemented by ϕ are essentially the same.

As motivated, authorities often have preferences over scores and diversity. To model this, we define the aggregate score index of any allocation as:

$$\bar{s}_h(\mu, \omega) = \int_\Theta \mu(s,m) h(s) \mathrm{d}F_\omega(s,m) \tag{5}$$

for some continuous, strictly increasing function $h : [0,1] \to \mathbb{R}_+$, which determines the extent to which the authority values agents with higher scores. To capture diversity, we compute the measure of agents of each group allocated the resource $x(\mu,\omega) = \{x_m(\mu,\omega)\}_{m\in\mathcal{M}}$ as:

$$x_m(\mu, \omega) = \int_{[0,1]} \mu(s,m) f_\omega(s,m) \mathrm{d}s \tag{6}$$

---

[10] Formally, we mean that $f_\omega(s,m) = \frac{\partial}{\partial s} F_\omega(s,m)$ exists for all $s \in [0,1]$ and $m \in \mathcal{M}$.

[11] Formally, μ is a measurable function with respect to the Borel $\sigma-$algebra of the product topology in Θ.

To separate the roles of scores and diversity, we impose that their utility function over these dimensions $\xi : \mathbb{R}^{|\mathcal{M}|+1} \to \mathbb{R}$ satisfies the following separability assumption:[12]

Assumption 1. *The authority's utility function can be represented as:*

$$\xi\left(\bar{s}_h, x\right) \equiv g\left(\bar{s}_h + \sum_{m \in \mathcal{M}} u_m(x_m)\right) \tag{7}$$

for some continuous, strictly increasing function $g : \mathbb{R} \to \mathbb{R}$ *and differentiable and concave functions* $u_m : \mathbb{R} \to \mathbb{R}$ *for all* $m \in \mathcal{M}$.

We also assume that the authority always prefers to allocate the entire resource.[13] The preference of the authority is a monotone transformation of a quasi-linear utility index comprised of scores and a diversity preference. Intuitively, u_m determines the preference for assigned agents of group m, with its concavity following from a preference for diversity.[14] The function g determines their risk preferences over their utility over scores and diversity across states of the world.

As we later show, this assumption allows for particularly simple functional forms for optimal mechanisms. We explore the robustness of our results to relaxing this assumption in Appendix B.4. We show that our results are essentially unchanged when preferences are non-separable over diversity, *i.e.,* when $\sum_{m \in \mathcal{M}} u_m(x_m)$ is replaced with $u(x)$. The essential assumption for our results is the separability of diversity and score preferences. When this fails, it is no longer possible to specify optimal mechanisms without explicitly conditioning the allocation on the realized distribution of agents. As a result, our analysis will not apply to situations in which there is complementarity or substitutability in preferences over match quality and diversity.

We define the value of a mechanism ϕ under distribution Λ as the authority's

[12] The assumption of separable preferences over scores and diversity is common in the literature on affirmative action concerns (see e.g., Athey, Avery, and Zemsky, 2000; Chan and Eyster, 2003; Ellison and Pathak, 2021).

[13] A necessary and sufficient condition for this is: $h(0) + u'_m(q) \geq 0$ for all $m \in \mathcal{M}$. This condition is clearly sufficient, the lowest utility the authority can get from allocating the resource is always positive. It is also necessary: if $h(0) + u'_m(q) < 0$ for some m, in the state of the world where there are only measure q of group m agents with uniform score distribution, the authority would prefer not to allocate a portion of the resource to the lowest-scoring agents.

[14] Note that u_m depends on m, so our specification allows the designer to have different preferences for allocating the resource to agents from different groups. For example, this allows for a designer with affirmative action motives who prefers to assign the resource to some particular group m: $u'_m(x) > u'_{m'}(x)$ for all x or a designer who prefers a balanced composition of allocated agents: $u'_m(x) = u'_{m'}(x)$ for all $m \in \mathcal{M}$.

expected utility of the allocations induced by that mechanism:

$$\Xi(\phi, \Lambda) = \int_{\Omega} \xi(\bar{s}_h(\phi(\omega), \omega), x(\phi(\omega), \omega))\mathrm{d}\Lambda(\omega) \tag{8}$$

We say that a mechanism is first-best optimal if it maximizes the authority's expected utility for all possible *distributions of* distributions of agents' characteristics.

Definition 11 (First-Best Optimality)**.** *A mechanism ϕ^* is first-best optimal if:*

$$\Xi(\phi^*, \Lambda) = \sup_{\phi} \Xi(\phi, \Lambda) \tag{9}$$

for all $\Lambda \in \Delta(\Omega)$.

This is a demanding property for a mechanism to possess. Indeed, as the example from Section 2.2 shows, priority and quota mechanisms can fail to be first-best optimal while APM can attain first-best optimality. In the remainder of this section, we formally define APM, show that (when suitably designed) they are first-best optimal, and characterize the conditions under which priorities and quotas are first-best optimal.

2.3.2 Adaptive Priority Mechanisms

Toward deriving a first-best optimal mechanism, we introduce APMs. To this end, we first introduce an *adaptive priority policy* $A = \{A_m\}_{m \in \mathcal{M}}$, where $A_m : \mathbb{R} \times [0, 1] \to \mathbb{R}$. The adaptive priority policy assigns priority $A_m(y_m, s)$ to an agent with score s in group m when measure y_m of agents of the same group is allocated the object. Given an adaptive priority policy, an APM implements allocations in the following way:

Definition 12 (Adaptive Priority Mechanism)**.** *An adaptive priority mechanism, induced by an adaptive priority A, implements an allocation μ in state ω if the following are satisfied:*

1. *Allocations are in order of priorities: $\mu(\theta) = 1$ if and only if for all θ' with $\mu(\theta') = 0$, we have that:*

$$A_{m(\theta)}(x_{m(\theta)}(\mu, \omega), s(\theta)) > A_{m(\theta')}(x_{m(\theta')}(\mu, \omega), s(\theta')) \tag{10}$$

2. *The resource is fully allocated:*

$$\sum_{m \in \mathcal{M}} x_m(\mu, \omega) = q \tag{11}$$

With some abuse of terminology, we will often refer to an APM as the adaptive priority A that induces it. By way of illustration, we provide a simple example of the flexibility of APM to act like a hybrid of priority and quota policies.

Example 5. *Let $\mathcal{M} = \{m, n\}$. Both groups have measure 0.5 and capacity is $q = 0.5$. We consider the adaptive priority policy $A = \{A_m, A_n\}$ given by:*

$$A_m(x, s) = s, \qquad A_n(x, s) = \begin{cases} s + 1 & \text{if } x \leq 0.1 \\ s + 0.1 & \text{if } x \in (0.1, 0.25) \\ s & \text{if } x \geq 0.25 \end{cases} \tag{12}$$

This leaves the score of group m agents unchanged and gives agents of group n a score boost of: 1 if less than measure 0.1 group n agents is assigned, 0.1 if between measure 0.1 and 0.25 group n agents is assigned, and no score boost at all if measure greater than 0.25 group n agents is assigned.

To understand the properties of this adaptive priority policy, observe that the highest-scoring measure 0.1 group n agents is guaranteed the resource, even in states where they score badly. Therefore, A_n practically embeds a quota of 0.1. For admissions levels between 0.1 and 0.25, the APM acts like a priority policy and boosts the scores of group n agents by 0.1, increasing the admissions of group n when group n agents score moderately well. For admissions levels beyond 0.25, group n agents are given no further advantage. Thus, when diversity is attained, this APM "phases out" and no longer advantages any group.

At this point, we have not established that a given APM implements any allocation at all, or that it implements a unique allocation. However, there is a natural subclass of APM that do implement a unique allocation: those that are monotone. An APM A is *monotone* when (i) $A_m(\cdot, s)$ is a decreasing function for all $m \in \mathcal{M}, s \in [0, 1]$ and (ii) $A_m(y_m, \cdot)$ is a strictly increasing function for all $m \in \mathcal{M}, y_m \in \mathbb{R}$.[15]

Proposition 7. *Any Monotone APM A implements an essentially unique allocation.*

Moreover, the unique outcome of a monotone APM can be implemented "greedily:"

Algorithm 1 (Greedy Algorithm for Implementation of APM)**.** *The greedy APM algorithm proceeds in the following four steps:*

[15]Observe that monotone adaptive priority mechanisms are fair in the sense that they preserve the ranking of agents within any group and assign higher priority to an agent whenever there are fewer agents from her group who are allocated the resource.

1. *For each θ, define*

$$\overline{x}(\theta) = \int_{s(\theta)}^{1} f_\omega(s, m(\theta))\mathrm{d}s \tag{13}$$

as the measure of agents who have higher scores than θ and belong to same group.

2. *Construct a ranking of the agents as*

$$R(\theta) = A_{m(\theta)}(\overline{x}(\theta), s(\theta)) \tag{14}$$

3. *Define the cutoff ranking for the agents as $\overline{R}$ by*

$$\int_\Theta \mathbb{I}\{R(\theta) \geq \overline{R}\}\mathrm{d}F_\omega(\theta) = q \tag{15}$$

4. *Allocate the resource to all θ with $R(\theta) \geq \overline{R}$.*

Intuitively, this algorithm works by ranking all agents by their score within each group m and assigning agents in order of their transformed scores evaluated at the measure of *already assigned* agents of the same group, conditional on their admission. Informally, the algorithm greedily moves down the ranking of agents until the resource is exhausted.

2.3.3 Adaptive Priority Mechanisms Achieve the First-Best

Having shown that monotone APM implement a unique allocation and provided an algorithm to compute this allocation, we now show that a certain, monotone APM is first-best optimal:

Theorem 1. *The following APM is monotone and first-best-optimal:*

$$A^*_m(y_m, s) \equiv h^{-1}(h(s) + u'_m(y_m)) \tag{16}$$

Moreover, if a mechanism is first-best-optimal, then it implements essentially the same allocations as A^*.

Observe that A^* is not only uniquely first-best optimal, it also requires only that the authority knows its preferences over scores h and diversity u_m. Importantly, it need have no knowledge of the underlying distribution of agents and can be fully specified even without any knowledge of the nature or extent of uncertainty, Λ. Moreover, this mechanism does not depend at all on the authority's across-state risk preferences,

g. This is because it achieves the *ex post* optimal allocation in all states and so there is no need to trade-off gains and losses across states (the preferences over which are exactly determined by g).

To gain intuition for the form of this mechanism, suppose that the authority has linear utility over scores $h(s) \equiv s$. In this case, $A^*_m(y_m, s) = s + u'_m(y_m)$, so an agent in group m is awarded a boost of $u'_m(y_m)$ when there are y_m higher-scoring agents of the same group, their direct marginal contribution to the diversity preferences of the authority. This is optimal because this boost precisely trades off the marginal benefit of additional diversity with the marginal costs of reduced scores. Moreover, failing to award this precise level of boost would result in a suboptimal allocation. Thus, any optimal mechanisms must be essentially identical to the optimal APM we have characterized. To generalize this beyond linear utility of scores, consider the following observation: we can map agents' scores from s to $h(s)$, and consider the optimal boost in this space. As h is monotone, this preserves the ordinal structure of the optimal allocation, and the authority has linear preferences over $h(s)$. Thus, in this transformed space, the optimal boost remains additive and given by $u'_m(y_m)$. To find the optimal transformed score in the original space, we simply invert the transformation h and apply it to the optimal score in the transformed space, yielding the formula for the optimal mechanism in Theorem 1.

2.3.4 (Sub)Optimality of Priorities and Quotas

We have shown that APM are optimal. However, the primary classes of mechanisms that have been used in practice are priority and quota mechanisms. Therefore, it is important to understand whether (and when) these mechanisms are also optimal. We now establish that APM generally provide a strict improvement over priority and quota mechanisms.

We first formally define priority and quota mechanisms. A *priority policy* $P : \Theta \to [0, 1]$ awards an agent of type $\theta \in \Theta$ a priority $P(\theta)$.

Definition 13 (Priority Mechanisms)**.** *A priority mechanism, induced by a priority policy P, allocates the resource in order of priorities until measure q has been allocated, with ties broken uniformly and at random.*

We define a *quota policy* as (Q, D), where $Q = \{Q_m\}_{m \in \mathcal{M}}$ and $D : \mathcal{M} \cup \{R\} \to \{1, 2, \ldots, |\mathcal{M}| + 1\}$ is a bijection. The vector Q reserves measure of the capacity Q_m for agents in group m, with residual capacity $Q_R = q - \sum_{m \in \mathcal{M}} Q_m$ open to agents of all types. The bijection D (often called the precedence order) determines the order in which the groups are processed.

Definition 14 (Quota Mechanisms). *A quota mechanism, induced by a quota policy (Q, D), proceeds by allocating the measure $Q_{D^{-1}(k)}$ agents from group $D^{-1}(k)$ (if there are sufficient agents from this group) to the resource in ascending order of k, and in descending order of score within each k. If there are insufficiently many agents of any group to fill the quota, the residual capacity is allocated to a final round in which all agents are eligible.*

We now characterize when priority and quota mechanisms are (sub)optimal by providing simple necessary and sufficient conditions on the preferences of the authority that allow us to characterize the optimality of priority and quota mechanisms. To do this, we first provide some definitions. Authority preferences are *non-trivial* if for all $m, n \in \mathcal{M}$, we have that:

$$h(1) + u_n'(0) > h(0) + u_m'(q) \tag{17}$$

Intuitively, the authority's preferences are non-trivial when their concerns for representation of certain groups do not always outweigh the consideration of scores.[16] The authority is *risk-neutral* over diversity if for all $m \in \mathcal{M}$, $u_m' : [0, q] \to \mathbb{R}$ is constant, *i.e.,* there are constant marginal returns to admitting more agents from all groups. If there are decreasing marginal returns, then the authority's preferences feature risk aversion. To define extremely risk-averse preferences, let $\tilde{u}$, $\tilde{h}$ and $\{x_m^{\text{tar}}\}_{m \in \mathcal{M}}$ be as follows: (i) $\tilde{u}_m'(x_m) = 0$ for all $x_m > x_m^{\text{tar}}$ (ii) $\tilde{u}_m'(x_m) \geq \tilde{h}(1) - \tilde{h}(0)$ for $x_m \leq x_m^{\text{tar}}$ and (iii) $\sum_{m \in \mathcal{M}} x_m^{\text{tar}} \leq q$. Intuitively, an authority whose preferences are represented by $\tilde{u}$ and $\tilde{h}$ is very risk-averse as the condition that $\tilde{u}_m'(x_m) \geq \tilde{h}(1) - \tilde{h}(0)$ implies that the loss from being below the target level for a group x_m^{tar} dominates any benefit from increased scores. Thus, they are infinitely risk-averse to failing to meet this target. We say that the authority is *extremely risk-averse* if the authority's preferences over the optimal allocations can be represented by $(\tilde{u}, \tilde{h})$.[17]

Theorem 2. *Suppose that the authority has non-trivial preferences. The following statements are true:*

1. *There exists a first-best optimal priority mechanism if and only if the authority is risk-neutral. Moreover, this mechanism is given by $P(s, m) = h^{-1}(h(s) + u_m')$.*

[16] Note that failure of non-triviality means there exists m and n such that $h(1) + u_n'(0) \leq +h(0) + u_m'(q)$, *i.e.,* a group n agent with the maximum score is less preferred than a group m agent with the minimum score even when all of the entire capacity is allocated to group m agents.

[17] More formally, this means that (u, h) are such that the optimal allocation under (u, h) is also optimal under $(\tilde{u}, \tilde{h})$ for all $\omega \in \Omega$.

2. *There exists a first-best optimal quota mechanism if and only if the authority is extremely risk-averse. Moreover, this mechanism is given by* $Q_m = x_m^{tar}$ *and* $D(R) = |\mathcal{M}| + 1$.

Theorem 2 provides precise conditions on preferences such that the inability of priorities and quotas to adapt to the state is not problematic. That risk-neutrality and high risk aversion are sufficient for the optimality of priority and quota mechanisms is intuitive. On the one hand, if the authority is risk-neutral over the measure of agents from different groups, then they can perfectly balance their score and diversity goals without regard for the state of the world. This is because, under risk-neutrality, there is a constant "exchange rate" between the two: how the authority compares any two agents does not depend on the final allocation and thus can be specified *ex ante* by a priority policy. On the other hand, if the authority is extremely risk-averse as to the prospect of failing to assign x_m^{tar} agents from group m, then a quota allows them to always achieve this target level of allocation in all states of the world while minimally sacrificing scores. It is less obvious that risk-neutrality and high risk aversion are necessary. We prove this result by constructing certain adversarial distributions of agents that render any priority or quota mechanism suboptimal unless the authority is risk-neutral or extremely risk-averse, respectively. Importantly, this result also shows that the only optimal quota mechanisms are those that process open slots last.

This result highlights the fragility of priority mechanisms to uncertainty absent the strong assumption of risk-neutrality over diversity. Intuitively, this is because they feature no guarantees as to how many agents of different groups will be assigned. Indeed, the unfortunate interaction between priority mechanisms and unforeseen market realizations has led to public backlash against priority mechanisms. For example, in the Vietnamese university admissions system, which combines exam scores with priority boosts for students from disadvantaged groups, a year of unexpectedly easy exams led to "top-scoring students missing out on the opportunity to attend their university of choice" and generated backlash against the system (Tuoi Tre News, 2017). Moreover, in the Boston Public Schools system, a priority policy that is set to award students bonus points for high school admissions based on the level of disadvantage of their middle school has made it impossible for students from certain middle schools to get into certain high schools, no matter their grades. This led the Boston Herald (2022) to write that "in Boston, hard work and good grades will only set your child back." As APM dynamically adjust priorities, they have the potential to remedy these practical deficiencies of priority mechanisms.

Moreover, our result highlights that quota mechanisms similarly fail to achieve the first-best away from high levels of risk aversion as they do not take advantage of the potential for positive selection. Our quantitative analysis in Section 2.6 in the context of quota mechanisms in Chicago Public Schools suggests that the substantial variation in the distribution of characteristics across years generates economically meaningful welfare gains from switching to APM.

To formalize the connection between uncertainty and the importance of the adaptability of APM, we consider a setting with *no uncertainty*, where Λ is a Dirac measure. In this context, we say that a mechanism is optimal without uncertainty if it is a utility maximizer.

Proposition 8. *If there is no uncertainty, then there exist optimal priority and quota mechanisms.*

This result shows that if an authority is certain about the market, then appropriately constructed priority and quota mechanisms would be optimal. This formalizes the idea that the suboptimality of priority and quota mechanisms stems from their inability to adapt to the state. Of course, in practice, an authority is always somewhat uncertain of the market they face. Thus, absent the strong conditions on authority preferences that we have characterized in Theorem 2, APM dominate priority and quota mechanisms.

2.4 Equilibrium Mechanisms with Multiple Authorities

The single-authority model is relevant for many resource allocation contexts, such as the medical resource allocation problem of a hospital. However, in other settings such as school or university admissions, multiple authorities must decide upon their admissions policies and rules. In this section, we generalize our single authority model to a setting with multiple authorities. We define *stability* in this setting and show that there is a unique stable allocation. Moreover, we show that a mechanism is consistent with stability if and only if it coincides with the single-authority-optimal APM. We then consider a model where agents sequentially apply to the authorities, who decide which agents to admit. We show that the optimal APM is a dominant strategy. Moreover, we show that in any equilibrium in which authorities use the optimal APM, the resulting allocation corresponds to the unique stable matching of the economy. Taken together, our results provide cooperative (stability) and non-cooperative (dominance) foundations for recommending the use of APM in multi-authority settings.

2.4.1 The Multi-Authority Model

There are authorities $c \in \mathcal{C} = c_0 \cup \bar{\mathcal{C}} = \{c_0, c_1, \ldots, c_{|\mathcal{C}|-1}\}$ with capacities q_c, where c_0 is a dummy authority that corresponds to an agent going unmatched. The agents differ in their authority-specific scores, the group to which they belong, and their preferences over the authorities, $\succ$. We index agents by their type $\theta = (s, m, \succ) \in [0,1]^{|\mathcal{C}|} \times \mathcal{M} \times \mathcal{R} = \Theta$, where $\mathcal{R}$ is set of all complete, transitive, and strict preference relations over $\mathcal{C}$ such that c_0 is less preferred than all $c \in \bar{\mathcal{C}}$. For each type θ, $s_c(\theta)$ denotes the score of θ at authority c and $m(\theta)$ denotes the group of θ. From now, to economize on notation, we suppress indexing states by $\omega \in \Omega$ and let the measure of types be F, with density f.[18] We assume that f has full support over Θ (*i.e.,* $f > 0$) and that the capacity of c_0 is greater $F(\Theta)$.

Each authority has preferences over the agents they are assigned of the form introduced in the previous section:

$$\xi_c(\bar{s}_{h_c}, x_c) = g_c\left(\bar{s}_{h_c} + \sum_{m \in \mathcal{M}} u_{m,c}(x_{m,c})\right) \tag{18}$$

where the extent to which they care about risk g_c, scores h_c, and diversity $\{u_{m,c}\}_{m \in \mathcal{M}}$ are potentially specific to each authority.

A matching is a function $\mu : \mathcal{C} \cup \Theta \to 2^{\Theta} \cup \mathcal{C}$ where $\mu(\theta) \in \mathcal{C}$ is the authority any type θ is assigned and $\mu(c) \subseteq \Theta$ is the set of agents assigned to authority c.[19] Given a matching μ, $\bar{s}_{h_c,c}(\mu)$ and $x_c(\mu) = \{x_{m,c}(\mu)\}_{m \in \mathcal{M}}$ denote the score indices and measures of agents from different groups matched to c at μ. We say that c *prefers* μ to μ', which we denote by $\mu \succ_c \mu'$, if $\xi_c(\bar{s}_{h_c,c}(\mu), x_c(\mu)) > \xi_c(\bar{s}_{h_c,c}(\mu'), x_c(\mu'))$.

Definition 15. *A matching μ is a cutoff matching if there exist cutoffs $S = \{S_{m,c}\}_{m \in \mathcal{M}, c \in \mathcal{C}}$ such that $\mu(\theta) = c$ if (i) $s_c(\theta) \geq S_{m(\theta),c}$ and (ii) for all c' with $c' \succ_\theta c$, $s_{c'}(\theta) < S_{m(\theta),c'}$.*

Given S, the *demand* of an agent θ is their favorite authority among those for which they clear the cutoff:

$$D^\theta(S) = \{c : s_c(\theta) \geq S_{m(\theta),c} \text{ and } c \succeq_\theta c' \text{ for all } c' \text{ with } s_{c'}(\theta) \geq S_{m(\theta),c'})\} \tag{19}$$

The *aggregate demand* for authority c is the set of agents who demand it $D_c(S) =$

[18] Formally, this density is given by $f(s, m, \succ) = \frac{\partial}{\partial s} F(s, m, \succ)$.

[19] The mathematical definition of a matching for the continuum economy we study follows Azevedo and Leshno (2016) and requires that μ satisfies the following four properties: (i) for all $\theta \in \Theta$, $\mu(\theta) \in \mathcal{C}$; (ii) for all $c \in \mathcal{C}$, $\mu(c) \subseteq \Theta$ is measurable and $F(\mu(c)) \leq q_c$; (iii) $c = \mu(\theta)$ iff $\theta \in \mu(c)$; (iv) (open on the right) for any $c \in \mathcal{C}$, the set $\{\theta \in \Theta : c \succ_\theta \mu(\theta)\}$ is open.

$\{\theta : D^\theta(S) = c\}$, while $\tilde{D}_c(S_{-c}) = D_c((0,\ldots,0), S_{-c})$ returns the set of all agents who would demand c if offered admission admission when other authorities' cutoffs are S_{-c}.

2.4.2 Characterization of Stable Allocations

We first characterize the set of stable allocations. Our context presents two challenges in this regard. First, the priorities which are typically used to define stability are not primitives of our model. Therefore, to define stability, we will use the preferences of the authorities induced by Equation 18. Second, unlike discrete models, a single agent does affect the preferences of an authority. Therefore, we need to consider a positive mass of agents to define blocking.

For each matching μ, authority $c \neq c_0$, and two sets of agents $\tilde{\Theta}$ and $\hat{\Theta}$, we let $\hat{\mu}_{(\hat{\Theta},\tilde{\Theta},c,\mu)}$ denote the matching that maps $\hat{\Theta}$ to c and $\tilde{\Theta}$ to c_0 and otherwise coincides with μ.[20] A set of agents $\hat{\Theta}$ *blocks* matching μ at authority c by $\tilde{\Theta}$ if (i) for all $\theta \in \hat{\Theta}$, $c \succ_\theta \mu(\theta)$, (ii) $\tilde{\Theta} \subseteq \mu(c)$, (iii) $F(\tilde{\Theta}) = F(\hat{\Theta})$, and (iv) $\hat{\mu}_{(\hat{\Theta},\tilde{\Theta},c,\mu)} \succ_c \mu$. A matching μ is *not blocked* if there does not exist such a $(\hat{\Theta}, \tilde{\Theta}, c)$. A matching μ satisfies *within-group fairness* if for all $\theta, \theta' \in \Theta$ such that $m(\theta') = m(\theta)$ and $s_{\mu(\theta)}(\theta') > s_{\mu(\theta)}(\theta)$, it holds that $\mu(\theta') \succeq_{\theta'} \mu(\theta)$.[21] A matching μ is *stable* if it satisfies within-group fairness, is not blocked, and all non-dummy authorities fill their capacity. The following result establishes that there exists a unique stable matching and that this is a cutoff matching.

Theorem 3. *There is a unique stable matching. This matching is a cutoff matching.*

This result extends Theorem 1.1 of Azevedo and Leshno (2016) to our setting in which the preferences of authorities are not exogenously fixed and rather depend on the composition of the admitted agents. The stable allocation is characterized by cutoffs S, where each agent θ is matched to $D^\theta(S)$.

To gain intuition for this result, first imagine that there is only one group of agents $|\mathcal{M}| = 1$, so that authorities' preferences are determined by the scores of the agents. Given a set of cutoffs S_{-c}, a cutoff t_c clears the market for c if $F(D_c(t_c, S_{-c})) = q_c$.

[20]Formally,

$$\hat{\mu}_{(\hat{\Theta},\tilde{\Theta},c,\mu)}(\theta) = \begin{cases} c_0 & \text{if } \theta \in \tilde{\Theta} \\ c & \text{if } \theta \in \hat{\Theta} \\ \mu(\theta) & \text{otherwise} \end{cases} \tag{20}$$

[21]Within-group fairness simply requires an authority not to reject a agent if it is admitting a agent from the same group with lower score. Under our assumption that authorities prefer higher scores (h_c is strictly increasing), within-group fairness is satisfied if there is no blocking in discrete models.

When $|\mathcal{M}| = 1$, for a given S_{-c}, there is a unique t_c that clears the market since a smaller cutoff will exceed the capacity while a larger one will leave a positive measure of the capacity empty. Define $T = \{T_c\}_{c \in \mathcal{C}}$, where $T_c(S)$ is the function that maps each S to the market-clearing cutoff t_c under S_{-c}. The result then follows from (i) showing the fixed points of T correspond to market-clearing cutoffs of stable matchings, (ii) establishing that T is monotone, (iii) applying Tarski's fixed point theorem to show that the set of market-clearing cutoffs is a lattice, and (iv) showing that there can only be one market-clearing cutoff as, if there were two, one would strictly exceed the capacities of at least one authority.

When $|\mathcal{M}| > 1$, there is a potential continuum of cutoffs that would clear the market for authority c. A selection from this set is provided by the cutoffs induced by the optimal APM, $A^*_{m,c}(y_m, s) \equiv h_c^{-1}(h_c(s) + u'_{m,c}(y_m))$. We show that the APM cutoffs are unique among the market-clearing cutoffs in being compatible with stability. This is because, for any other t'_c, there is a set $\hat{\Theta}$ of agents (with positive measure) who have scores lower than the cutoff for their group and a set $\tilde{\Theta}$ of agents (with positive measure) who have scores higher than the cutoff for their group, but the authority is strictly better of by admitting $\hat{\Theta}$ and rejecting $\tilde{\Theta}$. We define $T_c(S)$ as the market-clearing cutoffs induced by the optimal APM, show that the fixed points of T_c correspond to market-clearing cutoffs of stable matchings, and follow the same steps as above to demonstrate uniqueness.

This hints at a connection between the stable allocation and the allocation induced by all authorities pursuing the optimal APM, which we now make explicit. The demand set of c at μ, $D_c(\mu)$, is the set of agents who prefer c to their allocation under μ. A mechanism is *consistent with stability* if for all F with stable matching μ_F, it chooses $\mu_F(c)$ from $D_c(\mu_F)$. In other words, evaluated at the set of agents who demand an authority, this mechanism chooses the set of agents with which the authority is already matched. Moreover, we say that a mechanism ϕ is *equivalent* to ϕ' if it chooses the same agents under all full support distributions. We now establish that single-authority-optimal APMs (and equivalent mechanisms) comprise the full set of mechanisms that are consistent with stability.

Theorem 4. *A mechanism ϕ is consistent with stability if and only if it is equivalent to A^*_c.*

Thus, not only is the optimal APM A^* inherent to the structure of stable allocations, but it also characterizes stability in this setting in the sense that any deviation from A^* would result in a violation of stability.

2.4.3 APM Are Dominant Under Decentralized Admissions

We now consider a model where the agents apply to the authorities sequentially, who decide which agents to admit. We index the stage of the game by $t \in \mathcal{T} = \{1, \ldots, |\mathcal{C}|-1\}$. Each stage corresponds to a (non-dummy) authority $I(t)$, where $I : \mathcal{T} \to \mathcal{T}$. At each stage t, any unmatched agents choose whether to apply to authority $I(t)$. Given the set of applicants, authority $I(t)$ chooses to admit a subset of these agents, who are then matched to the authority. Given this, histories are indexed by the path of the measure of agents who have not yet matched, $h^{t-1} = (F, F_1, \ldots, F_{t-1}) \in \mathcal{H}^{t-1}$. Given each history h^{t-1} and set of applicants $\Theta_c^A \subseteq \Theta$, a strategy for an authority returns a set of agents $\Theta_c^G \subseteq \Theta$ whom they will admit such that $\Theta_c^G \subseteq \Theta_c^A$ and $F_t(\Theta_c^G) \leq q_c$ for each time at which they could move $t \in \mathcal{T}$, $a_{c,t} : \mathcal{H}^{t-1} \times \mathcal{P}(\Theta) \to \mathcal{P}(\Theta)$, where $\mathcal{P}(\Theta)$ is the power set over Θ.[22] A strategy for an agent returns a choice of whether to apply to authorities at each history and time for all agent types $\theta \in \Theta$, $\sigma_{\theta,t} : \mathcal{H}^{t-1} \to [0, 1]$.

Within this context, our notion of equilibrium is that of subgame perfect equilibrium:

Definition 16 (Equilibrium)**.** *A strategy profile $\Sigma = \{\{a_{c,t}\}_{c \in \bar{\mathcal{C}}}, \{\sigma_{\theta,t}\}_{\theta \in \Theta}\}_{t \in \mathcal{T}}$ is a subgame perfect equilibrium if $a_{c,t}$ maximizes authority utility given Σ for all $c \in \bar{\mathcal{C}}$ and $t \in \mathcal{T}$ and $\sigma_{\theta,t}$ is maximal according to agent preferences for all $\theta \in \Theta$ and $t \in \mathcal{T}$.*

We moreover say that a strategy $a_{\tilde{c},t}$ for an authority $\tilde{c}$ at time t is *dominant* if it maximizes authority utility regardless of the strategies of all other authorities and agents, $\{\{a_{c,t}\}_{c \in \bar{\mathcal{C}} \setminus \{\tilde{c}\}}, \{\sigma_{\theta,t}\}_{\theta \in \Theta}\}_{t \in \mathcal{T}}$, and the order in which authorities admit agents, I. Moreover, an equilibrium Σ is in *dominant strategies* if $a_{c,t}$ is dominant for all $c \in \bar{\mathcal{C}}$ and $t \in \mathcal{T}$. We denote the unique probabilistic allocation of agents to authorities induced by Σ as $\mu_\Sigma : \Theta \to \Delta(\mathcal{C})$. A probabilistic allocation μ_Σ is deterministic if $\mu_\Sigma(\theta)$ is a Dirac measure on some authority $c \in \mathcal{C}$ for all $\theta \in \Theta$. A deterministic allocation μ_Σ corresponds to a matching μ if $\mu_\Sigma(\theta)$ is a Dirac measure on $\mu(\theta)$ for all $\theta \in \Theta$.

We now establish that the single-authority optimal APM characterizes dominance.

Theorem 5. *A mechanism implements a dominant strategy for an authority if and only if it implements essentially the same allocations as A_c^*.*

The intuition behind this result is that each authority takes as given the set of agents that will accept it. Thus, given this measure of agents, they can do no better

[22] Formally, so that $F_t(\Theta_c^G)$ is well defined, we require that authorities' strategies be measurable in the Borel sigma algebra over Θ.

than to employ the same APM that a single authority would, which is A^*_c by Theorem 1.

Theorem 5 provides a powerful rationale for focusing on APMs in decentralized markets at both positive and normative levels. Normatively, this result allows an analyst to advise an authority regarding how it should conduct its admissions. This is important because any policy that does not coincide with the APM we derive — such as the popular priority and quota mechanisms outside of the cases delimited by Theorem 2 — will disadvantage an authority. Positively, this result allows a sharp prediction that the equilibrium matching between agents and authorities will be the unique stable matching (as per Theorem 3):

Proposition 9. *For all equilibria Σ^* where authorities use A^*, the allocation μ_{Σ^*} is deterministic and corresponds to the unique stable matching of this economy.*

The intuition for this result is that if an equilibrium matching under A^* was not the unique stable matching, then it must be that some agents are applying suboptimally and failing to select the best authority they can attend (according to their preferences), which contradicts that the outcome is an equilibrium.

2.5 Efficient Mechanisms with Multiple Authorities

We have so far characterized the decentralized outcome, but two natural questions remain. First, is the decentralized outcome, which corresponds to the stable allocation, efficient? Second, if not, what kind of centralized solution can remedy any inefficiency? We show that the decentralized outcome is generally inefficient and that a modified, centralized APM mechanism restores efficiency.

2.5.1 Inefficiency of the Decentralized Outcome

The notion of efficiency that we will consider is utilitarian efficiency over authorities. A mechanism in the multi-authority setting is a function $\phi : \Omega \to \mathcal{U}$, where $\mathcal{U}$ is the set of matchings (which, by definition, encodes feasibility requirement imposed in the single authority setting). We define the total authority value Ξ_T of a mechanism ϕ under distribution $\Lambda \in \Delta(\Omega)$ as the total expected utility of the allocations induced by that mechanism:

$$\Xi_T(\phi, \Lambda) = \sum_{c \in \vec{\mathcal{C}}} \Xi_c(\phi, \Lambda) \tag{21}$$

A mechanism is efficient if it maximizes total authority value for all possible distributions:

Definition 17 (Efficiency)**.** *A mechanism ϕ^* is efficient if:*

$$\Xi_T(\phi^*, \Lambda) = \sup_{\phi} \Xi_T(\phi, \Lambda) \tag{22}$$

for all $\Lambda \in \Delta(\Omega)$.

For the remainder of the paper, so that scores are directly comparable across authorities and allocations are interior, we impose the following assumption:

Assumption 2. *Scores and preferences are such that $s_c(\theta) = s_{c'}(\theta)$, $h_c = h$ and $g_c = Id$, where Id is the identity function, for all $c, c' \in \bar{\mathcal{C}}$ and $\theta \in \Theta$. Moreover, $\lim_{x \to +0} u'_{m,c}(x) = \infty$ and $u_{m,c}$ is strictly concave for all $m \in \mathcal{M}$ and $c \in \bar{\mathcal{C}}$.*

Assumption 2 makes scores a common numeraire across authorities and is akin to the standard quasi-linearity assumption in mechanism design. For example, it may be suitable in settings where the score is derived from a common index of academic attainment, such as in Chicago Public Schools. This assumption does not impose that all authorities have common marginal rates of substitution between scores and diversity, as they are allowed unrestricted heterogeneity in diversity preferences. We add the Inada condition for analytical tractability. We argue that it is also reasonable to assume that failing to admit any individuals from a given group is intolerable for authorities.

With the efficiency benchmark defined, we can now demonstrate that the decentralized equilibrium outcome can fail to be efficient.

Proposition 10 (Equilibrium Inefficiency)**.** *All authorities using the privately optimal APMs $\{A^*_c\}_{c \in \mathcal{C}}$ is not necessarily efficient.*

We prove this result via an explicit example with two authorities, c and c' of capacity $\frac{1}{4}$, and two groups of agents, m and m' of measure $\frac{1}{2}$. All agents in group m prefer c' to c and all agents in group m' prefer c to c'. Authority c values admitting group m agents more on the margin than group m' agents, and authority c' values admitting group m' agents more on the margin than group m agents. Using the optimal APMs, both authorities admit more agents of the group whose admissions they value relatively less than the efficient benchmark. The intuition for this is that both authorities "steal" the high-scoring agents of the group whom they relatively less value from the other authority, an externality that they do not internalize.

2.5.2 An Efficient Centralized Mechanism

The inefficiency of each authority using a decentralized APM stems from the implicit incompleteness of markets: if we added the ability for authorities to pay each other for agents in the equilibrium allocation, then they would have willingness-to-pay to do so. A centralized mechanism can remedy this issue by ensuring the cross-sectional allocation of agents to authorities is optimal.

We propose the following augmentation of an APM to solve this problem, an *adaptive priority mechanism with quotas* (APM-Q). The idea behind this hybrid mechanism is to use aggregate, market-level priorities with authority-specific quotas. To this end, an APM-Q comprises an aggregate non-separable APM $\tilde{A} = \{\tilde{A}_m\}_{m\in\mathcal{M}}$ with $\tilde{A}_m : \mathbb{R}^{|\mathcal{M}|} \times [0,1] \to \mathbb{R}$ and a profile of quota functions $Q = \{Q_{m,c}\}_{m,c\in\mathcal{M}}$ with with $Q_{m,c} : \mathbb{R}^{|\mathcal{M}|} \to \mathbb{R}_+$. Intuitively, the aggregate APM pins down the aggregate measures of allocations of each group to *any* authority $\{x_m\}_{m\in\mathcal{M}}$, where $x_m = \sum_{c\in\bar{\mathcal{C}}} x_{m,c}$. The non-separability of this APM simply means that the measures of all groups matter for the adaptive priority of any agent. Given the aggregate measure of allocation for group m, the quota function for authority c assigns $Q_{m,c}(\{x_m\}_{m\in\mathcal{M}})$ agents of type m to authority c.

Definition 18 (Adaptive Priority Mechanism with Quotas)**.** *An adaptive priority mechanism with quotas* $(\tilde{A}, Q)$ *comprises a non-separable APM* $\tilde{A}$ *and a quota function* Q*. An APM-Q implements allocation* μ *if the following are satisfied:*

1. *Aggregate allocations are in order or priorities:* $\mu(\theta) \in \bar{\mathcal{C}}$ *if and only if for all* θ' *with* $\mu(\theta') = c_0$*, we have that:*

$$\tilde{A}_{m(\theta)}(\{x_m(\mu)\}_{m\in\mathcal{M}}, s(\theta)) > \tilde{A}_{m(\theta')}(\{x_m(\mu)\}_{m\in\mathcal{M}}, s(\theta')) \tag{23}$$

2. *Authority-level allocations are given by the corresponding quota functions:*

$$x_{m,c}(\mu) = Q_{m,c}(\{x_m(\mu)\}_{m\in\mathcal{M}}) \tag{24}$$

3. *The resources are fully allocated:*

$$\sum_{m\in\mathcal{M}} x_{m,c}(\mu) = q_c \tag{25}$$

By appropriate choice of the APM and quota functions, we can derive an APM-Q that is efficient. To this end, define the optimally-allocated aggregate utility from diversity:

$$\begin{aligned} \tilde{u}(\{x_m\}_{m\in\mathcal{M}}) = \max_{\{x_{m,c}\}_{c\in\mathcal{C}}} \sum_{c\in\mathcal{C}} \sum_{m\in\mathcal{M}} u_{m,c}(x_{m,c}) \\ \text{s.t.} \sum_{c\in\mathcal{C}} x_{m,c} \leq x_m, \sum_{m\in\mathcal{M}} x_{m,c} \leq q_c, \ \forall m \in \mathcal{M}, c \in \mathcal{C} \end{aligned} \tag{26}$$

Moreover, define the marginal value of aggregate group m admissions $\tilde{u}^{(m)}(y) = \frac{\partial}{\partial y_m}\tilde{u}(y)$ and the marginal value of authority capacity $\tilde{u}_{q_c}(y) = \frac{\partial}{\partial q_c}\tilde{u}(y)$. Using these marginal values, we can design an efficient APM-Q that combines market-level APMs with authority-level quotas:

Theorem 6 (Efficient APM-Q)**.** *Every allocation induced by the following APM-Q* $(\tilde{A}, Q)$ *is efficient:*

1. *The non-separable APM is given by* $\tilde{A}_m(y,s) = h^{-1}\left(h(s) + \tilde{u}^{(m)}(y)\right)$
2. *The quota functions are given by* $Q_{m,c}(y) = \left(u'_{m,c}\right)^{-1}\left(\tilde{u}^{(m)}(y) + \tilde{u}_{q_c}(y)\right)$

The proof of this result constructs a fictitious aggregate authority in our single object setting. The claimed APM is optimal for this aggregate authority by a non-separable adaptation of Theorem 1. The substantial step in this proof establishes that $\tilde{u}$ is increasing, concave, and differentiable by employing the restrictions provided by Assumption 2. Then, given the allocation induced by this APM, we construct the quota function to optimally allocate the level of aggregate admissions induced by the APM.

Intuitively, this mechanism remedies inefficiency by "completing markets." There is a common "market price" for each group given by $\mathcal{P}_m = \tilde{u}^{(m)}(x)$ and an authority-level "shadow price of admissions" $\mathcal{P}_c = \tilde{u}_{q_c}(x)$. Authorities are allocated agents so that the marginal benefit of additional agents equals the sum of the market price and shadow price of admissions $u'_{m,c}(x_{m,c}) = \mathcal{P}_m + \mathcal{P}_c$. Hence, through the completion of markets, a centralized planner can allocate agents efficiently and internalize the externalities that prevented efficiency under the decentralized outcome. Notice that this market involves relatively few prices as it involves only $|\mathcal{M}| + |\mathcal{C}|$ shadow prices rather than the full set of $|\mathcal{M}| \times |\mathcal{C}|$ marginal values.

2.6 Benchmarking the Quantitative Gains from APM

We have so far shown theoretically that APM outperform conventional priority and quota mechanisms. In this section, we attempt to benchmark the magnitude of the gains from implementing APM relative to the reserve system employed by Chicago Public Schools (CPS). To do this, we use application and admission data from CPS for the 2013-2017 academic years. Estimating preference parameters to best rationalize the pursued reserve policy, we find that the gains from using the optimal APM are equivalent to removing 37.5% of the loss to CPS' payoffs from failing to admit a diverse class of students. Our analysis therefore suggests that the gains from APM are considerable.

2.6.1 Institutional Detail on Chicago Public Schools

We first describe the institutional context of CPS. Under current policy, CPS admits students to its selective exam schools based on two criteria. First, CPS ranks students according to a composite score which combines the results of a specialized entrance exam, prior standardized test scores, and grades in prior coursework. This composite score ranges from 0 to 900 and higher-scoring students are admitted before lower-scoring ones, reflecting CPS's desire to allocate seats in exam schools to the students with the best academic standing. In our model, these are the students' scores. Second, CPS divides the census tracts in the city into four *tiers* based on socioeconomic characteristics.[23] Tier 1 tracts are the most disadvantaged, while Tier 4 tracts are the most advantaged. This is reflected in the composite scores of students from Tier 1, who represent 25% of the city's population but comprise relatively few of the high-scoring students (Ellison and Pathak, 2021). As a result, Tier 1 students would have a very small share in the city's top exam schools without affirmative action. To ensure more equal representation across socioeconomic status in these schools, between 2013-2017 CPS implemented a quota policy that reserves 17.5% of the seats for each tier, yielding a total of 70% reserve seats and 30% merit slots that are open to students from all tiers. CPS allocates the seats by first assigning the highest-scoring students (regardless of their tier) to the merit slots and then the highest-scoring students from each tier to the 17.5% reserve seats.

We focus on the most selective CPS school, Walter Payton College Preparatory High School (Payton), which has the highest cutoffs for each tier in each year in our

[23]Concretely, 800 census tracts are divided into four tiers based on six characteristics of each census tract: (i) median family income, (ii) percentage of single-parent households, (iii) percentage of households where English is not the first language, (iv) percentage of homes occupied by the homeowner, (v) adult educational attainment, and (vi) average Illinois Standards Achievement Test scores for attendance-area schools. These characteristics are then combined to construct the socioeconomic score for the tract. Finally, the tracts are ranked according to socioeconomic scores and partitioned into 4 tiers with approximately the same number of school-age children. See Ellison and Pathak (2021) for a more detailed account of the CPS system.

Table 2.1: Admissions Cutoff Scores for Payton

Cutoff Score	2013	2014	2015	2016	2017
Tier 1	801	838	784	769	771
Tier 2	845	840	831	826	846
Tier 3	871	883	877	853	875
Tier 4	892	896	891	890	894

Notes: The table reports the score of the lowest-scoring student that was admitted to Payton in each of the four tiers and five years.

data and would have very few tier 1 students without affirmative action.[24] Table 2.1 presents the cutoff scores of each tier (*i.e.,* the composite score of the last admitted student from each tier).

We make two observations. First, the cutoff students from less advantageous tiers face is lower than the cutoff for students from more advantageous tiers. Therefore, CPS has a *revealed* preference (and not merely a stated preference) for a diverse student body. Second, cutoff scores vary across years. This implies that the distribution of applicant characteristics varies from year to year. Given this uncertainty and the fact that CPS uses a policy that processes quotas after open slots, we know by Theorem 2 that CPS' baseline policy cannot be rationalized as optimal (even if they are extremely risk-averse). Nevertheless, it is always possible that the gains from APM could be small.

2.6.2 Preferences and Estimation Methodology

We perform our analysis in two steps: establishing a parametric framework for evaluating welfare gains and losses and then estimating its parameters.

Preferences We assume a parametric form for CPS's preferences to evaluate the gains from APM. In particular, we impose that the preferences of CPS over the scores and diversity of the student body are represented by the following parametric utility

[24] This approach follows the analysis in Ellison and Pathak (2021), who focus on the two most competitive schools, Northside College Preparatory High School (Northside) alongside Payton. In the years we study, the cutoff scores for Northside are below some other schools frequently, which is why we restrict attention to Payton.

function[25]

$$\xi(\bar{s}, x; \beta, \gamma) = \bar{s} + \sum_{t=1}^{4} \beta |x_t - 0.25|^{\gamma} \tag{27}$$

where $\bar{s}$ is the average score of admitted students, x_t is the percentage of tier t students. Motivated by CPS' desire to allocate the highest-scoring students, ξ is increasing in $\bar{s}$. To model the diversity preferences of CPS, we assume that CPS loses as the gap from equal representation in each tier increases. We do this through the functional form $\beta|x_t - 0.25|^{\gamma}$. The parameters β and γ index the slope and curvature of utility in losses from unequal representation and are the two free parameters of our framework.

Estimation We estimate β and γ to best rationalize the choice of 17.5% reserves for each tier as optimal. We believe this to be a reasonable approach, as the size of the reserves is an important issue that is decided only after much deliberation.[26] Moreover, CPS has used the size of the reserves as a policy tool, increasing them from 15% to 17.5% in 2012 and is currently deliberating another change that would further boost the representation of tier 1 and tier 2 students (Chicago Public Schools, 2022).

Given our functional form, the optimality of the chosen reserve sizes yields moment conditions that we use to estimate the parameters β and γ. Formally, we index reserve mechanisms by the reserve sizes of the four socioeconomic tiers $r = (r_1, r_2, r_3, r_4)$. We let $\bar{s}(r, y)$ and $x(r, y)$ denote the average scores and tier percentages that would be obtained in year y, with distribution F_y, under reserve policy r. The payoff of the policymaker under reserve policy r is given by $\Xi(r, \Lambda; \beta, \gamma)$, as per Equation 8:

$$\Xi(r, \Lambda; \beta, \gamma) = \mathbb{E}_{\Lambda}[\xi(\bar{s}(r, y), x(r, y); \beta, \gamma)] \tag{28}$$

where the expectation is taken over distributions of agents' characteristics F_y under the subjective probability measure Λ. Define the expected marginal benefit of

[25] In Appendix B.6.2, we consider two other parametric utility functions that estimate separate coefficients for underrepresented and overrepresented tiers and only considers loss from underrepresented tiers.

[26] These points are emphasized in Dur, Pathak, and Sönmez (2020): "This change was made at the urging of a Blue Ribbon Commission (BRC, 2011), which examined the racial makeup of schools under the 60% reservation compared to the old Chicago's old system of racial quotas. They advocated for the increase in tier reservations on the basis it would be "improving the chances for students in neighborhoods with low performing schools, increasing diversity, and complementing the other variables."

increasing reserve i and decreasing reserve j as:

$$G_{ij}(r, \Lambda; \beta, \gamma) = \frac{\partial}{\partial r_i}\Xi(r, \Lambda; \beta, \gamma) - \frac{\partial}{\partial r_j}\Xi(r, \Lambda; \beta, \gamma) \tag{29}$$

Any (interior) optimal reserve policy r^* must equate the expected marginal benefit of increasing reserve i and decreasing reserve j at r^* to zero for all (i, j) pairs, *i.e.,* $G_{ij}(r^*, \Lambda; \beta, \gamma) = 0$ for all $\{i, j\} \subset \{1, 2, 3, 4\}$ such that $j > i$. These six first-order conditions yield six moments.

We take empirical analogs of the theoretical moments and estimate preference parameters by minimizing the sum of squared deviations of these moments from zero. We take CPS' pursued reserve policy as optimal, $\hat{r}^* = (0.175, 0.175, 0.175, 0.175)$. We estimate the empirical joint distribution of students' scores and tiers in CPS in each year $\hat{F}_y$ for $y \in \{2013, 2014, 2015, 2016, 2017\}$ and estimate $\hat{\Lambda}$ as a distribution that places equal probability on each of these five measured distributions. We plug these sample estimates into the theoretical moment functions. This yields six empirical moment functions that depend only on the preference parameters, $G_{ij}(\hat{r}^*, \hat{\Lambda}; \beta, \gamma)$. Motivated by the theoretical necessity of $G_{ij}(r^*, \Lambda; \beta, \gamma) = 0$, we estimate the preference parameters by minimizing the sum of squared deviations of the empirical moments from zero:

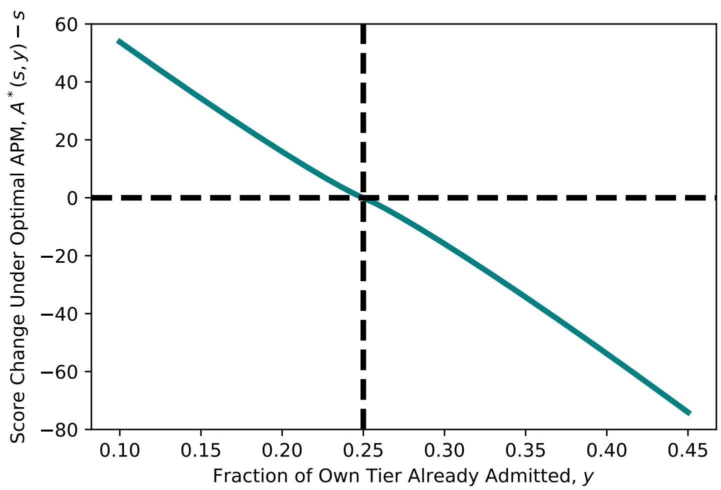
$$(\beta^*, \gamma^*) \in \arg\min_{\beta,\gamma} \sum_{i=1}^{4} \sum_{j>i} G_{ij}(\hat{r}^*, \hat{\Lambda}; \beta, \gamma)^2 \tag{30}$$

Performing this estimation yields estimated parameter values of $\beta^* = -209.5$ and $\gamma^* = 2.11$.

2.6.3 The Estimated Gains from APM

We now use our estimated model to quantify the welfare gains from using APM. To do this, we compare the empirical payoff $\Xi(\phi, \hat{\Lambda}, \beta^*, \gamma^*)$ under two mechanisms: the pursued quota policy, r^*, and the optimal APM from Theorem 1, A^*. In Figure 2-4, we illustrate how the estimated optimal APM changes students' scores to arrive at their ultimate priorities. In accordance with the preferences we have assumed, students receive a score boost when their tier is underrepresented and a score penalty when their tier is overrepresented. As we found $\gamma^* = 2.11$, the estimated diversity preference is very close to quadratic. Thus, the optimal APM is very close to linear. From Theorems 1 and 2, we know that this APM achieves the first-best allocation in each year while the implemented quota policy does not. However, our theorems do not guarantee that the gains from APM are economically meaningful.

Figure 2-4: The Estimated Optimal APM



Notes: This figure plots the change in a student's score when fraction y of students in their own tier has already been admitted under the estimated optimal APM, A^*. At $y = 0.25$ (the vertical dashed black line), the score in unchanged. For $y < 0.25$, students receive a score boost. For $y > 0.25$, students receive a score penalty. The range of the x-axis, $[0.1, 0.45]$, is chosen to cover the full range of fractions of admitted students under both the optimal and the CPS reserve policy from all tiers in all years of our sample (see Figure 2-5).

The empirical payoff under APM is 876.9, while it is 874.8 under the CPS reserve policy. Thus, the gains from APM are equivalent to increasing average scores by 2.1, holding diversity fixed. To benchmark the size of the gains, we require units in which they can be meaningfully expressed. We define the *loss from underrepresentation* as the payoff lost by CPS under its baseline policy from not admitting a fully balanced class, while holding fixed the average score of the class. This is equal to 5.6 points under our estimated parameters. Thus, the gains from APM are equal to 37.5% of the loss from underrepresentation incurred under the CPS policy.[27] We define *score cost of diversity* as the difference between the average scores of admitted students without any affirmative action (893.8) and the average score under the CPS policy (880.5). Thus, the gains from APM are equal to 15.7% of the score cost of diversity. Finally, we can compare the gains from switching to the optimal APM to the gains from the 2012 (the year before our sample) CPS reform, which increased the size of all reserves from 15% to 17.5%. Under the estimated preferences, the empirical payoff under the 15% reserve rule is 873.9, and so the gains from the reform are equivalent to increasing average scores by 0.9. Thus, the gains from switching to the optimal APM are 2.3 times larger than the gains from this recent reform.

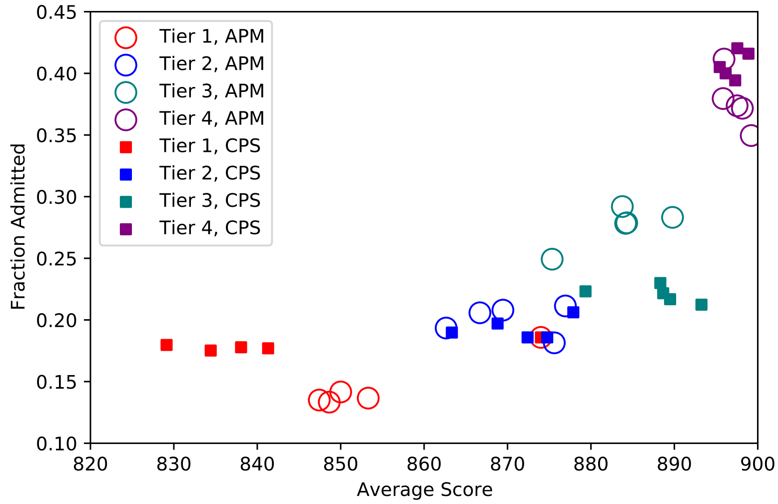
These estimates suggest that the gains from APM are economically meaningful. These gains stem from the variation across years in the joint distribution of student scores and tiers. This can be seen in Table 2.1, which shows the variability in the scores of the marginally admitted students from tiers 1, 2 and 3. More systematically, we visualize the difference in outcomes under CPS' reserves and the optimal APM by plotting the average scores and fraction admitted for each tier for each year under both mechanisms in Figure 2-5. There are two main differences between the allocations. First, the APM allocates systematically fewer tier 1 and tier 4 students and more tier 3 students. Second, the APM admits a greater fraction of students from each tier (especially tiers 1 and 3) in the years in which that tier scores well. This positive selection generates the welfare gains.

Robustness We now explore the robustness of APM to the three core assumptions of our analysis: (i) that CPS has the correct beliefs about the distribution of distributions of students, (ii) that CPS has preferences that lie in the assumed parametric family, and (iii) that CPS separately optimizes the sizes of all four tiers.

Our baseline analysis took the beliefs of CPS to be the true empirical distribution

[27] This is equivalent to increasing the percentage of students from tier 1 from 0.179 to 0.21 and decreasing the percentage of students from tier 4 from 0.407 to 0.378. This corresponds to swapping 8.7 students from the most overrepresented group (tier 4) for the most underrepresented group (tier 1) each year.

Figure 2-5: Comparing Admissions under the Optimal APM and the CPS Policy



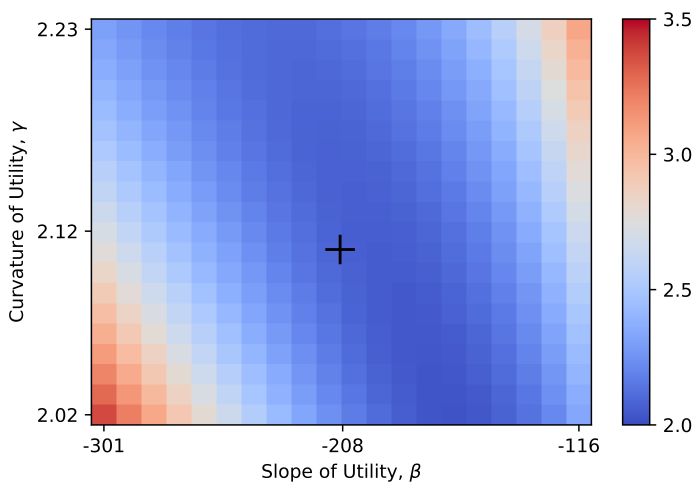
Notes: Each point corresponds to one of the four tiers of students in one of the five years under either the optimal APM or the CPS policy. The x-axis corresponds to the average score of those admitted from that tier in that year under that policy. The y-axis corresponds to the fraction of admitted students from that tier in that tier under that policy.

of student distributions over years. To test robustness to this assumption, we take $\hat{\Lambda}$ as a Dirac distribution on the realized distribution for each of the five years of our data and re-estimate the preference parameters. In Figure 2-6, we plot the difference in welfare under the optimal APM and CPS reserve policies over the full range of these re-estimated parameters (*i.e.,* we take the minimum and maximum of the estimated parameters across years as the ranges for the axes). We find that the gains from APM range from 2.0 to 3.5, while our baseline estimate was 2.1. Thus, the point estimate of our welfare gains from APM appears to be conservative by this metric.

To gauge robustness to the functional form we have assumed for CPS' preferences, in Appendix B.6.2, we estimate two different parametric specifications of utility. First, we consider a utility function that includes a loss term only for underrepresented tiers (and does not penalize overrepresentation of any tier). Second, we allow for CPS to care differentially about underrepresentation and overrepresentation by considering a utility function with separate coefficients for underrepresented and overrepresented tiers. We find that under these specifications, the improvement from APM corresponds to 9.7% and 8.7% of the loss from underrepresentation, which is attenuated relative to our baseline, but remains considerable.

To study the robustness of our findings to the assumption that CPS separately

Figure 2-6: Robustness of the Gains from APM



Notes: This chart plots the difference in empirical payoffs from the optimal APM and CPS reserve policy under alternative parameter values, with the shaded colors corresponding to the numerical value of the gains from APM, ranging from 2.0 to 3.5. The black '+' indicates our baseline parameter values. The ranges for the axes are obtained by estimating β and γ separately for each year of our data and separately taking the minimum and maximum estimated values of each set of estimated parameters.

optimizes the size of all four tiers, in Appendix B.6.1 we consider a setting where CPS sets a *single* reserve size for all tiers. As we now have only one moment condition, we vary γ over the interval [1,10], estimate $\beta^*(\gamma)$ as the exact solution to the moment condition, and compute the gains from APM as a function of γ. The *minimum* gain from APM over the estimated range is 1.98 points, which corresponds to 26.2% of the loss from underrepresentation under that parameterization. This is slightly smaller than our baseline estimate but still considerable.

Limitations Finally, we state some limitations of our analysis. First, even though we argue that our functional form assumptions are reasonable and parsimonious in this setting, there are possibly many other parametric utility functions that might represent the preferences of CPS. Unfortunately, with the available data, richer methods of preference estimation that allow for higher dimensionality of preference parameters are severely limited. This is because admissions rules are only set once and we can therefore use only six first-order conditions. Richer data in which we observe choices of students from different applicant sets (either in practice or in hypothetical choice settings) would allow for a more detailed analysis. Second, one of the main aims of the tier system employed by CPS is to increase racial diversity in the prestigious exam schools. Indeed, the pursued tier system is a race-neutral alternative that replaced the previous race-based system following two Supreme Court Rulings in 2003 and 2007 (see Ellison and Pathak, 2021, for a summary). Because of this, CPS uses tiers based on socioeconomic status instead of race and so we estimate their preferences over tiers. This notwithstanding, if admission rules could depend on race, then one could perform a similar analysis in which APM simply prioritize based on race rather than tier.

2.7 Conclusion

Motivated by the use of priority and quota policies in resource allocation settings with diversity concerns, we consider an authority that has separable preferences over scores and diversity. We introduce Adaptive Priority Mechanisms (APM) and characterize an APM that is both optimal and can be specified solely in terms of the preferences of the authority. We also study the priority and quota policies that are used in practice and show that they are optimal if and only if the authority is either risk-neutral or extremely risk-averse over diversity. Analyzing a setting with multiple authorities that dynamically admit agents, we show that the optimal APM is a dominant strategy. Thus, one could potentially advise authorities to follow an optimal APM with confidence (under our maintained assumptions on preferences) that they could do no better. Moreover, all authorities following the optimal APM implements the unique stable allocation. This notwithstanding, the stable allocation can fail to be efficient for the authorities. We propose a centralized adaptive priority mechanism with quotas to remedy this.

Our analysis has potential implications for improving the design of real-world allocation mechanisms. First, we show that while both priorities and quotas can be better than one another (depending on the risk preferences of the authority), they are generally suboptimal. Second, we show how to improve upon these mechanisms using APM that harness the strengths of these policies: APM benefit both from the guarantee effect of quotas in ensuring certain levels of admissions from various groups and the positive selection effect of priorities in expanding affirmative action when it is least costly. Our quantitative analysis using CPS data suggests that the use of APM can yield considerable welfare gains over the *status quo*. On the basis of our analysis, we conclude that APM may have a real-world use in delivering more desirable allocations of resources. Moreover, by virtue of their generality, APM could be applied in many settings, including the allocation of seats at schools, places at universities, and medical resources to patients.

Chapter 3

Best Response Dynamics in the Boston Mechanism

3.1 Introduction

The Boston Mechanism (BM) is a centralized assignment mechanism that has been widely used in many parts of the world to assign students to schools. The most important feature of BM is that it is not strategy-proof: students lose their priority in schools they rank lower in their preference lists to students who rank those schools higher. Starting with the characterization of equilibria of BM (Ergin and Sönmez, 2006), a large body of literature has analyzed BM and compared it to the widely-used Deferred Acceptance Mechanism (DA), which implements the Student Optimal Stable Matching (SOSM) as a dominant strategy equilibrium.

This paper contributes to the literature on BM by asking the following question: (when) should we expect an equilibrium of BM to be played? Motivated by the fact that BM is used repeatedly across periods in many markets,[1] I set up a multi-period model where each period students submit rankings of the schools to a centralized clearinghouse which uses BM to determine the allocation. In the baseline model, which I call the Repeated Boston Mechanism (RBM), in each period, students play a best response to the admission cutoffs of the previous period, with the exception of the initial period, where they apply truthfully. This paper focuses on characterizing the (non-)convergence of the best response dynamics under RBM and its modifications.

My analysis builds on a surprisingly close connection between the steps of DA

[1] Examples of settings where BM or variants are used include college admissions in China (Chen and Kesten, 2017), public school systems in Charlotte (Hastings, Kane, and Staiger, 2009), Beijing (He, 2015) and Barcelona (Calsamiglia, Fu, and Güell, 2020). See Agarwal and Somaini (2018) for more examples and details.

and the periods of RBM.[2] It turns out that each period of RBM is analogous to the corresponding step of a slightly modified but equivalent version of DA in the sense that, the set of students who are (tentatively) matched to each school is identical. My main result, Theorem 7, shows that the matching implemented in RBM converges in finitely many periods to the Student Optimal Stable Matching (SOSM), which is the dominant-strategy outcome of DA and the Pareto-dominant equilibrium of BM. Theorem 7 provides a foundation to the equilibrium analysis of BM based on best response dynamics by showing that if BM is repeatedly used in the same market in consecutive periods and students form their strategies by myopically best responding to previous period's cutoffs, the implemented matching converges to a stable matching in finitely many periods. The rest of the paper relaxes some assumptions of RBM and to analyzes convergence in more general settings.

Section 3.4 considers an environment with unsophisticated students, who always report their true preferences, and sophisticated students, who strategize as they do under RBM. Pathak and Sönmez (2008) characterize the equilibria of BM in this setting, show that unsophisticated students lose their priorities to sophisticated ones and sophisticated students prefer the Pareto-dominant Nash equilibrium of BM to DA. I show that RBM with unsophisticated students converge in finitely many periods to the Pareto-dominant equilibrium of BM with unsophisticated students, providing a foundation to their equilibrium characterization and their equilibrium selection for the comparison between BM and DA.

Section 3.5 relaxes my assumptions on the first period behavior and allows students to best reply to some initial cutoff scores, instead of applying truthfully. The initial cutoff scores determine the optimism of students about their admission chances: if cutoff scores are low, they believe that their score is enough to obtain a place in most schools, while if cutoffs are high, they believe most schools are not achievable. I show that if students are initially optimistic enough about their admission chances, then RBM converges to the SOSM, while if students are pessimistic (in particular, if they believe their school under SOSM is not achievable), then it is possible that RBM does not converge and implements matchings that are Pareto-dominated by the SOSM in each period. These non-convergence results are in line with the observations of non-equilibrium play under BM in real-world markets (He, 2015; Kapor, Neilson, and Zimmerman, 2020; Song, Tomoeda, and Xia, 2020).

In the baseline model, I assume that same, finite market is repeated in each period.

[2] In both BM and DA, students submit a ranking of schools to the mechanism, which then proceeds iteratively and returns a matching of students to schools as the outcome after finitely many steps.

However, it is plausible that there is some randomness in the market across periods. To deal with this, Section 3.6 extends the analysis to large, random markets. First I extend results to a market with continuum students (Abdulkadiroğlu, Che, and Yasuda, 2015; Azevedo and Leshno, 2016). Then, I study the best response dynamics when a different and random finite market is sampled from a stationary continuum market each period and show that for each period, the matching implemented in the finite random market under RBM converges to the matching implemented in the continuum market as the finite market grows large. This shows that the results are robust to uncertainty in large markets.

Related Literature. This paper contributes to the literature on school choice. Abdulkadiroğlu and Sönmez (2003) formalize and study the Boston Mechanism and compare it with its alternatives, DA and the Top Trading Cycles (TTC) mechanisms. The equilibria of BM is characterized in Ergin and Sönmez (2006) and this characterization is extended to unsophisticated students in Pathak and Sönmez (2008). Chen and Kesten (2017) characterize a parametric family of mechanisms, the application-rejection assignment mechanisms, where BM and DA constitute two limiting cases and empirically demonstrate that authorities move away from BM towards stable mechanisms such as DA. Dur, Pathak, Song, and Sönmez (2022) study the assignment mechanism in Taiwan, which is a hybrid of BM and DA. Akbarpour, Kapor, Neilson, van Dijk, and Zimmerman (2022) show that when students have the same ordinal preferences and some have outside options, students with outside options prefer manipulable BM to DA. Abdulkadiroğlu, Che, and Yasuda (2011) study BM from an ex-ante perspective where students do not know their priorities and have identical ordinal preferences and show that the equilibria of BM generate higher ex-ante welfare for each student than the dominant strategy outcome of DA. In a similar setting, Babaioff, Gonczarowski, and Romm (2018) show that it is possible that non-trivial fraction of students prefer to be sincere rather than strategic under BM.

Dur, Hammond, and Kesten (2021) study a related model where students sequentially submit their preferences to BM and DA mechanisms and compare the efficiency of equilibria. Most relatedly, they also consider a model where students can update their submitted preferences as many times as they want and observe the latest submission of each student while they are resubmitting their preferences. When all students rank the best achievable school first, they show that this process converges to SOSM. My results complement their work by considering a setting where students observe the last period's cutoff scores and submit their preferences once, rather than a setting where students can update their preferences observing what others have done. Thus, this paper manages to uncover the relationship between repeated BM and DA, show that convergence may be attained under repeated application of BM even if students submit preference once and demonstrate how convergence may depend on the beliefs of students and their sophistication.

Zhang (2021) studies the equilibria of BM in a setting where students engage different levels of strategic reasoning, using the level-k model. He defines Fast DA, a modified version of DA that I refer as the Modified DA (MDA) in this paper, and shows that under certain assumptions about students beliefs and common knowledge of preferences and priorities, the matchings in the rounds of Fast DA are analogous to the equilibrium matchings under corresponding levels of strategic reasoning. Proposition 11 in this paper is similar to his result as I establish a similar relationship between the rounds of MDA and periods of the RBM, extending the connection between DA and BM to best response dynamics.

This paper also builds on the empirical literature that study manipulable mechanisms, and BM in particular. He (2015) examines choice data from Beijing where the Boston mechanism is used, and show that parents are overcautious and play safe strategies too often. Agarwal and Somaini (2018) find evidence that students engage in strategic behavior under BM. Song, Tomoeda, and Xia (2020) analyze college admissions in China and find that equilibrium is not being played. They show that two different types of behavioral students, unsophisticated students who reveal their preferences truthfully and cautious students, who are pessimistic about their admission chances, make up a large portion of student populations. The extensions that show non-convergence to SOSM under unsophisticated or initially pessimistic students are in line with their results. Kapor, Neilson, and Zimmerman (2020) demonstrates that students' beliefs about admissions chances differ from rational expectations values and affect their choices. They evaluate the effects of switching to DA, and of improving households' belief accuracy, and find that both would improve welfare.

3.2 Model

Let $I = \{i_1, ..., i_n\}$ denote the set of students and $C = \{c_1, ..., c_m\}$ denotes the set of schools. Schools' capacities are given by $Q = \{q_{c_1}, ..., q_{c_m}\}$. Each student has a strict preference over the set of schools and being unmatched, denoted by $\succ_I = (\succ_{i_1}, ..., \succ_{i_n})$.[3] Let $\succeq_i$ denote the "at least as good as" relation induced by $\succ$. Σ denotes the set of all strict student preferences. Each student has score $s_c(i) \in [0, 1]$ in school c, where $s_c(i) \neq s_c(j)$ for all i, j and c. The scores of the lowest and highest scoring

[3] Being unmatched is denoted by i_k for student i_k.

students in each school are normalized to 0 and 1, respectively. Strict school priorities $\succ_C = (\succ_{c_1}, ..., \succ_{c_m})$ are derived from scores by ranking students with respect to their scores. A market is a tuple $\omega = \{I, C, Q, \succ_I, \succ_C\}$.

A *matching* is a function $\mu : I \cup C \to 2^I \cup C$, where $\mu(i) \in C \cup \{i\}$ for all $i \in I$, $\mu(c) \subseteq I$, $|\mu(c)| \leq q_c$, and $\mu(i) = c$ if and only if $i \in \mu(c)$. $\mathcal{U}$ is the set of all matchings. A matching μ is *blocked* by student i and school c if i prefers c to $\mu(i)$ and either c prefers i to some $i' \in \mu(c)$ or $|\mu(c)| < q_c$. A matching μ is individually rational if $\mu(i) \succeq_i i$ for all i. A matching μ is *stable* if it is individually rational and is not blocked.

A mechanism $\phi : \Sigma \to \mathcal{U}$ produces a matching given preference reports from the students. $\phi(\sigma) = \mu_\phi(\sigma, \cdot)$ for all $\sigma \in \Sigma$ where $\mu_\phi(\sigma, i)$ and $\mu_\phi(\sigma, c)$ denote the school and the set of students i and c are matched to under ϕ. Next, I describe the Deferred Acceptance and Boston Mechanisms.

The Deferred Acceptance Mechanism (DA)

Step 1: Students apply to their first choice school. Schools reject the lowest-ranking students in excess of their capacity. All other offers are held temporarily.

Step t: If a student is rejected in Step $t-1$, he applies to the next school on his rank-order list. If all remaining schools are below the outside option, he applies nowhere. Schools consider both new offers and the offers held from previous rounds. They reject the lowest ranked students in excess of their capacity. All other offers are held temporarily.

Stop: The algorithm stops when no rejections are issued. Each school is matched to the students it is holding at the end.

The Boston Mechanism (BM)

Step 1: Students apply to their first choice school. Schools reject the lowest-ranking students in excess of their capacity. All other offers are immediately accepted and become permanent matches. School capacities are adjusted accordingly.

Step t: If a student is rejected in Step $t-1$, he applies to the next school on his rank-order list. If all remaining schools are below the outside option, he applies nowhere. Schools reject the lowest ranked students in excess of their capacity. All other offers become permanent matches. School capacities are adjusted accordingly.

Stop: The algorithm stops when no rejections are issued.

It is well known that DA is strategy proof (Dubins and Freedman, 1981; Roth, 1982), and its outcome is the student optimal stable matching (SOSM), which is individually rational and stable, but is not efficient for students. BM, on the other hand, is neither strategy proof nor stable, but is efficient with respect to the submitted preferences.

To make the connection between DA and RBM clear, I consider the following slightly modified version of DA, previously defined by Zhang (2021).

> **The Modified Deferred Acceptance Mechanism (MDA)**
> **Step 1**: Students apply to their first choice school. Schools reject the lowest-ranking students in excess of their capacity. All other offers are held temporarily.
> **Step t**: If a student is rejected in Step $t-1$, he *applies to the highest school on his rank-order list within the schools that either (i) did not fill its capacity in the last round or (ii) temporarily hold the offer of a student who has lower ranking in that school.* If there are no schools in his list that satisfies either (i) or (ii), or that school is ranked below the outside option, he applies nowhere. Schools consider both new offers and the offers held from previous rounds. They reject the lowest ranked students in excess of their capacity. All other offers are held temporarily.
> **Stop**: The algorithm stops when no rejections are issued. Each school is matched to the students it is holding at the end.

Under MDA, students skip applications to schools that already have enough applicants to fill their capacity with more preferred students. MDA is useful in understanding the relationship between DA, the student optimal stable matching and BM under repeated play. The following lemma shows that this mechanism is equivalent to DA.

Lemma 1. *Both DA and MDA terminate in finite time and yield the student optimal stable matching.*

3.3 Best Response Dynamics under Boston Mechanism

I now describe the setting for the repeated application of the Boston Mechanism. Time is discrete and denoted by $t \in \mathbb{N}$. These periods correspond to the consecutive years where students are matched to schools. In each period, each student in I

submits a ranking of schools in C and being unmatched. A strategy $\sigma \in \Sigma$ is a best response to σ_{-i} under mechanism ϕ if

$$\mu_\phi(\sigma, \sigma_{-i}, i) \quad \succeq_i \quad \mu_\phi(\sigma', \sigma_{-i}, i) \text{ for all } \sigma' \in \Sigma \tag{31}$$

While modelling the behavior of the students, it is not very realistic to consider a setting where students best reply to the strategy used by other students in the previous period for a couple of reasons. First, the strategy of all students is a high dimensional and complicated object. Therefore, it is not feasible the school district to report this statistic and students to process it to calculate their best responses. Second, disclosing the rank-order lists of the students in previous years would be challenging for the school board from a legal perspective. However, the cutoff scores of each school, which correspond to the score of the lowest scoring student who is assigned to a school, is a much simpler object which is reported in various settings.

I use $S^c \in [0,1]$ to denote the cutoff in school c, and $S = \{S^1, \ldots, S^n\}$. Let $S_t^c(\sigma)$ denote the score of the lowest scoring student who is assigned to school c in the first step of BM in period t if the school fills its capacity in that step (set $S_t^c(\sigma) = 0$ otherwise). Moreover, let $S_t = \{S_t^{c_1}, \ldots, S_t^{c_m}\}$. A strategy σ is *compatible* with a vector of cutoffs S if the cutoff scores in the first step of BM under σ is S, that is $S_t(\sigma) = S$. Let $\mathcal{S}$ denote set of all cutoffs that are compatible with at least one $\sigma \in \Sigma$. These are the cutoffs that might arise in the first step of the implementation of BM. A school c is *achievable for i at cutoffs S* if $s_c(i) \geq S^c$. Moreover, let $FC_i(S)$ denote the set of all strategies that (i) ranks the most preferred school that is achievable at S as the first choice (if there is no achievable school, then students rank their most preferred school first), (ii) ranks all schools preferred to being unmatched above being unmatched and (iii) ranks all schools less preferred to being unmatched below being unmatched. For the rest of the paper, I assume that ϕ is the Boston Mechanism and suppress dependence on ϕ.

Lemma 2. *If $\sigma_i \in FC_i(S)$, then it is a best response to all σ' such that σ' is compatible with S.*

This lemma shows that students do not need to observe σ_{-i} to determine their best response. Rather, observing S and determining the most preferred school that is achievable under S, a simple exercise, is sufficient for a student to compute a best response.[4] I study the following best response process, which I call *Repeated Boston Mechanism* (RBM):

[4]Some other strategies that does not rank the most preferred achievable school first might be best

- In period 1, all students apply truthfully, $\sigma_i^1 = \succ_i$. BM is used to determine the allocation.

- In period t, all students choose a strategy $\sigma_i^t \in FC(S_{t-1})$. BM is used to determine the allocation.

Let T denote the step where MDA terminates. We have the following result.

Proposition 11. *The set of students who are accepted by school c in the first step of the Boston Mechanism in period t of RBM is identical to the set of students who are tentatively accepted by school c in step t of the MDA for all $t \leq T$.*

This proposition shows that there is a close connection between periods of RBM, where students myopically best respond using information on the cutoffs from the previous period, and the steps in MDA. To get intuition about the result, suppose that the result holds for some t. If student i is tentatively accepted to school c at step t of MDA or applies to c in step $t+1$ of MDA, then that student has already either been rejected by or skipped more preferred schools in previous rounds of MDA. Given that the result holds for period t, in both cases, we know (i) the period t first step cutoffs of all schools that i prefers to c are higher than i's score at those schools and (ii) the period t first step cutoff of school c is lower than i's score at c. Therefore, i applies to c in step 1 of period $t+1$ of RBM. Thus, the set of applicants for each school is identical in the first step of BM in period $t+1$ of RBM and step $t+1$ of MDA, which means that the set of students who are accepted in the former and tentatively accepted in the latter are the same. Moreover, Proposition 11 implies that for in all periods after T, the BM terminates at step 1 and the outcome is SOSM:

Theorem 7. *The matching implemented in RBM converges in finitely many periods to the student optimal stable matching.*

As BM is not strategy-proof, to make any predictions about its outcome, one needs to study the equilibria of a preference revalation game where students submit their preferences to the mechanism. Ergin and Sönmez (2006) show that the set of Nash equilibria under BM correspond to the set of stable matchings and interpret their result as evidence in favor of DA. First, the Pareto-dominant equilibrium of BM is SOSM, which is attained under the DA when students report their preferences truthfully. Therefore, switching from the BM to DA cannot harm any student, but

responses. However this is only possible if all schools ranked above the the most preferred achievable school are not achievable and the most preferred achievable school has a cutoff of 0 under $\mathcal{S}$.

can potentially improve the outcomes of some. Second, as being truthful is a dominant strategy under DA, DA does not require the students to strategic about their applications.

Theorem 7 complements Ergin and Sönmez (2006) by providing a foundation to their analysis based on best response dynamics by showing that if BM is repeatedly used in the same market, and students form their strategies by myopically best responding previous period's cutoffs, the implemented matching converges to a stable matching in finitely many periods. Therefore, the set of equilibria characterized in Ergin and Sönmez (2006) can be reached under best response dynamics with minimal information about play in the previous period. Moreover, it selects SOSM as the equilibrium. Even though this result indicates the BM would converge to its Pareto-dominant equilibrium, DA still has a couple of advantages over BM. First, the convergence to SOSM is not immediate, while under DA it is reached in the first period. Moreover, even after the convergence is reached, students still need to be strategic about their applications in each period, while under DA they can rank the schools truthfully.

The rest of the paper studies extensions of the model to analyze convergence of RBM under different conditions.

Unsophisticated Students. Under RBM, all students play a best response to the previous period's cutoffs. An important question is the following: what happens if some subset of students are not able to play strategically, but report their preferences truthfully? Pathak and Sönmez (2008) develop a framework to allow for unsophisticated students who report their preferences truthfully, extending the analysis of Ergin and Sönmez (2006), while Song, Tomoeda, and Xia (2020) empirically demonstrate that a non-trivial fraction of students behave this way under BM. In Section 3.4, I extend the analysis in this section by allowing some students to be unsophisticated and show that RBM converges to the the Pareto-dominant equilibrium of BM that Pathak and Sönmez (2008) characterize, providing foundation for their equilibrium characterization as well as their focus on the Pareto-dominant equilibrium when analyzing preferences of sophisticated students over these two mechanisms.

First Period Behavior and Optimism. Under RBM, all students are truthful in the first period. However, it still possible students are actually strategic in the initial period, depending on their beliefs about their admission chances in different schools. In particular, if students are best responding to some initial cutoffs, truthful revelation of preferences correspond to a setting where they believe all schools are achievable, in other words, these students are optimistic about their admission chances. In Section 3.5, I relax the assumption of truthful revelation in the first period and show that if students are optimistic enough, then RBM converges to SOSM, while if students are pessimistic (in particular, if they believe their outcome under SOSM is not achievable), the implemented matching may not converge but instead cycle between matchings that are Pareto-dominated by SOSM.

Large Random Markets. Under RBM, in each period, the preferences and scores of students are the same, which is a reasonable assumption for large student assignment markets. In Section 3.6, I first study a continuum matching model (Abdulkadiroğlu, Che, and Yasuda, 2015; Azevedo and Leshno, 2016) and extend the results to that setting. Next, I study a setting where n students are sampled independently from a given continuum market each period and prove the convergence of RBM is obtained asymptotically (as $n \to \infty$), showing that if a market is large and stationary and students best respond to the previous period's cutoffs, then the implemented matching converges to the SOSM.

3.4 Sophisticated and Unsophisticated Students

The set of students is given by $I = I_S \cup I_U$. If $i \in I_s$, then in each period, i behaves strategically as they did under RBM, while if $i \in I_U$, then $\sigma_i^t = \succ_i$, that is, students in I_U apply truthfully in every period. Following Pathak and Sönmez (2008), we refer to these students as sophisticated and unsophisticated students, respectively.

Given a market $\omega = \{I_S, I_U, C, Q, \succ_I, \succ_C\}$, construct the augmented preferences $\tilde{\succ}_C$ where each school c ranks students as follows

> Rank all sophisticated students and unsophisticated students who ranks c first according to $\succ_c$
> Rank all unsophisticated students who ranks c second according to $\succ_c$
> $\vdots$
> Rank all unsophisticated students who ranks c last according to $\succ_c$

Under $\tilde{\succ}_C$, unsophisticated students lose their priorities to sophisticated students in all schools apart from the one they rank first. Pathak and Sönmez (2008) define the *augmented market*, $\tilde{\omega} = \{I, C, Q, \succ_I, \tilde{\succ}_C\}$ and show that (i) the Nash equilibria of BM where unsophisticated students mechanically submit their true preferences correspond to the set of stable matchings of the augmented market $\tilde{\omega}$ and (ii) the Pareto-dominant Nash equilibrium of this game corresponds to the SOSM of the augmented market $\tilde{\omega}$ and (iii) unsophisticated students become better off if they become sophisticated.

The following proposition shows that RBM with unsophisticated students converges to the Pareto-dominant equilibrium of BM.

Proposition 12. *The matching implemented in RBM with unsophisticated students converges in finitely many periods to the student optimal stable matching under the augmented market* $\tilde{\omega}$.

Proposition 12 provides a foundation for the equilibria characterized in Pathak and Sönmez (2008) based on best response dynamics. Moreover, Pathak and Sönmez (2008) also show that sophisticated students prefer the Pareto-dominant equilibrium of BM to the dominant strategy outcome under DA. This indicates that BM favors sophisticated parents if the Pareto-dominant Nash equilibrium is played. Their result explains why some parents were in favor of BM and provides formal support for Boston Public School's position on changing their student assignment system to level the playing field for students do not have resources to be strategic about their applications.[5] Proposition 12 complements their result by showing that the Pareto-dominant equilibrium would be obtained under best response dynamics, providing further justification for focusing on the Pareto-dominant equilibrium instead of other equilibria, which may not be preferred by the sophisticated students.

3.5 Initial Conditions and First Period Play

In this section, I study how the outcome of RBM depends on the behavior in the initial period, which turns out to be an important determinant of the convergence of RBM. Let S_0 denote the initial cutoffs to which students best reply in the first period of RBM. If $S_0 = \{0, \ldots, 0\}$, then students are optimistic in the sense that they believe all schools are achievable and in the first period and they rank first their most preferred school, in other words, they are truthful. If initial cutoffs are higher, then some schools become unachievable for some students, and students become more pessimistic about their admission chances in the initial period. In particular, if the initial cutoffs are above the DA cutoffs, then there are students who are pessimistic enough that they believe their match under SOSM is not achievable. The following example shows that if the initial cutoffs are above the DA cutoffs (denoted by S_{DA}),

[5] See Pathak and Sönmez (2008) for a detailed discussion of this topic. In particular, they note that the BPS Strategic Planning Team recommended the implementation of a strategy-proof algorithm by saying:

> A strategy-proof algorithm "levels the playing field" by diminishing the harm done to parents who do not strategize or do not strategize well.

then they may stay above the DA cutoffs in each period, which implies that (i) the cutoffs and the matchings do not converge and (ii) the realized matchings are Pareto dominated by SOSM.

Example 6. *There are three schools with unit capacity and three students. The priorities and preferences of students are as follows.*

Students	s_{c_1}	s_{c_2}	s_{c_3}	$\succ$
i_1	3	1	2	$c_2 \succ c_3 \succ c_1$
i_2	2	3	1	$c_3 \succ c_1 \succ c_2$
i_3	1	2	3	$c_1 \succ c_2 \succ c_3$

In the first step of DA and MDA, all students apply to their most preferred school and are accepted. Thus both mechanisms terminate in the first period and the cutoffs are $S_{DA} = \{1,1,1\}$. However, the matchings where all students obtain their second choice and their third choice are also stable, with cutoffs $\{2,2,2\}$ and $\{3,3,3\}$ respectively. Let $S_0 = \{2,1,1\}$. The following table shows the applications and cutoffs in the first step of first 4 periods of RBM, as well as the implemented matching in each period.

	Applications			*cutoffs*			*Matching*		
	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
P1		$i_1, \underline{i_3}$	i_2	1	2	1	i_1	i_3	i_2
P2	i_3		$\underline{i_1}, i_2$	1	1	2	i_3	i_2	i_1
P3	$\underline{i_2}, i_3$	i_1		2	1	1	i_2	i_1	i_3
P4		$i_1, \underline{i_3}$	i_2	1	2	1	i_1	i_3	i_2

In the first step of the BM in the first period of RBM, i_1 and i_3 apply to c_2 and i_2 apply to c_3. The cutoffs for next period then becomes $S_1 = \{1,2,1\}$ and i_3 applies to c_1 while i_1 and i_2 apply to c_3. This results in $S_2 = \{1,1,2\}$ and in the next period, i_1 applies to c_2 while i_2 and i_3 apply to c_1, which results in the cutoffs $S_3 = \{2,1,1\}$. Since $S_0 = S_3$, period 4 is identical to period 1, and the the matchings implemented under RBM cycles between matchings implemented in periods 1, 2 and 3 without converging to a matching. In each period, one student is assigned to her most preferred school, one student is assigned to her second most preferred school while one student is assigned to her least preferred school. This outcome is dominated by SOSM, which is attained under DA.

The close connection between MDA and RBM, and the fact that under MDA the cutoffs are increasing and converges to SOSM cutoffs suggests that under RBM, if initial cutoffs are below the DA cutoffs, then the convergence could be attained. The following example shows that convergence may not be attained even in that case.

Example 7. *There are two schools with unit capacity and five students. The preferences and scores of the students are as follows.*

Students	s_{c_1}	s_{c_2}	$\succ$
i_1	5	3	$c_2 \succ_i c_1$
i_2	3	5	$c_1 \succ_i c_2$
i_3	2	4	$c_1 \succ_i c_2$
i_4	0	1	$c_2 \succ_i c_1$
i_5	1	0	$c_2 \succ_i c_1$

The applications and admissions in each step of MDA, as well as the realized cutoffs at the end on each step are given in the following table:

	Applications		*Cutoffs*	
	c_1	c_2	c_1	c_2
S1	$\underline{i_2}, i_3, i_5$	$\underline{i_1}, i_4$	2	2
S2	$\underline{i_2}$	$i_1, \underline{i_3}$	2	3
S3	$\underline{i_1}, i_2$	$\underline{i_3}$	4	2
S4	$\underline{i_1}$	$\underline{i_2}, i_3$	4	4
S5	$\underline{i_1}$	$\underline{i_2}$	4	4

By Proposition 11, if $S_0 = (0, 0)$, then the students who are accepted by the schools in the step 1 of first 5 periods of RBM are given by the table above. Moreover, in later periods, the outcome is the same as the outcome of step 5. Suppose that $S_0 = (4, 2)$. RBM proceeds as follows:

	Applications		*Cutoffs*	
	c_1	c_2	c_1	c_2
P1S1	$\underline{i_5}$	$i_1, \underline{i_2}, i_3, i_4$	1	4
P2S1	$\underline{i_1}, i_2, i_3, i_5$	$\underline{i_4}$	4	1
P3S1	$\underline{i_5}$	$i_1, \underline{i_2}, i_3, i_4$	1	4
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Under initial cutoffs $(4,2)$*, most students think that* c_2 *is achievable, while* c_1 *is not. They apply to school* c_2*, increasing its cutoff score, while decreasing the cutoff of* c_1*. However, both schools are unachievable for* i_4 *and* i_5*, who apply to their most preferred school and are matched to that school in even and odd periods, respectively. Thus, the neither the cutoffs nor the matchings under RBM converges to the SOSM, but cycles among matchings that are not stable.*

In Example 7, the cutoffs start below S_{DA} and stay below S_{DA} in all periods but do not converge to S_{DA}. Next proposition shows this result is true in general.

Proposition 13. *If* $S_0 \leq S_{DA}$*, then* $S_t \leq S_{DA}$ *for all* t*. Therefore,* S_t *either converges to* S_{DA} *in finitely many periods, or cycles below* S_{DA}*.*

Thus, if students are optimistic enough that they believe their match under the student optimal stable matching is achievable, then they believe this would be the case in all future periods. Unlike Example 6, the reason of non-convergence is not that students do not apply to the school they will receive in the SOSM, thinking it is unachievable. The convergence does not happen because students may be optimistic in each period and believe that schools more preferred to their match in SOSM are achievable, and never apply to that school.

Given my result on possible non-convergence of RBM, the next question is the efficiency of the matchings that are implemented in a cycle. The matchings that constitute the cycle in Example 7 neither Pareto dominate, nor are Pareto dominated by SOSM. Without any restrictions on strategies, *i.e.*, if students can submit any ranking which is a best response to previous period's cutoffs, a matching implemented in a cycle of RBM can also be Pareto dominated by the SOSM, as students can be matched to suboptimal schools after the first step in each period. However, if students preserve the relative ranking of schools conditional on best replying, then this cannot happen. Formally, σ_i is an *order-preserving best reply* to cutoffs S if σ_i ranks the most preferred achievable school first and preserves the relative rankings of other schools (as well as the outside option) under $\succ_i$.

Proposition 14. *If* $S_0 \leq S_{DA}$ *and students' strategies are order-preserving best replies, then any matching implemented in an RBM cycle cannot be Pareto dominated by the SOSM.*

The reason behind this result is that when students use order-preserving best replies, if a student is matched to a school worse than their match under SOSM, then there must be at least one student who has become strictly better off compared to

SOSM. I now present an example where all matchings in a cycle Pareto dominate SOSM.

Example 8. *There are four schools and five students. The preferences and scores of the students are as follows.*

Students	s_{c_1}	s_{c_2}	s_{c_3}	s_{c_4}	$\succ$	*SOSM*
i_1	1	2	3	0	$c_1 \succ_i c_3 \succ_i c_2$	c_2
i_2	3	1	0	0	$c_2 \succ_i c_1$	c_1
i_3	0	0	2	3	$c_3 \succ_i c_4$	c_4
i_4	0	0	4	2	$c_4 \succ_i c_3$	c_3
i_5	2	0	1	1	$c_3 \succ_i c_4 \succ_i c_1$	i_5

SOSM is inefficient, as i_1 and i_2 as well as i_3 and i_4 can exchange their schools and become better off. However, these exchanges are blocked by i_5 and i_1, respectively. Suppose that $S_0 = (1, 1, 4, 1)$. RBM proceeds as follows:

	Applications				*Cutoffs*				*Matching*			
	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4	c_1	c_2	c_3	c_4
P1S1	$\underline{i_1}$	$\underline{i_2}$	$\emptyset$	$\underline{i_3}, i_4, i_5$	1	1	1	4	i_1	i_2	i_4	i_3
P2S1	$\underline{i_1}$	$\underline{i_2}$	$i_3, \underline{i_4}, i_5$	$\emptyset$	1	1	4	1	i_1	i_2	i_4	i_3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

As the cutoffs at the end of period 2 are identical to the initial cutoffs, $S_t = (1, 1, 4, 1)$ for odd t and $S_t = (1, 1, 1, 4)$ for even t, while the same matching is implemented in each period. In this matching, students i_1 and i_2 are strictly better off compared to SOSM, while all other students get the same outcome.[6] *Thus, in each period of RBM, the implemented matching Pareto dominates the student optimal stable matching.*

The examples and propositions in this section show how convergence may not be attained and how students may be worse or better off under BM. I now turn to the following question: when does RBM converge to the student optimal stable matching? It turns out that MDA is useful in understanding when convergence happens. Let $\tilde{S}_t^c$ denote the score of the lowest scoring student who is matched to school c in period t of MDA.[7] As at each step of MDA, the students who are tentatively admitted

[6] This more efficient matching is not stable, as i_5 would block it. However, i_5 applies to either c_3 or c_4 in each period, and does not initiate a rejection cycle between i_1, i_2 and i_5.

[7] I set $\tilde{S}_0^c = 0$ as there are no student matched to any school at the start of the algorithm and $\tilde{S}_t^c = 0$ if c has empty seats in step t.

to any school apply to that school, the cutoffs under MDA, $\tilde{S}_t^c$, are increasing in t. Intuitively, as the algorithm proceeds, each school replaces lower scoring tentatively admitted students with higher scoring ones and students move towards less preferred schools. Therefore, $\tilde{S}_t^c$ can be interpreted as an index of optimism regarding a student's admission chances at school c. A cutoff S_0 is *compatible* with MDA if there exists k such that for all c, $S_0^c \in [\tilde{S}_k^c, S_{k+1}^c]$. This means that the initial cutoffs of all schools are between the cutoffs in steps k and $k+1$ of MDA. The following proposition extends the close connection between MDA and RBM to this setting.

Proposition 15. *If S_0 is compatible with MDA, RBM converges to student optimal stable matching in finitely many periods.*

To prove this result, I first show that if S_t is compatible with round k cutoffs of MDA, then S_{t+1} is compatible with round $k+1$ cutoffs of MDA (Lemma 18). To see why, first note that when the cutoffs of other schools are higher, more students (in set inclusion sense) demand school c and next period cutoff of school c increase. Moreover, when other schools cutoffs are exactly $\tilde{S}_k^{-c}$ (regardless of school c's cutoff, as long as it is compatible), next period cutoff of c under RBM is $\tilde{S}_{k+1}^c$. Thus, whenever S_t is between $\tilde{S}_k$ and $\tilde{S}_{k+1}$, then next period cutoffs under RBM is between $\tilde{S}_{k+1}$ and $\tilde{S}_{k+2}$. As the cutoffs under MDA converges to S_{DA} in finitely many rounds, so does the cutoffs under RBM.

The main take-away from Proposition 15 is that, if students' levels of optimism about their admissions chances at different schools are similar in the sense that $S_0^c \in [\tilde{S}_k^c, \tilde{S}_{k+1}^c]$ for all $c \in \mathcal{C}$ for some k, then RBM converges to SOSM. Given S_0, the students are *optimistic enough* if $S_0 \leq \tilde{S}_1$. As this implies compatibility, we have the following theorem:

Theorem 8. *If students are optimistic enough, then RBM converges to the student optimal stable matching in finitely many periods.*

Theorem 8 reinterprets truthfulness in the first period of RBM as optimism regarding admissions chances and further emphasizes the importance of beliefs of students over the behavior of best response dynamics under BM.

Finally, I analyze a special case of the model to understand the reasons behind non-convergence. Given $\{s_c\}_{c \in \mathcal{C}}$, *scores are common across schools* if $s(i) \equiv s_c(i) = s_{c'}(i)$ for all i, c and c'.[8] It is well known that when scores are common across schools, DA

[8]Examples where this property holds include Taiwanese high school admissions (Dur, Pathak, Song, and Sönmez, 2022), Chicago Public schools (Dur, Pathak, and Sönmez, 2020) and Turkish high school admissions.

is equivalent to the Serial Dictatorship Mechanism and both mechanisms return the unique stable matching of the market.

Proposition 16. *When scores are common across schools, then the cutoffs and the allocation under RBM converges to S_{DA} and the unique stable allocation in finitely many periods.*

This proposition shows that the non-convergence is due to heterogeneity of scores between schools. This heterogeneity can create cycles where students remain pessimistic and do not apply to the schools that are attainable in the stable matching (Example 6) or remain optimistic and they keep applying to schools that are not attainable in the stable matching (Example 7). The proof of the proposition shows that once a student gets their stable matching in a period, they will be matched to that school in every other period and in each period, at least one additional student will be matched to their school under the stable matching. Therefore, when the scores are common across schools, such cycles cannot exist.

Sections 3.4 and 3.5 highlight the limitations of Theorem 7. First, if some students do not behave strategically, then RBM does not converge to SOSM. Second, students' initial beliefs are important and RBM may not converge to SOSM if they are initially pessimistic. Moreover, even if convergence eventually happens, it is not immediate and may take time. These results are in line with the empirical observations of non-equilibrium behavior in BM such as truthful reporting of preferences and overcautious strategies (Song, Tomoeda, and Xia, 2020).

3.6 Large Random Markets

Under RBM, the same market is repeated in every period. In this section, I analyze the robustness of the results to this assumption by building a large market model. First, I consider a continuum matching model based on Abdulkadiroğlu, Che, and Yasuda (2015) and Azevedo and Leshno (2016). Second, I consider the setting where in each period, n students from the continuum market are randomly drawn and matched using BM. I show that the results about convergence of RBM to SOSM continue to hold in the first case and hold asymptotically (as $n \to \infty$) in the second case.

3.6.1 Continuum Markets

There are a finite set of schools, denoted by $\mathcal{C} = \{c_0, c_1, \ldots, c_n\}$ and a unit measure of students, where c_0 is a dummy school which denotes being unmatched. Let $\theta =$

$(\succ_\theta, \{s_c^\theta\}_{c\in\mathcal{C}})$ denote the type of a student whose preferences over the set of schools is $\succ_\theta$ and has score $s_c^\theta \in [0,1]$ in school c. The set of student types is denoted by Θ, over which there is a probability measure η, which admits a full support density.[9] $Q = (q_0, q_1 \dots, q_n)$ denotes the capacities of schools. There are no capacity constraints for being unmatched, that is, $q_0 \geq 1$. A matching in this environment is a function $\mu : \mathcal{C} \cup \Theta \to 2^\Theta \cup \mathcal{C}$ where $\mu(\theta) \in \mathcal{C}$ is the school any type θ is assigned and $\mu(c) \subseteq \Theta$ is the set of students assigned to authority c such that no school is assigned to a measure of students larger than its capacity.[10]

A student-school pair (θ, c) *blocks* a matching μ if the student prefers c to its match and either school c does not fill its quota or school c is matched to another student who has a strictly lower score than θ. Formally, (θ, c) blocks μ if $c \succ_\theta \mu(\theta)$ and either $\eta(\mu(c)) < q_c$ or there exists $\theta' \in \mu(c)$ with $s_c^{\theta'} < s_c^\theta$. A matching μ is *stable* if it is not blocked by any student-school pair.

Stable matchings can be represented by cutoffs for n non-dummy schools.[11] For each $S \in [0,1]^n$, $\tilde{D}_\theta(S)$ denotes the *demand* of student θ at S, which is the most preferred school such that θ clears its threshold. $\tilde{D}_c(S)$ denotes the set of students who demand c. Given S, define the assignment $\mathcal{M}(S) = \nu$ induced by S as $\nu(\theta) = \tilde{D}_\theta(S)$. An assignment ν is a matching if $\eta(\tilde{D}_c(S)) \leq q_c$ for all c.

The definitions of the mechanisms are almost identical previous definitions, with the appropriate changes to adapt them to the continuum model. The formal definitions are provided in Appendix C.2. As η has full support, there is a unique stable matching in this market (Theorem 1 in Azevedo and Leshno (2016)). Moreover, this matching can be represented by a set of cutoffs (S_{DA}) where each student θ is matched to $\tilde{D}_\theta(S_{DA})$. As in Section 3.2, $\tilde{S}_t^c$ denotes the step t cutoffs in the MDA mechanism. The following proposition shows that MDA implements the stable matching.

Proposition 17. *$\tilde{S}_t$ converges to S_{DA}. The offers each school hold in step t of MDA converge to the unique stable matching as $t \to \infty$.*

The main difference of this result from the convergence results in finite markets is the fact that convergence may not happen in finitely many rounds when there is a continuum of students. Although the measure of students who are rejected in each

[9] Formally, the score distribution of students with each preference profile $\succ$ has a full support density.

[10] The mathematical definition of a matching for the continuum market we study follows Azevedo and Leshno (2016) and requires that μ satisfies the following four properties: (i) for all $\theta \in \Theta$, $\mu(\theta) \in \mathcal{C}$; (ii) for all $c \in \mathcal{C}$, $\mu(c) \subseteq \Theta$ is measurable and $\eta(\mu(c)) \leq q_c$; (iii) $c = \mu(\theta)$ iff $\theta \in \mu(c)$; (iv) (open on the right) for any $c \in \mathcal{C}$, the set $\theta \in \Theta : c \succ_\theta \mu(\theta)$ is open.

[11] The cutoff for c_0 is always 0 as the outside option of being unmatched is always available.

step of MDA (and also, DA) converges to 0, it is possible that a positive measure of students are rejected in each step.[12]

Let R_0 denote the initial cutoffs that students best respond in the initial period of RBM, while R_t denote the step 1 cutoffs in the period t of RBM. We can extend Proposition 15 and Theorem 8 to the continuum setting.

Proposition 18. *If R_0 is compatible with MDA, then RBM converges to SOSM.*

Theorem 9. *If students are optimistic enough at R_0, then allocation under RBM converges to the unique stable matching.*

These results serve two purposes. First, they show that the convergence properties of RBM are also true in continuum markets. Second, and more importantly, they are useful in proving the convergence in finite large markets drawn from a stationary distribution, showing the assumption that the repetition of the same market is not necessary in large markets.

3.6.2 Large Random Markets

To extend the results to large finite markets sampled from a continuum market, I first define finite markets in the continuum setting following Azevedo and Leshno (2016). A finite market $F = [\tilde{\Theta}, \tilde{Q}]$ specifies a finite set of students $\tilde{\Theta} \subset \Theta$ and an integer vector of capacities $\tilde{q}_c > 0$, where $\tilde{q}_{c_0} \geq |\tilde{\Theta}|$. A matching for a finite market is a function $\tilde{\mu} : \mathcal{C} \cup \Theta \to 2^{\Theta} \cup \mathcal{C}$ such that (i) for all $\theta \in \tilde{\Theta}$, $\tilde{\mu}(\theta) \in \mathcal{C}$, (ii) for all $c \in \mathcal{C}$, $\tilde{\mu}(c) \in 2^{\tilde{\Theta}}$ and $|\tilde{\mu}(c)| \geq \tilde{Q}_c$ and (iii) for all $\theta \in \tilde{\Theta}$ and $c \in \mathcal{C}$, $\tilde{\mu}(\theta) = c$ iff $\theta \in \tilde{\mu}(c)$. The definition of blocking pairs, as well as the definition of stability is the same as in Section 3.6.1. A finite market $F = [\tilde{\Theta}, \tilde{Q}]$ is associated with the following empirical distribution of types

$$\eta = \sum_{\theta \in \tilde{\Theta}} \frac{1}{|\tilde{\Theta}|} \delta_\theta$$

where δ_θ denotes the probability distribution that places probability one on the point θ. The supply of seats per student is given by $Q = \tilde{Q}/|\tilde{\Theta}|$. Either $[\tilde{\Theta}, \tilde{Q}]$ or $[\eta, Q]$ uniquely determine a discrete market F.[13] Fix a continuum market with full support, (η, Q). (η, Q) has a unique stable matching μ with cutoffs $S_{DA} \in [0, 1]^{\mathcal{C}}$. To remove any confusion, I use $\tilde{R}_t$ to denote the first step cutoffs of BM in period t. $\tilde{R}_0$ denotes the initial cutoffs.

I study the following repeated implementation of BM where k students are drawn from the continuum market (η, Q):

[12]See Azevedo and Leshno (2016) for an example.

[13]To see why, note that $\tilde{\Theta} = \text{support}(\eta)$ and $\tilde{Q} = Q|\tilde{\Theta}|$.

- In period 1, k students are independently and randomly drawn from η and the vector of capacities is $\tilde{Q} = Qk$. All students choose a strategy that is a best response to the $\tilde{R}_0$. BM is used to determine the allocation and $\tilde{R}_1$ denotes the first step cutoffs in the BM.

- In period t, k students are independently and randomly drawn from η and the vector of capacities is $\tilde{Q} = Qk$. All students choose a strategy that is a best response to the $\tilde{R}_{t-1}$. The BM is used to determine the allocation and $\tilde{R}_t$ denotes the first step cutoffs in the BM.

For a given k, the period t cutoffs of RBM, $\tilde{R}_t(k)$ is a random variable distributed in $[0,1]^{\mathcal{C}}$ and its realization depends on the previous period's cutoffs as well as the random market drawn from the distribution. The following proposition shows that the behavior of RBM in random finite markets converges to its behavior in the continuum market the random markets are sampled from.

Proposition 19. *Suppose that $R_0 = \tilde{R}_0$. Then $\lim_{k\to\infty} \tilde{R}_t(k)$ converges in probability to R_t. If $\tilde{R}_0$ is compatible with MDA, then $\lim_{t\to\infty} \lim_{k\to\infty} \tilde{R}_t(k)$ converges in probability to S_{DA}.*

Given Proposition 19 and Theorem 9, we conclude that RBM converges to the unique stable matching in large random markets.

Theorem 10. *If students are optimistic enough, then the matching implemented in period t of RBM converges to the unique SOSM of the continuum market as the market grows large and $t \to \infty$.*

This theorem shows that repetition of the same market is not necessary for the result. In large markets with a stationary distribution of students and scores across periods, the the cutoffs and the allocation in period t of RBM converges to the cutoffs and tentative allocation in step t of MDA, which converges to the unique stable matching of the continuum market. Therefore, it is reasonable to expect that if BM is used in a market and students get can obtain information about previous periods, then the outcome converges to the student optimal stable matching.

3.7 Conclusion

This paper studies best response dynamics under repeated application of the Boston Mechanism. When students best respond to the admission cutoffs in the previous period (with the exception of the initial period, where they are truthful), the implemented matching converges to the student optimal stable matching, which is the

Pareto-dominant equilibrium of the Boston Mechanism and dominant strategy outcome of the competing Deferred Acceptance Mechanism. This result provides a foundation for the equilibrium analysis of the Boston Mechanism based on best response dynamics. Extending the model to include unsophsiticated students who always apply truthfully and allow for student's initial beliefs to determine their first period behavior, I show how the student optimal stable matching may not be reached if some students cannot strategize or are initially pessimistic about their admissions chances.

Appendix A

Appendix to Diversity Preferences, Affirmative Action and Choice Rules

A.1 Omitted Proofs

A.1.1 Proof of Proposition 1

Follows from Proposition 2 if we define $\triangleright$ and $\trianglerighteq$ as follows: $I \trianglerighteq I'$ if and only if $\tau(I) = \tau(I')$ and $\triangleright$ is the empty relation.

A.1.2 Proof of Proposition 2

Define the relation $>_C$ as follows: $I >_C I'$ if $I \in C(\hat{I})$ and $I' \subset \hat{I}$. Let $>_C \cup \trianglerighteq$ denote the union of $>_C$ and $\trianglerighteq$ and $\text{tran}(>_C \cup \trianglerighteq)$ denote the transitive closure of this relation. We say that C satisfies $\triangleright$-congruence if (i) $I \ \text{tran}(>_C \cup \trianglerighteq) I'$ and $I' \in C(\hat{I})$ imply $I \in C(\hat{I})$ for every $\hat{I}$ that contains I and (ii) $I \ \text{tran}(>_C \cup \trianglerighteq) I'$ imply not $I' \triangleright I$. The following lemma follows from the finiteness of $\mathcal{I}$.

Lemma 3. *C satisfies $\triangleright$-congruence if and only if C does not admit a score-choice cycle.*

Proof. Assume that C admits a score-choice cycle $I_1, \ldots, I_n, I_1$. Then for each $i \leq n-1$, either $I_i >_C I_{i+1}$ or $I_i \trianglerighteq I_{i+1}$. Thus, $I_1 \ \text{tran}(>_C \cup \trianglerighteq) I_n$. Moreover, we also have either $I_n \triangleright I_1$ or $I_n \in C(\hat{I})$, $I_n \notin C(\hat{I})$ and $I_1 \subset \hat{I}$ for some $\hat{I}$, both of which cause failure of $\triangleright$-congruence.

Next, assume that C does not satisfy $\triangleright$-congruence and let I_1 and I_n denote the sets that cause the violation. Then $I_1 \ \text{tran}(>_C \cup \trianglerighteq) I_n$. As $\mathcal{I}$ is finite, all q element subsets of $\mathcal{I}$ is also finite, which means that there exists $I_1, I_2, \ldots, I_n$ such that for each $i < n$, either (i) there exists an $\hat{I}_i$ such that $I_i \in C(\hat{I}_i)$ and $I_{i+1} \subset \hat{I}_i$ or (ii) $I_i \trianglerighteq I_{i+1}$.

Moreover, we also have either either (i) there exists $\hat{I}_n$ such that $I_n \in C(\hat{I}_n)$, $I_1 \subset \hat{I}_n$ and $I_1 \notin C(\hat{I}_n)$ or (ii) $I_n \triangleright I_1$, which completes the proof. □

The result then follows from Theorem 7 in Nishimura, Ok, and Quah (2016).

A.1.3 Proof of Proposition 3

Let $\mathcal{I}$ denote set of individuals where there are q individuals from each $\theta \in \Theta$ and $\succeq$ denote the preferences that does not take intersectionality into account. For the rest of the proof, I will refer individuals who belong to group j in dimension k as (j,k) individuals. Formally, i is a (j,k) individual if $\theta_k(i) = j$.

Without loss of generality, let the first dimension to be one of the dimensions with fewest available groups, *i.e.*, $|\Theta_1| \leq |\Theta_l|$ for all l. Let D^* denote the set of all optimal marginal distributions. Formally, $d \in D^*$ if there exists I' such that $I' \in C(\mathcal{I})$ and $M(I') = d$. Let d_1 denote an element of D^* with highest number of $(1,1)$ individuals and let m_{11} denote the number of $(1,1)$ individuals at d_1. Let D_1^* denote the set of all optimal group distributions where the number of $(1,1)$ individuals is m_{11}. Next, let d_{11}^* be a group distribution in D_1^* with the highest number of $(2,1)$ individuals. m_{21} denotes the number of $(2,1)$ individuals at d_{11}^*.

We say that a set of individuals I is *compatible with* marginal distributions d^* if there exists I' such that $M(I \cup I') = d^*$. If $M(I \cup I') = d^*$, then I' is a *complement* of I for d^*. Let $M_{ij}(I)$ denote the number of group i individuals in dimension j in I.

Lemma 4. *Let d denote a marginal distribution and let $M_{ij}(d)$ denote the number of group i individuals in dimension j. If $M_{ij}(I) \leq M_{ij}(d)$ for all i and j, then I is compatible with d.*

Proof. Note that since d is a marginal distribution, $|d| \equiv \sum_i M_{ij}(d)$ for all j. Moreover, $\sum_i M_{ij}(I) = |I|$ for all j.

First, if $|I| = |d|$, then $M_{ij}(I) \leq M_{ij}(d)$ implies $M_{ij}(I) = \theta_{ij}(d)$ and I is compatible with d. If $|I| < |d|$, then for each dimension i, there exists a group j such that $M_{ij}(I) < M_{ij}(d)$. Let t denote an individual who belongs to group j that satisfies this condition. Then the set $I \cup \{t\}$ still satisfies $M_{ij}(I) \leq M_{ij}(d)$ and repeating this procedure yields a $\tilde{I}$ such that $M_{ij}(\tilde{I}) = M_{ij}(d)$. Letting $I' = \tilde{I} \setminus I$, we have $M(I \cup I') = d^*$ and therefore I is compatible with d. □

Claim 1. *$m_{11} < q$ and $m_{21} < q$.*

Proof. If either $m_{11} = q$ or $m_{21} = q$, then there exists $I' \in C(I)$ with no $(1,2)$ or $(2,2)$ individuals, which is a contradiction to interior optimality since $M(I')$ is at boundary. □

We are now ready prove the result. There are two cases, either $m_{11} \leq m_{21}$ or $m_{11} > m_{21}$.

Case 1: $m_{11} \leq m_{21}$.

Claim 2. *There exists* $I_{11} = \{i_1^{11}, \ldots, i_{m_{11}}^{11}\}$ *where all* $i \in I_{11}$ *are* $(1,1)$ *and* $(2,1)$ *individuals and* I_{11} *is compatible with* d_{11}^*.

Proof. First, note that since $m_{11} < q$ and groups $(1,1)$ and $(2,1)$ have (weakly) more individuals at d_{11}^*, one can choose the groups of individuals in I_{11} to satisfy $M_{ij}(I_{11}) \leq M_{ij}(g_{11}^*)$ for all i and j. Then the result follows from Lemma 4. □

Let I' denote a complement of I_{11} at d_{11}^* that includes an individual who is neither a $(1,1)$ nor $(2,1)$ individual.[1] Take $j \in I_{11}$ and $k \in I'$, where k is not a $(1,1)$ or $(2,1)$ individual. Define $\tilde{j}$ and $\tilde{k}$ as

$$\theta_1(\tilde{j}) = \theta_1(j), \theta_\ell(\tilde{i}) = \theta_\ell(k) \text{ for all } \ell \neq 1$$

$$\theta_1(\tilde{k}) = \theta_1(k), \theta_\ell(\tilde{k}) = \theta_\ell(j) \text{ for all } \ell \neq 1$$

Let $\tilde{I}_{11} = I_{11} \setminus \{j\} \cup \tilde{j}$ and $I'' = I' \setminus \{k\} \cup \tilde{k}$. Note that I'' is a complement of $\tilde{I}_{11}$ at d_{11}^*.

Claim 3. I' *and* I'' *does not have any* $(1,1)$ *individuals. Moreover,* I'' *has* $m_{21} - m_{11} + 1$ *group* $(2,1)$ *individuals.*

Proof. The first part of the result follows from the fact that at any optimal I cannot have more than m_{11} $(1,1)$ individuals and I' and I'' are complements of I_{11} and $\tilde{I}_{11}$ at d_{11}^*. Second part follows from the fact that I'' is a complement of $\tilde{I}_{11}$ at d_{11}^* and $\tilde{I}_{11}$ has $m_{11} - 1$ $(2,1)$ individiuals. □

Let $\bar{I} = I_{11} \cup I' \cup \tilde{I}_{11} \cup I''$. First, note that $I_{11} \cup I' \in C(\bar{I})$ and $\tilde{I}_{11} \cup I'' \in C(\bar{I})$, since $M(I_{11} \cup I') = d_{11}^*$ and $M(\tilde{I}_{11} \cup I'') = d_{11}^*$.

Lemma 5. *There does not exist an* $I^* \in C(\bar{I} \setminus I')$ *such that* $I_{11} \subset I^*$

Proof. For a contradiction, assume such an I^* exists. Then it must be that, $\tilde{I}_{11} \cap I^* = \emptyset$, since otherwise there will be more than m_{11} $(1,1)$ individuals at I^*, which will be a contradiction. However, this means that $I^* = I_{11} \cup I''$. But I^* has $m_{21} + 1$ $(2,1)$ individuals and m_{11} $(1,1)$ individuals, which contradicts the optimality of I^* as d_{11}^* is a group distribution in D_1^* with the highest number of $(2,1)$ individuals and has m_{21} such individuals. Since $\tilde{I}_{11} \cup I''$ is available and optimal, this is a contradiction. □

[1] This is possible since the preferences satisfy interior optimum and all individuals in I_{11} are both $(1,1)$ and $(2,1)$ indviduals.

The result then follows from the fact that I_{11} is chosen from $\bar{I}$, but not from $\bar{I} \setminus I'$.

Case 2: $m_{11} > m_{21}$ Let $n = m_{11} - m_{21}$.

Claim 4. *There exists* $I_{12} = \{i_1^{11}, \ldots, i_{m_{21}}^{11}, i_1^1, \ldots, i_n^1\}$ *where the first* m_{21} *elements are* $(1,1)$ *and* $(2,1)$ *individuals, rest are* $(1,1)$ *individuals and* I_{12} *is compatible with* d_{11}^*.

Proof. First, note that since $m_{21} < q$ and groups $(1,1)$ and $(2,1)$ have (weakly) more individuals at d_{11}^*, one can choose the groups of individuals in I_{12} to satisfy $M_{ij}(I_{12}) \leq M_{ij}(d_{11}^*)$ for all i and j. Then the result follows from Lemma 4. □

Let I' denote a complement of I_{12} at d_{11}^* that includes an individual who is neither a $(1,1)$ nor $(2,1)$ individual.[2] Take $j \in I_{12}$ and $k \in I'$, where k is not a $(1,1)$ or $(2,1)$ individual. Define $\tilde{j}$ and $\tilde{k}$ as

$$\theta_2(\tilde{j}) = \theta_2(j), \theta_\ell(\tilde{i}) = \theta_\ell(k) \text{ for all } \ell \neq 1$$

$$\theta_2(\tilde{k}) = \theta_2(k), \theta_\ell(\tilde{k}) = \theta_\ell(j) \text{ for all } \ell \neq 1$$

Let $\tilde{I}_{12} = I_{12} \setminus \{j\} \cup \tilde{j}$ and $I'' = I' \setminus \{k\} \cup \tilde{k}$. Note that I'' is a complement of $\tilde{I}_{12}$ at d_{11}^*.

Claim 5. I' *and* I'' *does not have any* $(2,1)$ *individuals. Moreover,* I'' *has* 1 *group* $(1,1)$ *individual.*

Proof. First, note that $I_{12} \cup I'$ and $\tilde{I}_{12} \cup I''$ have m_{11} group $(1,1)$ individuals. First part then follows since any optimal I that has m_{11} group $(1,1)$ individuals cannot have more than m_{21} $(2,1)$ individuals and I' and I'' are complements of I_{12} and $\tilde{I}_{12}$ at d_{11}^*. Second part follows from the fact that I'' is a complement of $\tilde{I}_{12}$ at d_{11}^* and $\tilde{I}_{12}$ has $m_{11} - 1$ $(1,1)$ individuals. □

Let $\bar{I} = I_{12} \cup I' \cup \tilde{I}_{12} \cup I''$. First, note that $I_{12} \cup I' \in C(\bar{I})$ and $\tilde{I}_{12} \cup I'' \in C(\bar{I})$, since $M(I_{12} \cup I') = d_{11}^*$ and $M(\tilde{I}_{12} \cup I'') = d_{11}^*$.

Lemma 6. *There does not exist an* $I^* \in C(\bar{I} \setminus I')$ *such that* $I_{12} \subset I^*$

Proof. First, note that since $\tilde{I}_{11} \cup I''$ is available and optimal, I^* must also be optimal. For a contradiction, assume such an I^* exists. Then it must be that, $\tilde{I}_{12} \cap I^* = \emptyset$.

To see why, assume there exist a $t \in \tilde{I}_{12} \cap I^*$. If $t = \tilde{k}$, then there are more than m_{21} group $(2,1)$ individuals and at least group m_{11} group $(1,1)$ individuals in I^*,

[2] This is possible since the preferences satisfy interior optimum and all individuals in I_{12} are both $(1,1)$ and $(2,1)$ indviduals.

which contradicts optimality of I^*. If $t \neq \tilde{k}$, then there are more than m_{11} group $(1,1)$ individuals in I^*, which contradicts optimality of I^*.

However, this means that $I^* = I_{11} \cup I''$. But then I^* has $m_{11}+1$ $(1,1)$ individuals, which contradicts the optimality of I^* as d^*_{11}. □

The result then follows from the fact that I_{11} is chosen from $\bar{I}$, but not from $\bar{I} \setminus I'$.

A.1.4 Proof of Proposition 4

The proof is follows from following the steps in the proof of Proposition 3 with minor modifications. Assume that all individuals have the highest scores, $\bar{s}$ and replicate the steps. For Case 1, consider I'_s where all individuals in I' have strictly lower scores than the highest score $\bar{s}$.

Let $\bar{I} = I_{11} \cup I' \cup \tilde{I}_{11} \cup I''$ and $\bar{I}_s = I_{11} \cup I'_s \cup \tilde{I}_{11} \cup I''$. First, note that $I_{11} \cup I' \in C(\bar{I})$ and $\tilde{I}_{11} \cup I'' \in C(\bar{I})$, since $M(I_{11} \cup I') = d^*_{11}$ and $M(\tilde{I}_{11} \cup I'') = d^*_{11}$.

Lemma 7. *There does not exist an $I^* \in C(\bar{I}_s)$ such that $I_{11} \subset I^*$.*

Proof. For a contradiction, assume such an I^* exists.

Claim 6. $I^* \cap I_s = \emptyset$.

Proof. For a contradiction, assume that $I^* \cap I_s \neq \emptyset$. Since $I^* \in C(\bar{I}_s)$, $I^* \succeq \tilde{I}_{11} \cup I''$. Define I^*_s by increasing the scores of all individuals in I^* to $\bar{s}$. From monotonicity, $I^*_s \succ I^* \succeq \tilde{I}_{11} \cup I''$, which is a contradiction as $\tilde{I}_{11} \cup I''$ is an optimal group distribution when all individuals have scores $\bar{s}$. □

Given Claim 6, following the same steps in Lemma 5 yields the result. □

The proof of Case 1 then follows from the fact that in $\bar{I}_s$ the scores of individuals in I_{11} are same as $\bar{I}$, scores of all other individuals are weakly lower than $\bar{I}$ and I_{11} is chosen from $\bar{I}$, but not from $\bar{I}_s$.

To prove the Case 2, Let $\bar{I} = I_{12} \cup I' \cup \tilde{I}_{12} \cup I''$ and $\bar{I}_s = I_{12} \cup I'_s \cup \tilde{I}_{12} \cup I''$. First, note that $I_{12} \cup I' \in C(\bar{I})$ and $\tilde{I}_{12} \cup I'' \in C(\bar{I})$, since $M(I_{12} \cup I') = d^*_{11}$ and $M(\tilde{I}_{12} \cup I'') = d^*_{11}$.

Lemma 8. *There does not exist an $I^* \in C(\bar{I}_s)$ such that $I_{12} \subset I^*$.*

Proof. For a contradiction, assume such an I^* exists.

Claim 7. $I^* \cap I_s = \emptyset$.

Proof. Follows from the same steps in Claim 6. □

Given Claim 7, following the same steps in Lemma 6 yields the result. □

The proof of Case 2 then follows from the fact that in $\bar{I}_s$ the scores of individuals in I_{12} are same as $\bar{I}$, scores of all other individuals are weakly lower than $\bar{I}$ and I_{12} is chosen from $\bar{I}$, but not from $\bar{I}_s$.

A.1.5 Proof of Proposition 5

Suppose that C satisfies gross substitutes* and acyclicty. First, I extend $>_C$ to include comparisons of individuals from same groups.

Lemma 9. *Suppose that gross substitutes and acyclicty are satisfied. Fix an s, θ and $n \geq 1$, suppose that for any $n' > n$, $>_C$ does not include*

$$(s, \theta, n') >_C (s, \theta, n'') \tag{32}$$

Define $\tilde{>}_C$ by adding $(s, \theta, n) \tilde{>}_C (s, \theta, n+1)$ to $>_C$. Then $\tilde{>}_C$ also satisfies acyclicity.

Proof. Assume for a contradiction $\tilde{>}_C$ does not satisfy acyclicity. This means that there exists a cycle under $\tilde{>}_C$ where $(s'_1, \ldots, s'_m)$ is a permutation of $(s_1, \ldots, s_m)$ and $(d'_1, \ldots, d'_m)$ is a permutation of $(d_1, \ldots, d_m)$.

Claim 8. *In this cycle, for some i, we have $(s_i, d_i) = (s, \theta, n)$ and $(s'_i, d'_i) = (s, \theta, n+1)$.*

Proof. Suppose that this is not the case. Then replacing $\tilde{>}_C$ with $>_C$ we still have a cycle as apart from $(s, \theta, n) \tilde{>}_C (s, \theta, n+1)$, the relations are the same. This contradicts that $>_C$ satisfies acyclicity and proves the result. □

Note that since $(s, \theta, n+1)$ is in the RHS of the cycle, there exists $(\hat{s}, \theta, n+1)$ in the LHS for some $\hat{s} \in \mathcal{S}$. Let $(\hat{s}, \theta, n+1) \tilde{>}_C (s', \theta', n')$ denote the element of the cycle that corresponds to $(\hat{s}, \theta, n+1)$. As $>_C$ does not include any relation characterized in Equation 32 for $n' > n$, $(\hat{s}, \theta, n+1) \tilde{>}_C (s', \theta', n')$ implies that $\theta' \neq \theta$, which in turn implies $(\hat{s}, \theta, n+1) >_C (s', \theta', n')$.

Claim 9. $(\hat{s}, \theta, n) >_C (s', \theta', n')$.

Proof. As $(\hat{s}, \theta, n+1) >_C (s', \theta', n')$, there exists I, j and k such that $s(j) = \hat{s}$, $\theta(j) = \theta$, $N_\theta(I \cup \{j\}) = n+1$, $s(k) = s'$, $\theta(k) = \theta'$, $N_{\theta'}(I \cup \{k\}) = n'$ and $k \in R(I \cup \{j, k\})$. Moreover, since $\theta(j) \neq \theta(k)$, we have that $(s(j), \theta(j), N_{\theta(j)}) >_C (s(k), \theta(k), N_{\theta(k)})$.

Now consider $\hat{I} = I \cup \{j, k, k'\}$ where $\theta(k) = \theta(k')$ and $s(k) = s(k')$. For a contradiction, assume that $j \in R(I \cup \{j, k, k'\})$. Let $I^* \in C(\hat{I})$ where $\{j, l\} \notin I^*$ for some other l. If $\theta(l) = \theta(k)$ and $s(l) = s(k)$, this means that $I^* = I \cup \{k\} \in C(\hat{I})$. However as I^* was available when the choice set was $I \cup \{j, k\}$, and $I \cup \{j\}$ was chosen, while I^* was not chosen, this creates a choice cycle, which is a contradiction. If $t(l) \neq t(k)$ or $s(l) \neq s(k)$, then $k \notin R(\hat{I})$. However, this violates gross substitutes as letting $I_{\theta(k)}$ denote all individuals in $\hat{I}$ with type $\theta(k)$, $I_{\theta(k)} \setminus \{k'\} \in C(\hat{I})$, but $I_{\theta(k)} \setminus \{k'\} \notin C(\hat{I} \setminus \{k'\})$. Therefore, $j \notin R(I \cup \{j, k, k'\})$

As $n \geq 1$, at I, there is another individual with type $\theta(j)$, which I denote by j'. Consider $(I \cup \{j, k, k'\} \setminus \{j'\}$ and let $I_{\theta(j)}$ denote all individuals in $I \cup \{j\} \setminus \{j'\}$. As $I_{\theta(j)}$ is chosen under $I \cup \{j, k, k'\}$, by gross substitutes*, it is also chosen from $I \cup \{j, k, k'\} \setminus \{j'\}$. This also implies that at least one of k and k' are rejected, which means that $(s(j), \theta(j), N_{\theta(j)} - 1) >_C (s(k), \theta(k), N_{\theta(k)})$, proving the claim. □

Now, we can replace $\hat{s}, \theta, n+1$ with $\hat{s}, \theta, n+1$ in the RHS and s, θ, n with $s, \theta, n+1$. Note that the line with $s, \theta, n > s, \theta, n+1$ is now $s, \theta, n+1 > s, \theta, n+1$. Removing this line, we still have a cycle on $\tilde{>}_C$, which does not use $s, \theta, n \tilde{>}_C s, \theta, n+1$, and therefore is also a cycle in $>_C$, which is a contradiction. This proves the result. □

Using Lemma 9 repeatedly, we arrive at an acyclic $\tilde{>}_C$ relation that satisfies $s, \theta, n > s, \theta, n+1$ for all $n \geq 1$, s and θ.

Lemma 10. *There exists u and h such that $s, d \tilde{>}_C s', d'$ implies $u(s) + h(d) > u(s') + h(d')$.*

Proof. Follows from Fishburn Theorem 4.1. □

Claim 10. *$h(\theta, n) > h(\theta, n+1)$ for all θ and n.*

Proof. Immediate from the construction of $\tilde{>}_C$, which includes $(s, \theta, n) >_C (s, \theta, n+1)$ for all n. □

Now, define $h_\theta(n) = \sum_{i=1}^{n} h(\theta, n)$. From the previous claim, we know that h_θ is concave. The following Lemma finishes the proof of the first part.

Lemma 11. *$U(I)$ where*

$$U(I) = \sum_{i \in I} u(s(i)) + \sum_{\theta \in T} h_\theta(N_\theta(I))$$

rationalizes C.

Proof. For a contradiction, assume it does not rationalize C. Then there exists I and I' such that $\tau(I) \neq \tau(I')$, $U(I) > U(I')$, $I' \in C(\hat{I})$ for some $\hat{I}$ that includes I. Moreover, we can take I to be a maximizer of $U(\tau(I)) = \max_{\tilde{I} \in 2^{\hat{I}}_q} U(\tau(\tilde{I}))$, which exists by the finiteness of $\hat{I}$.

First, if there exists $i \in I \setminus I'$ and $j \in I' \setminus I$ such that $t(i) = t(j)$. If this is the case, Let $\tilde{I} = I' \setminus \{j\} \cup \{i\}$ Clearly, this does not change the outcome of the choices or the utility difference.[3] We can repeat this until there does not exist any $i \in I \setminus I'$ and $j \in I' \setminus I$ such that $t(i) = t(j)$, and denote $\tilde{I}$ as this updated set. Now, choose an arbitrary $i \in \tilde{I} \setminus I$. Since C satisfies gross substitutes*, there exists I_C such that $I_C \in C(I \cup \{i\})$ and $i \in I_C$. Thus, there exists $j \in I$ such that $j \notin i_C$. As $t(i) \neq t(j)$, this shows that $s(i), \theta(i), N_{\theta(I)}(I_C) >_C s(i), \theta(i), N_{\theta(I)}(I_C) + 1$, which implies that

$$u(s(i)) + h(\theta(i), N_{\theta(i)}(I_C)) > u(s(j)) + h(\theta(j), N_{\theta(j)}(I_C) + 1)$$

However, above equation indicates $U(I \cup \{i\} \setminus \{j\}) > U(I)$, which is a contradiction as I maximizes utility in $\hat{I}$, which includes $I \cup \{i\}$. $\square$

To prove the second part, given $h_\theta(n)$, define $h(\theta, n) = h_\theta(n) - h_\theta(n-1)$. Assume for a contradiction there exists a cycle. This means that for each (s_i, d_i) and (s'_i, d'_i), $u(s_i) + h(d_i) > u(s'_i) + h(d'_i)$, which implies $\sum_i (s'_i, d'_i)$, $u(s_i) + h(d_i) > \sum_i u(s'_i) + h(d'_i)$, which is a contradiction (s'_i, d'_i) is a permutation of (s_i, d_i).

Next, for a contradiction, assume C_U violates gross substitutes*. Then there exists i, j, j' and a set of individuals I such that $s(j) > s(j')$, $\theta(j) = \theta(j')$, $i \in \hat{I}$ for some $\hat{I} \in C(I \cup \{i, j\})$ but either (i) there does not exist $\tilde{I} \in C(I \cup \{i, j'\})$ such that $i \in \tilde{I}$ or (ii) there does not exist $\tilde{I} \in C(I \cup \{i\})$ such that $i \in \tilde{I}$. Note that since U is increasing in scores, we have the following.

Claim 11. *If $i \in \hat{I}$ for some $\hat{I} \in C(I)$, $\theta(j) = \theta(i)$ and $s(j) > s(i)$, then there exists $j \in \hat{I}$.*

Both (i) and (ii) are proved in the same way. There are two subcases. Either for all $\tilde{I} \in C(I \cup \{i, j'\})$, there are fewer $\theta(i)$ individuals in $\tilde{I}$ compared to $\hat{I}$, or there exists a $\tilde{I} \in C(I \cup \{i, j'\})$ such that the number of $\theta(i)$ individuals in $\tilde{I}$ is weakly higher than $\hat{I}$. In the second subcase, from Claim 11, an individual who has same identity as i and weakly lower score must be selected, which means there exists $\tilde{I}' \in C(I \cup \{i, j'\})$ such that i is selected, which is a contradiction. In the first subcase, let $\tilde{I} \in C(I \cup \{i, j'\})$.

[3] Formally, the statement $\tau(I) \neq \tau(\tilde{I})$, $U(I) > U(\tilde{I})$, $\tilde{I} \in C(\hat{I})$ for some $\hat{I}$ that includes I still holds.

Then there exists another type θ' such that θ' has a higher number of individuals in $\tilde{I}$ compared to $\hat{I}$. Let k and k' denote the lowest scoring θ' individuals chosen under $\hat{I}$ and $\tilde{I}$. Therefore, from Claim 11, $s(k) \geq s(k')$. Moreover, we have

$$\begin{aligned} s(i) + h_{\theta(i)}(N_{\theta(i)}(\tilde{I}) + 1) - h_{\theta(i)}(N_{\theta(i)}(\tilde{I})) &\geq s(i) + h_{\theta(i)}(N_{\theta(i)}(\hat{I})) - h_{\theta(i)}(N_{\theta(i)}(\hat{I}) - 1) \\ &\geq s(k) + h_{\theta(k)}(N_{\theta(k)}(\hat{I}) + 1) - h_{\theta(i)}(N_{\theta(i)}(\hat{I})) \\ &\geq s(k) + h_{\theta(k)}(N_{\theta(k)}(\tilde{I})) - h_{\theta(i)}(N_{\theta(i)}(\tilde{I}) - 1) \end{aligned}$$

where first and third inequalities hold by from concavity of h_θ for all θ and second holds as $i \in \hat{I}$. However, this shows that switching i with k' does not decrease utility, and therefore there exists $\tilde{I} \setminus \{k'\} \cup \{i\}$ is also chosen, which is a contradiction. This completes the proof.

A.2 Diversity Reporting Figures

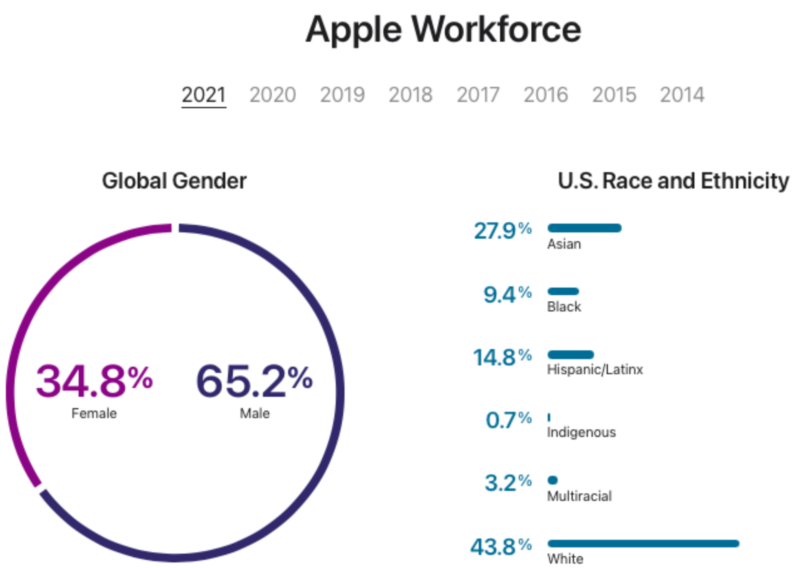


Figure A-1: Apple Diversity Statistics.

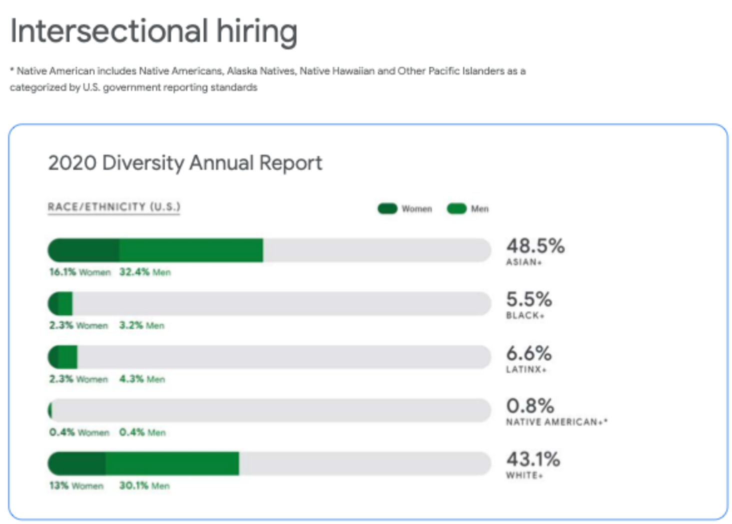


Figure A-2: Google Diversity Statistics.

Appendix B

Appendix to Adaptive Priority Mechanisms

B.1 Omitted Proofs

B.1.1 Proof of Proposition 6

Proof. Part (i): In state ω the payoff from admitting the highest-scoring minority students of measure $x(\omega)$ is:

$$q\omega + (1+\gamma-\omega)x(\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)x(\omega)^2 \tag{33}$$

Thus, the $x(\omega)$ that solves the FOC is given by:

$$x(\omega) = \frac{\kappa(1+\gamma-\omega)}{1+\kappa\gamma\beta} \tag{34}$$

Under our maintained assumptions, we have that:

$$x(\omega) = \frac{\kappa(1+\gamma-\omega)}{1+\kappa\gamma\beta} \leq \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega}-\underline{\omega}) < \min\{\kappa, q\} \tag{35}$$

and:

$$x(\omega) = \frac{\kappa(1+\gamma-\omega)}{1+\kappa\gamma\beta} \geq \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta} > \kappa(1-\underline{\omega}) \geq 0 \tag{36}$$

Thus, this level of minority admissions is feasible. Substituting, we have that:

$$V^* = q\mathbb{E}[\omega] + \frac{1}{2}\frac{\mathbb{E}[\kappa(1+\gamma-\omega)^2]}{1+\kappa\gamma\beta} \tag{37}$$

Consider now the APM $A(y) = \gamma(1-\beta y)$. Agents are allocated the resource if their modified scores exceed ω, with a uniform lottery over students with score exactly ω. Thus, in state ω, this policy admits measure $y(\omega)$ minorities that solve the fixed point equation:

$$y(\omega) = \min\left\{\kappa\int_0^1 \mathbb{I}[s + A(y(\omega)) \geq \omega]\mathrm{d}s, q\right\} = \min\{\kappa\left(1 - \max\{\omega - A(y(\omega)), 0\}\right), q\} \tag{38}$$

Denote the RHS of this fixed point equation by the function $\mathrm{RHS}(y,\omega)$, which is continuous and decreasing in y. Moreover, $\mathrm{RHS}(0,\omega) = \min\{\kappa(1-\max\{\omega-\gamma, 0\}), q\} > 0$ and $\mathrm{RHS}(\min\{\kappa, q\}, \omega) < \min\{\kappa, q\}$. The second of these inequalities is true because the condition $\underline{\omega} > \gamma(1 - \beta\min\{q,\kappa\})$ follows from our assumption that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$. Thus, there exists a unique $y(\omega)$ implemented by the APM. Moreover, let $y_A(\omega)$ denote the unique solution to the equation 38, which gives the measure of admitted minority students under APM A at state ω.

$$\begin{aligned} y_A(\omega) &= \kappa(1 - (\omega - \gamma(1 - \beta y_A(\omega)))) \\ &= \kappa(1 - \omega + \gamma) - \kappa\gamma\beta y_A(\omega) \\ &= \frac{\kappa(1-\omega+\gamma)}{1+\kappa\gamma\beta} \end{aligned} \tag{39}$$

Thus, A implements the optimal level of minority admissions characterized in equation 34 and $V_A = V^*$.

Part (ii): First, if we admit all minority students over some threshold $\hat{s}$, the total score of admitted minority students is $\kappa\int_{\hat{s}}^1 s\mathrm{d}s$. Moreover, when we admit measure x minority students where $x \leq \min\{\kappa, q\}$, this admissions threshold is defined by $x = \kappa\int_{\hat{s}}^1 \mathrm{d}s = \kappa(1-\hat{s})$. Thus, we have that $\hat{s} = 1 - \frac{x}{\kappa}$. Finally, the residual measure $q - x$ admitted majority students all score ω. Thus, the total score is given by $\bar{s} = q\omega + (1-\omega)x - \frac{1}{2\kappa}x^2$ for $0 \leq x \leq \min\{\kappa, q\}$. As both quota and priority policies always admit the highest-scoring minority students, the authority's utility is given by:

$$\mathcal{U} = q\mathbb{E}[\omega] + \mathbb{E}[(1+\gamma-\omega)x] - \frac{1}{2}\left(\frac{1}{\kappa} + \gamma\beta\right)\mathbb{E}[x^2] \tag{40}$$

We now derive the admitted measure of minority students. In the absence of a priority or quota policy, $\alpha = 0$ or $Q = 0$, we have that $x = \kappa(1-\omega)$ measure minority students is admitted. Thus, under a quota policy Q, measure $x = \max\{Q, \kappa(1-\omega)\}$ minority students are admitted. Under a priority policy, the measure of admitted minority students is $x = \kappa \int_{\omega-\alpha}^{1} \mathrm{d}x = \kappa(1+\alpha-\omega)$. In each case x is capped by $\min\{\kappa, q\}$ and floored by 0.

The expected utility function over quotas is given by one of four cases. First, $Q > \min\{\kappa, q\}$ and:

$$\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\min\{\kappa, q\} - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\min\{\kappa, q\}^2 \tag{41}$$

Second, $Q \in [\kappa(1-\underline{\omega}), \min\{\kappa, q\})$ and:[1]

$$\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])Q - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^2 \tag{42}$$

Third, $Q \in (\kappa(1-\overline{\omega}), \kappa(1-\underline{\omega}))$ and:

$$\begin{aligned}\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] &+ \int_{1-\frac{Q}{\kappa}}^{\overline{\omega}} \left((1+\gamma-\omega)Q - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^2\right)\mathrm{d}\Lambda(\omega) \\ &+ \int_{\underline{\omega}}^{1-\frac{Q}{\kappa}} \left((1+\gamma-\omega)\kappa(1-\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)(\kappa(1-\omega))^2\right)\mathrm{d}\Lambda(\omega)\end{aligned} \tag{43}$$

Finally, $Q \leq \kappa(1-\overline{\omega})$ and:

$$\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] + \mathbb{E}\left[(1+\gamma-\omega)\kappa(1-\omega)\right] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}\left[(\kappa(1-\omega))^2\right] \tag{44}$$

We claim that the optimum lies in the second case. See that in case two the strict maximum is attained at $Q^* = \frac{1+\gamma-\mathbb{E}[\omega]}{\frac{1}{\kappa}+\gamma\beta} \in (\kappa(1-\underline{\omega}), \min\{\kappa, q\})$, by our assumptions that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega}-\underline{\omega})$ and $\kappa(1-\underline{\omega}) < \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta}$. Moreover, in case three, the first derivative of the payoff is given by:

$$\mathcal{U}_Q'(Q) = \int_{1-\frac{Q}{\kappa}}^{\overline{\omega}} \left((1+\gamma-\omega) - \left(\frac{1}{\kappa}+\gamma\beta\right)Q\right)\mathrm{d}\Lambda(\omega) \tag{45}$$

Thus, checking that the sign of this is positive amounts to verifying that for all $Q \in (\kappa(1-\overline{\omega}), \kappa(1-\underline{\omega}))$, we have that:

$$Q < \frac{1+\gamma-\mathbb{E}[\omega|\omega \geq 1-\frac{Q}{\kappa}]}{\frac{1}{\kappa}+\gamma\beta} \tag{46}$$

[1] By our maintained assumptions we have that this interval has non-empty interior.

As the RHS is an increasing function of Q, it suffices to show that:

$$\kappa(1-\underline{\omega}) < \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta} \tag{47}$$

which we have assumed. Moreover, the expected utility in the first case equals $\mathcal{U}_Q(\kappa(1-\underline{\omega}))$, thus is lower than the optimum of the second case. The expected utility in the fourth case equals $\mathcal{U}_Q(\kappa(1-\overline{\omega}))$, thus is lower than the optimum of the third case. We therefore have that:

$$V_Q = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])Q^* - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^{*2} \tag{48}$$

We now turn to characterizing the value of priorities. There are three cases to consider. First, when $\kappa(1+\alpha-\overline{\omega}) \geq \min\{\kappa, q\}$ we have that $x = \min\{\kappa, q\}$ and:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\min\{\kappa, q\} - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\min\{\kappa, q\}^2 \tag{49}$$

Second, when $\kappa(1+\alpha-\underline{\omega}) \geq \min\{\kappa, q\} \geq \kappa(1+\alpha-\overline{\omega})$ we have that:

$$\begin{aligned}\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] &+ \int_{\underline{\omega}}^{1+\alpha-\min\{\frac{q}{\kappa},1\}} \left((1+\gamma-\omega)\min\{\kappa, q\} - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\min\{\kappa, q\}^2\right) \mathrm{d}\Lambda(\omega) \\ &+ \int_{1+\alpha-\min\{\frac{q}{\kappa},1\}}^{\overline{\omega}} \left((1+\gamma-\omega)\kappa(1+\alpha-\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)[\kappa(1+\alpha-\omega)]^2\right) \mathrm{d}\Lambda(\omega)\end{aligned} \tag{50}$$

Finally, when $\min\{\kappa, q\} \geq \kappa(1+\alpha-\underline{\omega})$, we have that:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + \mathbb{E}[(1+\gamma-\omega)\kappa(1+\alpha-\omega)] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}[(\kappa(1+\alpha-\omega))^2] \tag{51}$$

We claim that the optimum under our assumptions lies only in the third case. First, we argue that there is a unique local maximum in the third case. Second, we show the value in the second case is decreasing in α. By continuity, the unique optimum then lies in the third case.

First, it is helpful to write $\bar{x}(\alpha) = \kappa(1+\alpha-\mathbb{E}[\omega])$ and $\varepsilon = \kappa(\mathbb{E}[\omega]-\omega)$. The value in the third case can then be re-expressed as:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + \mathbb{E}[(1+\gamma-\omega)\,(\bar{x}(\alpha)+\varepsilon)] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}[(\bar{x}(\alpha)+\varepsilon)^2]$$
$$= q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\bar{x}(\alpha) - \mathbb{E}[\omega\varepsilon] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\bar{x}(\alpha)^2 - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}[\varepsilon^2] \tag{52}$$

Finally, we have that $\mathbb{E}[\varepsilon^2] = \kappa^2\text{Var}[\omega]$ and $\mathbb{E}[\omega\varepsilon] = \text{Cov}[\omega,\varepsilon] = -\kappa\text{Var}[\omega]$. Thus:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\bar{x}(\alpha) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\bar{x}(\alpha)^2 + \frac{\kappa}{2}\left(1-\kappa\gamma\beta\right)\text{Var}[\omega] \tag{53}$$

We then see that the optimal α^* in this range sets $\bar{x}(\alpha^*) = Q^* < \min\{\kappa, q\}$. It remains only to check that this optimal α^* indeed lies within this case, or equivalently that $\kappa(1+\alpha^*-\underline{\omega}) \leq \min\{\kappa, q\}$. To this end, see that $\kappa(1+\alpha^*-\mathbb{E}[\omega]) = Q^*$, and:

$$\kappa(1+\alpha^*-\underline{\omega}) = Q^* + \kappa(\mathbb{E}[\omega]-\underline{\omega}) \leq Q^* + \kappa(\overline{\omega}-\underline{\omega})$$
$$\leq \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega}-\underline{\omega}) < \min\{\kappa, q\} \tag{54}$$

where the final inequality follows by our assumption that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$.

Second, in the second case we have that the first derivative of the payoff in α is given by:

$$\mathcal{U}_P'(\alpha) = \int_{1+\alpha-\min\{\frac{q}{\kappa},1\}}^{\overline{\omega}} \frac{d}{d\alpha}\left((1+\gamma-\omega)\kappa(1+\alpha-\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)[\kappa(1+\alpha-\omega)]^2\right)\mathrm{d}\Lambda(\omega)$$
$$= \kappa\int_{1+\alpha-\min\{\frac{q}{\kappa},1\}}^{\overline{\omega}} \left((1+\gamma-\omega) - \left(\frac{1}{\kappa}+\gamma\beta\right)(\bar{x}(\alpha)+\varepsilon(\omega))\right)\mathrm{d}\Lambda(\omega) \tag{55}$$

Checking that the sign of this is negative for all α such that $\kappa(1+\alpha-\underline{\omega}) \geq \min\{\kappa, q\} \geq \kappa(1+\alpha-\overline{\omega})$ then amounts to checking that:

$$\bar{x}(\alpha) > \frac{1+\gamma-\mathbb{E}[\omega|\omega \geq 1+\alpha-\min\{\frac{q}{\kappa},1\}]}{\frac{1}{\kappa}+\gamma\beta} - \mathbb{E}\left[\varepsilon(\omega)|\omega \geq 1+\alpha-\min\{\frac{q}{\kappa},1\}\right] \tag{56}$$

for all $\bar{x}(\alpha) \in [\min\{\kappa, q\} - \kappa(\mathbb{E}[\omega]-\underline{\omega}), \min\{\kappa, q\} - \kappa(\mathbb{E}[\omega]-\overline{\omega})]$. So it suffices to check that the minimal possible value of the LHS exceeds the maximal possible value of the RHS. A sufficient condition for this is that:

$$\min\{\kappa, q\} - \kappa(\mathbb{E}[\omega] - \underline{\omega}) > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} - \kappa(\mathbb{E}[\omega] - \overline{\omega}) \tag{57}$$

Which holds as we assumed that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$. Substituting the optimal priority policy $\bar{x}(\alpha) = Q^*$ in equation 51, we obtain

$$V_P = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])Q^* - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^{*2} + \frac{\kappa}{2}(1-\kappa\gamma\beta)\,\mathrm{Var}[\omega] \tag{58}$$

We have now established that:

$$\Delta = V_P - V_Q = \frac{\kappa}{2}(1-\kappa\gamma\beta)\,\mathrm{Var}[\omega] \tag{59}$$

Part (iii): We have V^*, V_Q, V_P. Thus, we can compute the loss from restricting to quota policies:

$$\mathcal{L}_Q = \frac{1}{2}\frac{\kappa\mathrm{Var}[\omega]}{1+\kappa\gamma\beta} \tag{60}$$

To find the loss from restricting to priority policies, we compute:

$$\mathcal{L}_P = \mathcal{L}_Q - \Delta = \frac{1}{2}(\kappa\gamma\beta)^2\frac{\kappa\mathrm{Var}[\omega]}{1+\kappa\gamma\beta} \tag{61}$$

Enveloping over these losses yields the claimed formula. □

B.1.2 Proof of Proposition 7

Proof. Adapting Definition 15 to single object setting, we say that a matching μ admits a cutoff structure if there exists $S(\omega) = \{S_m(\omega)\}_{m\in\mathcal{M}}$ such that $\mu(s, m; \omega) = 1$ if and only if $s \geq S_m(\omega)$. A mechanism admits a cutoff structure if it admits a cutoff structure at every ω. We will first prove that any monotone APM admits a cutoff structure.

Lemma 12. *A monotone APM admits a cutoff structure.*

Proof. For a contradiction, assume it does not. Then there exists ω and matching μ implemented by the monotone APM such that for some $m \in \mathcal{M}$, $s > s'$, $\mu(s, m; \omega) = 0$ but $\mu(s', m; \omega) = 1$. Let x_m denote the measure of group m agents allocated the resource at μ. Since A is a monotone APM and $s > s'$, we have that $A(x_m, s) > A(x_m, s')$, which contradicts that A implements μ. □

We now use Lemma 12 to show that a monotone APM implements a unique allocation. Assume for a contradiction that $A_m(y_m, s)$ is monotone and implements two different allocations, μ and μ'. Let x_m and x'_m denote the measure of type m students assigned the resource at μ and μ'. First, we prove that if μ and μ' admit the same measure of students from each group, *i.e.*, $x_m = x'_m$ for all m, then the average score of admitted students are the same. Let s_m and s'_m denote the score of the lowest-scoring type m students assigned the resource at μ and μ'.

Claim 1. *If $x_m = x'_m$ for all $m \in \mathcal{M}$, then $\bar{s}_h(\mu, \omega) = \bar{s}_h(\mu', \omega)$*

Proof. Fix an m. Without loss of generality, let $s_m \geq s'_m$. First, since APM has cutoff structure and $x_m = x'_m$, we have that

$$\int_{\Theta} \mathbb{I}\{s(\theta) \in [s'_m, s_m], m(\theta) = m\} \mathrm{d}F_\omega(s, m) = 0 \tag{62}$$

Note that this holds regardless of $m' \in \mathcal{M}$ and whether $s_m \geq s'_m$ or $s'_m \geq s_m$. Therefore,

$$\begin{aligned}
\bar{s}_h(\mu, \omega) &= \int_{\Theta} \mu(s, m) h(s) \mathrm{d}F_\omega(s, m) \\
&= \sum_{m \in \mathcal{M}} \int_{\Theta} \mathbb{I}\{s(\theta) \geq s_m, m(\theta) = m\} h(s(\theta)) \mathrm{d}F_\omega(s, m) \\
&= \sum_{m \in \mathcal{M}} \int_{\Theta} \mathbb{I}\{s(\theta) \geq s'_m, m(\theta) = m\} h(s(\theta)) \mathrm{d}F_\omega(s, m) \\
&= \int_{\Theta} \mu(s, m) h(s) \mathrm{d}F_\omega(s, m) \\
&= \bar{s}_h(\mu', \omega)
\end{aligned} \tag{63}$$

where line equation holds from Equation 62 and all others are by definition. This finishes the proof of the claim. ◻

Therefore, if μ and μ' do not yield identical measures, then there are m and n such that $x_m > x'_m$ and $x'_n > x_n$. Since $x_m > x'_m$, it follows that $s'_m > s_m$. Likewise $x'_n > x_n$ implies that $s_n > s'_n$. Note that these imply: (i) $\mu'(s'_n, n) = 1$ while $\mu'(s'_m, n) = 0$ and (ii) $\mu(s_m, m) = 1$ while $\mu(s'_n, n) = 0$. Thus, the following inequalities hold:

$$A_n(s'_n, x'_n) > A_m(s'_m, x'_m) \geq A_m(s_m, x_m) > A_n(s'_n, x_n) \geq A_n(s'_n, x'_n) \tag{64}$$

where the first inequality follows from (i), the second inequality follows from the fact that $x'_m < x_m$ and A is monotone, the third inequality follows from (ii) and the fourth inequality follows from the fact that $x_n < x'_n$ and A is monotone. This equation yields $A_n(s'_n, x'_n) > A_n(s'_n, x'_n)$, which is a contradiction. Therefore, all allocations implemented by A yield the same x. Thus, from Lemma 12, if a monotone APM A implements μ and μ', both allocations admit the highest-scoring measure x_m agents from group m and can differ (at most) on a measure 0 set, proving the essential uniqueness of A. □

B.1.3 Proof of Theorem 1

Proof. We characterize the optimal allocation for each $\omega \in \Omega$ and show that the claimed adaptive priority mechanism implements the same allocation. Fix an $\omega \in \Omega$ and suppress the dependence of F_ω and f_ω thereon, and define the utility index of a score as $\tilde{s} = h(s)$ with induced densities over $\tilde{s}$ given by $\tilde{f}_m$ for all $m \in \mathcal{M}$. Let the measure of agents from any group $m \in \mathcal{M}$ that is allocated the resource be $x_m \in [0, \bar{x}_m]$ where $\bar{x}_m = \int_{h(0)}^{h(1)} \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s}$. Observe that, conditional on fixing the measures of agents from each group that are allocated the resource $x = \{x_m\}_{m\in\mathcal{M}}$, there is a unique optimal allocation (*i.e.*, ξ-maximal μ up to measure zero transformations). In particular, as g and h are continuous and strictly increasing, the optimal allocation conditional on x satisfies $\mu^*(\tilde{s}, m; x) = 1 \iff \tilde{s} \geq \underline{\tilde{s}}_m(x_m)$ for some thresholds $\{\underline{\tilde{s}}_m(x_m)\}_{m\in\mathcal{M}}$ that solve:

$$\int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} = x_m \tag{65}$$

We can then express the problem of choosing the optimal $x = \{x_m\}_{m\in\mathcal{M}}$ as:

$$\max_{x_m\in[0,\bar{x}_m],\,\forall m\in\mathcal{M}} \sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} + \sum_{m\in\mathcal{M}} u_m(x_m) \quad \text{s.t.} \sum_{m\in\mathcal{M}} x_m \leq q \tag{66}$$

where a solution exists by compactness of the constraint sets and continuity of the objective. We can derive necessary and sufficient conditions on the solution(s) to this problem by considering the Lagrangian:

$$\begin{aligned} \mathcal{L}(x, \lambda, \overline{\kappa}, \underline{\kappa}) = &\sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} + \sum_{m\in\mathcal{M}} u_m(x_m) \\ &+ \lambda\left(q - \sum_{m\in\mathcal{M}} x_m\right) + \sum_{m\in\mathcal{M}} \overline{\kappa}_m(\bar{x}_m - x_m) + \sum_{m\in\mathcal{M}} \underline{\kappa}_m x_m \end{aligned} \tag{67}$$

The first-order necessary conditions to this program are given by:

$$\frac{\partial \mathcal{L}}{\partial x_m} = -\underline{\tilde{s}}'_m(x_m)\underline{\tilde{s}}_m(x_m)\tilde{f}_m(\underline{\tilde{s}}_m(x_m)) + u'_m(x_m) - \lambda - \overline{\kappa}_m + \underline{\kappa}_m = 0 \tag{68}$$

$$\lambda \frac{\partial \mathcal{L}}{\partial \lambda} = \lambda \left(q - \sum_{m \in \mathcal{M}} x_m \right) = 0 \tag{69}$$

$$\overline{\kappa}_m \frac{\partial \mathcal{L}}{\partial \overline{\kappa}_m} = \overline{\kappa}_m (\bar{x}_m - x_m) = 0 \tag{70}$$

$$\underline{\kappa}_m \frac{\partial \mathcal{L}}{\partial \underline{\kappa}_m} = \underline{\kappa}_m x_m = 0 \tag{71}$$

for all $m \in \mathcal{M}$. By implicitly differentiating Equation 65, we obtain that:

$$-\underline{\tilde{s}}'_m(x_m)\tilde{f}_m(\underline{\tilde{s}}_m(x_m)) = 1 \tag{72}$$

Thus, we can simplify Equation 68 to:

$$\frac{\partial \mathcal{L}}{\partial x_m} = \underline{\tilde{s}}_m(x_m) + u'_m(x_m) - \lambda - \overline{\kappa}_m + \underline{\kappa}_m = 0 \tag{73}$$

Observe that all constraints are linear. Thus, if the objective function is strictly concave, the first-order conditions are also sufficient. Observe by Equation 72 that $\underline{\tilde{s}}_m(x_m)$ is a strictly decreasing function of x_m, and all cross-partial derivatives are zero. Therefore, the first summation is strictly concave. Moreover u'_m is a decreasing function of x_m by virtue of the assumption that u_m is concave for all $m \in \mathcal{M}$. Therefore, the second summation is concave. Thus, the objective function is strictly concave and the optimal allocation is unique.

Thus, to verify that our claimed adaptive priority mechanism is a first-best mechanism, it suffices to show that the allocation it implements satisfies Equations 68 to 71. The adaptive priority mechanism $A_m(y_m, s) = h^{-1}\left(h(s) + u'_m(y_m)\right)$ in the transformed score space yields transformed scores $h\left(A_m(y_m, s)\right) = \tilde{s} + u'_m(y_m)$. Define x_m as the admitted measure of agents from group m under this mechanism. Agents in group $m \in \mathcal{M}$ are allocated the resource if and only if $\tilde{s} + u'_m(x_m) \geq s^C$ for some threshold s^C that solves:

$$\sum_{m \in \mathcal{M}} \int_{\max\{\min\{s^C - u'_m(x_m), h(1)\}, h(0)\}}^{h(1)} \tilde{f}_m(\tilde{s}) \mathrm{d}\tilde{s} = q \tag{74}$$

We can therefore partition $\mathcal{M}$ into three sets that are uniquely defined: (i) interior

$\mathcal{M}_I = \{m \in \mathcal{M} | s^C - u'_m(x_m) \in (h(0), h(1))\}$; (ii) no allocation $\mathcal{M}_0 = \{m \in \mathcal{M} | s^C - u'_m(x_m) \geq h(1)\}$; (iii) full allocation $\mathcal{M}_1 = \{m \in \mathcal{M} | s^C - u'_m(x_m) \leq h(0)\}$. For all $m \in \mathcal{M}_0$, we implement $x_m = 0$. For all $m \in \mathcal{M}_1$, we implement $x_m = \bar{x}_m$. For all $m \in \mathcal{M}_I$, we implement $x_m \in (0, \bar{x}_m)$. For any $m \in \mathcal{M}_I$, the allocation threshold is $\underline{\tilde{s}}_m(x_m) = s^C - u'_m(x_m)$. For any $m \in \mathcal{M}_0$, the allocation threshold is $h(1)$. For any $m \in \mathcal{M}_1$, the allocation threshold is $h(0)$.

We now verify that this outcome satisfies the established necessary and sufficient conditions. For all $m \in \mathcal{M}_I$, by the complementary slackness conditions we have that $\underline{\kappa}_m = \overline{\kappa}_m = 0$. Substituting the above into Equation 68 for all $m \in \mathcal{M}_I$ we obtain that:

$$s^C - \lambda = 0 \tag{75}$$

which is satisfied for $\lambda = s^C$. As $q = \sum_{m \in \mathcal{M}} x_m$, the complementary slackness condition for λ is then satisfied. For all $m \in \mathcal{M}_0$, by complementary slackness we have that $\overline{\kappa}_m = 0$ and Equation 68 is satisfied by:

$$\underline{\kappa}_m = \lambda - h(1) - u'_m(0) \tag{76}$$

For all $m \in \mathcal{M}_1$, by complementary slackness we have that $\underline{\kappa}_m = 0$ and Equation 68 is satisfied by:

$$\overline{\kappa}_m = h(0) + u'_m(\bar{x}_m) - \lambda \tag{77}$$

This completes the proof of first-best optimality of A^*. Moreover, as the optimal allocation is unique for all ω, any allocation that differs from the allocation implemented by the optimal APM at any ω would not be first-best optimal. Therefore, any first-best-optimal mechanism must implement essentially the same allocation as A^*. □

B.1.4 Proof of Theorem 2

Proof. First, we prove the if parts of the results.

Part (i): When u_m is linear, u'_m is constant and the first-best optimal adaptive priority mechanism is a priority mechanism $P(s, m) = h^{-1}(h(s) + u'_m)$. Part (ii): When $\tilde{u}'_m(x_m) \geq k$ for $x_m \leq x_m^{\text{tar}}$ and $\tilde{u}'_m(x_m) = 0$ for $x_m > x_m^{\text{tar}}$ and $\sum_{m \in \mathcal{M}} x_m^{\text{tar}} < q$, observe that the optimal mechanism admits $x_m \geq x_m^{\text{tar}}$ for all $m \in \mathcal{M}$ in all states of the world, but conditional on $x_m \geq x_m^{\text{tar}}$ for all $m \in \mathcal{M}$ admits the highest-scoring set of agents. A quota $Q_m = x_m^{\text{tar}}$ and $Q_R = q - \sum_{m \in \mathcal{M}} x_m^{\text{tar}}$, with $D(R) = |\mathcal{M}| + 1$ implements this allocation and is first-best optimal for any authority that is extremely risk-averse.

Second, we prove the only if parts of the results.

Part (i): Assume the utility functions are not linear and let m denote a group where u'_m is not constant in $[0, q]$. We say that a state ω has full support if f_w has full support. A state ω has *full support in m and n* if $f_w(\cdot, m) > 0$ and $f_w(\cdot, n) > 0$ for some m and n and positive measures of only m and n. Let ω be a state that has full support in m and n. Moreover, assume both groups have a measure q of agents. We first establish that in any optimal allocation, agents from both groups are allocated the resource.

Claim 2. *If preferences are non-trivial, then the optimal allocation has $x_n, x_m > 0$.*

Proof. Toward a contradiction, suppose without loss of generality that $x_n = 0$. This implies that $x_m = q$. By the necessary first-order condition from Theorem 1 (combing Equations 68 and 71), we have that:

$$u'_m(q) + h(0) = u'_n(0) + h(1) + \underline{\kappa}_n \geq u'_n(0) + h(1) \tag{78}$$

where the inequality follows as $\underline{\kappa}_n \geq 0$. Thus, we have that:

$$u'_m(q) - u'_n(0) \geq h(1) - h(0) > u'_m(q) - u'_n(0) \tag{79}$$

where the first inequality follows by rearranging Equation 78 and the second follows by the definition of non-triviality of preferences. This is a contradiction, thus $x_n, x_m > 0$ in any optimal allocation. □

We now establish an equation relating x_n and x_m that will be useful in the steps to come.

Claim 3. *Let ω have full support in m and n, μ denote a cutoff matching with cutoffs s_m and s_n. Let x_m and x_n denote the measures of agents who are allocated the object at μ. μ is optimal if and only if $u'_m(x_m) + h(s_m) = u'_n(x_n) + h(s_n)$ and $x_n + x_m = q$.*

Proof. By the necessary and sufficient first-order conditions from Theorem 1, we again have that:

$$u'_m(x_m) + h(s_m) - \overline{\kappa}_m + \underline{\kappa}_m = u'_n(x_n) + h(s_n) - \overline{\kappa}_n + \underline{\kappa}_n \tag{80}$$

By Claim 2, we have $x_m, x_n > 0$. Thus, by the complementary slackness conditions

(Equations 70 and 71), we have that $\overline{\kappa}_m = \underline{\kappa}_m = \overline{\kappa}_n = \underline{\kappa}_n = 0$. Thus, we obtain:

$$u'_m(x_m) + h(s_m) = u'_n(x_n) + h(s_n) \tag{81}$$

together with $x_n + x_m = q$, we have characterized the optimal allocation as claimed. □

Continue to let x_m and x_n denote the measures of group m and n agents at the optimal allocation under ω, and s_m and s_n denote the cutoff scores for admission. There are now two cases to consider: (i) $u'_m(x_m)$ and $u'_n(x_n)$ are locally constant. (ii) $u'_m(x_m)$ or $u'_n(x_n)$ are not locally constant. If we are in case (i), we will construct an ω' with a unique optimal allocation x'_m and x'_n where $u'_m(x'_m)$ or $u'_n(x'_n)$ is not locally constant, and then show jointly how we arrive at a contradiction in both cases (i) and (ii).

To this end, suppose that we are in case (i). Let x^*_m and x^*_n denote the measures that are closest to x_m and x_n such that $u'_m(x_m)$ and $u'_n(x_n)$ are not locally constant, *i.e.*,:

$$\begin{aligned} x^*_k = \arg\min_{x'_k} \Bigg\{ |x_k - x'_k| \,\Big|\, & u'_k(x_k) = u'_k(x'_k) \text{ and for all } \varepsilon > 0 \\ & \text{either } u'_k(x'_k - \varepsilon) > u'_k(x_k) \text{ or } u'_k(x'_k + \varepsilon) > u'_k(x_k) \Bigg\} \end{aligned} \tag{82}$$

As u'_k is continuous, this minimum is attained and x^*_k is well-defined. Without loss of generality, assume $|x_m - x^*_m| \leq |x_n - x^*_n|$ and define both $\hat{x}_m = x^*_m$ and $\hat{x}_n = q - x^*_m$. We now construct a state ω' such that $\hat{x}$ is optimal:

Claim 4. *Define ω' where $F_m(1) - F_m(s_m) = \hat{x}_m$, $F_n(1) - F_n(s_n) = \hat{x}_n$ and ω' has full support in m and n. The allocation that admits the highest-scoring $\hat{x}_m$ group m agents and the highest-scoring $\hat{x}_n$ group n agents is the unique optimal allocation.*

Proof. By Claim 3, as $\hat{x}_m + \hat{x}_n = q$ by construction, $\hat{x}$ is optimal if and only if Equation 81 holds. To this end, observe that if we admit $\hat{x}$, then the cutoff scores are the same as under x as $F_m(1) - F_m(s_m) = \hat{x}_m$ and $F_n(1) - F_n(s_n) = \hat{x}_n$, by construction. Thus, we have that:

$$\begin{aligned} u'_m(\hat{x}_m) + h(s_m) &= u'_m(x_m) + h(s_m) \\ &= u'_n(x_n) + h(s_n) \\ &= u'_n(\hat{x}_n) + h(s_n) \end{aligned} \tag{83}$$

where the first equality holds by construction as $\hat{x}_m = x_m^*$ and $u_m'(x_m^*) = u_m'(x_m)$, the second equality holds by optimality of x, and the third equality holds as $|x_m - x_m^*| \leq |x_n - x_n^*|$, which implies that $u_n'(\hat{x}_n) = u_n'(x_n^*)$. Thus, Equation 81 holds, and $\hat{x}$ is optimal, as claimed. □

Observe that this construction also applies trivially in case (ii) with $x_m^* = x_m$. Thus, using this construction, we can now study cases (i) and (ii) together. In state ω', to implement this optimal allocation, we must have that $P(s,m) < P(s_n,n)$ for all but a measure zero set of s such that $s < s_m$. We will now construct another state ω'' such that any priority mechanism with this property is suboptimal.

First, suppose that $x_m^* \leq x_m$ and fix some $\varepsilon \in (0, x_m^*)$. Define $\tilde{s}_m$ as solving the following equation:

$$u_m'(\hat{x}_m - \varepsilon) + h(\tilde{s}_m) = u_n'(\hat{x}_n + \varepsilon) + h(s_n) \tag{84}$$

We then have that:

$$\begin{aligned} \tilde{s}_m &= h^{-1}\left(h(s_n) + u_n'(\hat{x}_n + \varepsilon) - u_m'(\hat{x}_m - \varepsilon)\right) \\ &< h^{-1}\left(h(s_n) + u_n'(\hat{x}_n) - u_m'(\hat{x}_m)\right) \\ &= s_m \end{aligned} \tag{85}$$

where the first equality rearranges Equation 84 and the second inequality uses the facts that $u_n'(\hat{x}_n) \leq u_n'(\hat{x}_n + \varepsilon)$ and $u_m'(\hat{x}_m) < u_m'(\hat{x}_m - \varepsilon)$. We now construct a state ω'' such that $(\hat{x}_m - \varepsilon, \hat{x}_n + \varepsilon)$ is optimal.

Claim 5. *Define ω'' where $1 - F_m(\tilde{s}_m) = \hat{x}_m - \varepsilon$, $1 - F_n(s_n) = \hat{x}_n + \varepsilon$ with full support in m and n. The allocation that admits the highest-scoring $(\hat{x}_m - \varepsilon, \hat{x}_n + \varepsilon)$ agents is the unique optimal allocation.*

Proof. Following the same steps as Claim 4, and the fact that Equation 84 holds by construction, we have that the claim holds. □

Observe that to implement this optimal allocation a priority mechanism must set $P(s,m) \geq P(s_n,n)$ for all but zero measure $s > \tilde{s}_m$. However, since $\tilde{s}_m < s_m$, this contradicts the optimality condition for state ω' that $P(s,m) < P(s_n,n)$. This is because for all but measure zero $s \in (\tilde{s}_m, s_m)$, which we have established is non-empty, we have that:

$$P(s,m) \geq P(s_n,n) > P(s,m) \tag{86}$$

which is a contradiction. To complete the proof, we need only consider the case that $x^*_m > x_m$. In this case, we can apply essentially the same steps and the result follows. Concretely, instead increasing $\hat{x}_m$ by ε and following the same steps yields the required contradiction.

We have now constructed three states $\omega, \omega', \omega''$ such that no priority mechanism can be optimal in each state when the authority is not risk-neural, completing the proof.

Part (ii): Assume that a quota policy is optimal, we now show that the authority's preferences must be extremely risk-averse. For each group $m \in \mathcal{M}$, let $c_m \in [0,1]$ and $c_m \neq c_n$ if $m \neq n$. Let ω be such that the scores of agents from each group m are uniformly distributed between $[c_m, c_m + \epsilon]$, where ϵ is chosen to be small so that there is no overlap of these supports and each group has measure q agents. Let m_ω denote the group with the highest c_m at ω. Now, compute the optimal allocation at ω and denote the measure of admitted agents from each group at the optimal allocation by $\{x^*_m(\omega)\}_{m \in \mathcal{M}}$. We first show that under any optimal quota policy, the level of the quotas must be set equal to the optimal allocation for all but the highest-scoring group:

Claim 6. *If a quota policy Q attains the optimal allocation, then for each $m \neq m_\omega$, $Q_m = x^*_m(\omega)$.*

Proof. If $Q_m > x^*_m(\omega)$, then we admit $x_m \geq Q_m > x^*_m(\omega)$, which is suboptimal as there is a unique optimal allocation by Theorem 1. If $Q_m < x^*_m(\omega)$ and $m \neq m_\omega$, then $x_m = Q_m$ as $c_{m_\omega} > c_m + \varepsilon$ and no agent from group m can claim a merit slot. This is suboptimal. Thus, $Q_m = x^*_m(\omega)$ for all $m \neq m_\omega$. □

Next, create ω' by changing the highest-scoring group, *i.e.*, $m_\omega \neq m_{\omega'}$. Let $x^*_{m_\omega}(\omega')$ denote the measure of admitted agents from group m_ω under ω'. Applying Claim 6, If Q attains the optimal allocation, then it must be that $Q_{m_\omega} = x^*_{m_\omega}(\omega')$. Define Q^*_m by $Q^*_m = x^*_m$ for all $m \in \mathcal{M} \setminus \{m_\omega\}$ and $Q^*_{m_\omega} = x^*_{m_\omega}(\omega')$.

Now, we have proved that if Q is an optimal policy, then $Q_m = Q^*_m$ for all $m \in \mathcal{M} \setminus \{m_\omega\}$ and $Q_{m_\omega} = Q^*_{m_\omega}$. We now establish that merit slots must be processed after any positive measure quota slots if the merit slots are of positive measure:

Claim 7. *If there is a quota policy that attains the first-best, Q, then $Q_m = Q^*_m$ and either $\sum_{m \in \mathcal{M}} Q^*_m = q$, i.e., there are no merit slots (thus the order of processing the merit slots does not matter), or $\sum_{m \in \mathcal{M}} Q^*_m < q$ and merit slots are processed after any positive measure quota slots.*

Proof. We have already proved $Q_m = Q_m^*$. If $\sum_{m\in\mathcal{M}} Q_m^* = q$, there are no merit slots and any processing order yields the same result. If $\sum_{m\in\mathcal{M}} Q_m^* < q$, for a contradiction, assume merit slots are processed before quota slots for group m and $Q_m^* > 0$. There are two cases, $m \neq m_\omega$ and $m = m_\omega$. We start with the first case. Note that there is a cutoff s_m for group m with $s_m < c_m + \epsilon$ and all agents from group m who score above s_m are allocated the resource. Create ω'' by taking measure $x_m/2$ of these agents who are allocated the resource and give them scores above c_{m_ω} (the highest-scoring group at ω). The scores of the remaining $x_m/2$ agents are distributed uniformly at $[c_m, c_m + \epsilon]$.

We now observe that the optimal allocations at ω and ω'' are the same. This is because increasing the scores of already admitted agents does not change the preferences of the authority of whom to admit. Moreover, the optimal allocation at ω'' cannot be attained if the quota slots for group m are processed after the merit slots. This follows as, if merit slots are processed before quota slots for group m, a strictly positive measure of them would go to group m agents at ω'' since now they have a measure of agents with the highest scores, which violates optimality.

This proves the claim for $m \neq m_\omega$. To prove the result for $m = m_\omega$, replicate the above steps with ω' where m_ω is not the highest-scoring group. □

We now use these claims to establish that if quotas are first-best optimal, then (u, h) must agree with $(\tilde{u}, \tilde{h})$ on optimal allocations.

Claim 8. *The quota first policy with $Q_m = x_m^{tar}$ maximizes the utility with respect to $\tilde{u}, \tilde{h}$.*

Proof. This is clear since for $\tilde{u}, \tilde{h}$, diversity utility dominates until x_m^{tar} and has no effect after. □

This proves the result since if there exists a first-best optimal quota policy, then it is rationalized by $(\tilde{u}, \tilde{h})$ with $x_m^{tar} = Q_m^*$. Hence, if there is a first-best quota mechanism, the authority is extremely risk-averse. □

B.1.5 Proof of Proposition 8

Let x_m^* denote the measure of group m agents in the optimal allocation, with $x^* = \{x_m^*\}_{m\in\mathcal{M}}$. A priority policy $P(s, m) = h^{-1}(h(s) + u'_m(x_m^*)) = A_m(x_m^*, s)$ implements the same allocation as the optimal adaptive priority mechanism and by Theorem 1, is optimal. A quota mechanism with (Q, D) where $Q_m = x_m^*$ implements x^* for all D, and is therefore optimal.

B.1.6 Proof of Theorem 3

Proof. We first prove the following lemma.

Lemma 13. *Any stable matching is a cutoff matching.*

Proof. Assume that μ is a stable matching. Let $S_{m,c} = \inf_\theta\{s_c(\theta) : m(\theta) = m, \mu(\theta) = c\}$. Since μ satisfies within-group fairness, for all m and $s' > S_{m,c}$, if $m(\theta) = m$ and $s_c(\theta) = s'$, $\mu(\theta) \succeq_\theta c$. Moreover, from part (iv) of the definition of matching, this extends to the case where $s' = S_{m,c}$. Concretely, suppose that $\mu(\theta) \neq c$, $c \succ_\theta \mu(\theta)$ and $s_c(\theta) = S_{m,c}$. Consider a sequence of types $\{\theta_k\}_{k\in\mathbb{N}}$ with common group m and scores $\{s_c(\theta_k)\}_{k\in\mathbb{N}}$ such that $s_c(\theta_k) > S_{m,c}$ for all $k \in \mathbb{N}$ and $s_k(\theta) \to S_{m,c}$. Define the set $\Theta^E = \{\theta \in \Theta : c \succ \mu(\theta)\}$, which must be open by part (iv) of the definition of a matching. We have that $\theta_k \notin \Theta^E$ for all $k \in \mathbb{N}$ but $\lim_{k\to\infty} \theta_k \in \Theta^E$, which contradicts that Θ^E is open. Thus, if μ is stable, then it is also a cutoff matching. □

Therefore, to characterize stable matchings, it is enough to characterize cutoffs that induce a stable matching, which we call *stable cutoffs.*

Definition 19. *A vector S is a market-clearing cutoff if it satisfies the following:*

1. $D_c(S) \leq q_c$ *for all* c.
2. $D_c(S) = q_c$ *if* $S_{m,c} > 0$ *for some* $m \in \mathcal{M}$.

Since an authority can admit different measures of agents from different groups, there is a continuum of cutoffs that clear the market given S_{-c}, as long as $\{(0, \dots, 0)\}$ is not the only market-clearing cutoff. Let $I(S_{-c})$ denote the set of market-clearing cutoffs. Let $I^*(S_{-c}) \subseteq I(S_{-c})$ denote the unique (by Lemma 12) cutoffs that implement the outcome under APM A_c^* when the authority faces the induced type distribution over the set $\tilde{D}_c(S_{-c})$. Define the map $T_c : [0,1]^{|\mathcal{M}|\times|\mathcal{C}|} \to [0,1]^{|\mathcal{M}|}$ as $T_c(S) = I_c^*(S_{-c})$ with $T : [0,1]^{|\mathcal{M}|\times|\mathcal{C}|} \to [0,1]^{|\mathcal{M}|\times|\mathcal{C}|}$ given by $T = \{T_c\}_{c\in\mathcal{C}}$. We first show that the set of fixed points of T equals the set of stable cutoffs:

Claim 9. *The set of fixed points of T equals the set of stable cutoffs.*

Proof. If S^*, with corresponding matching μ^* (by Lemma 13), is a fixed point of T, then each $c \neq c_0$ admits their most preferred measure q_c agents in $\tilde{D}_c(S^*_{-c})$ (by Theorem 1). Note that any $\hat{\Theta}$ that can block the matching must prefer c to their allocation at μ^* and therefore $\hat{\Theta} \subset \tilde{D}_c(S^*_{-c})$. Then there cannot be a $\hat{\Theta}$ that blocks μ^* at c since c already attains the first-best utility under $\tilde{D}_c(S^*_{-c})$ from the definition of $T_c(S)$ and Theorem 1. Conversely, if S^*, with corresponding matching μ^*, is a not fixed point of T, then there exists c such that $T_c(S^*) \neq S_c^*$. Let $\hat{\Theta}$ denote the set of agents who are not matched to c at μ^* but have scores greater than $T_c(S^*)$, and $\tilde{\Theta}$ denote the set of agents who are matched to c at μ^* but have scores lower than $T_c(S^*)$. From optimality of A_c^* (by Theorem 1), $\hat{\Theta}$ blocks μ^* at c with $\tilde{\Theta}$. □

We now show that T is increasing.

Claim 10. *T is increasing.*

Proof. Fix an arbitrary $c \in \mathcal{C}$ and suppose that $S'_{-c} \geq S_{-c}$. Toward a contradiction suppose that there exists $m \in \mathcal{M}$ such that $t'_{c,m} = T_{c,m}(S') = I_c^*(S'_{-c}) < I_c^*(S_{-c}) = T_{c,m}(S) = t_{c,m}$, *i.e.*, the admissions threshold for group m at authority c goes down. Let f and f' be the induced joint densities of agents over scores at c and groups by the sets $\tilde{D}_c(S_{-c})$ and $\tilde{D}_c(S'_{-c})$, respectively. Let $\{x_{m,c}\}_{m\in\mathcal{M}}$ and $\{x'_{m,c}\}_{m\in\mathcal{M}}$ denote the measure agents who score above $t_{m,c}$ for their group (*i.e.*, admitted under A_c^*) under $\tilde{D}_c(S_{-c})$ and $\tilde{D}_c(S'_{-c})$, respectively. As $S'_{-c} \geq S_{-c}$, we have that $D^c(S_{-c}, S_c) \subseteq D^c(S'_{-c}, S_c)$ for all $S_c \in [0,1]^{|\mathcal{M}|}$. It follows that $f'(\theta_c) \geq f(\theta_c) > 0$ for all $\theta_c = (s_c, m_c) \in [0,1] \times \mathcal{M}$. As $t'_{c,m} < t_{c,m}$, f' has full support, and $f' \geq f$, we have that the measure of admitted group m agents under increases $x'_{c,m} > x_{c,m}$. But as $\sum_{k\in\mathcal{M}} x'_k = \sum_{k\in\mathcal{M}} x_k = q$, we know that there exists an $m' \in \mathcal{M}$ such that $x'_{c,m'} < x_{c,m'}$. It follows that $t'_{c,m'} > t_{c,m'}$, otherwise, if $t'_{c,m'} \leq t_{c,m'}$, then $x'_{c,m'} \geq x_{c,m'}$. But now we have shown the following:

$$h_c(t'_{c,m'})+u'_{m',c}(x'_{c,m'}) > h_c(t_{c,m'})+u'_{m',c}(x_{c,m'}) \geq h_c(t_{c,m})+u'_{m,c}(x_{c,m}) > h_c(t'_{c,m})+u'_{m,c}(x'_{c,m}) \tag{87}$$

where the first inequality follows by $t_{c,m'} < t'_{c,m'}$, $x_{c,m'} > x'_{c,m'}$, concavity of u_m and strictly increasing h_c. The second inequality follows by optimality. This is because the facts that $t_{c,m} > 0$ and $t'_{c,m'} < 1$ imply that $\overline{\kappa}_m = \underline{\kappa}_{m'} = 0$ and so Equation 73 implies that:

$$h_c(t_{c,m'}) + u'_{m',c}(x_{c,m'}) - \overline{\kappa}_{m'} = h_c(t_{c,m}) + u'_{m,c}(x_{c,m}) + \underline{\kappa}_m \tag{88}$$

with $\overline{\kappa}_{m'}, \underline{\kappa}_m \geq 0$. The final inequality follows as $t'_{c,m} < t_{c,m}$ and $x'_{c,m} > x_{c,m}$. But this contradicts the optimality condition for APM (Theorem 1), which implies that $T_c \neq I_c^*$, which is a contradiction. Hence, for all c and $m \in \mathcal{M}$, $T_{c,m}$ is an increasing function. □

As $T : [0,1]^{|\mathcal{M}|\times|\mathcal{C}|} \to [0,1]^{|\mathcal{M}|\times|\mathcal{C}|}$ is monotone and $[0,1]^{|\mathcal{M}|\times|\mathcal{C}|}$ is a lattice under the elementwise order $\geq$, Tarski's fixed point theorem implies that the set of stable matching cutoffs is a non-empty lattice.

Finally, we use the fact that the set of stable cutoffs is a complete lattice to argue that there is a unique cutoff consistent with stability.

Claim 11. *The stable matching cutoffs are unique.*

Proof. Assume that there are multiple stable cutoffs. As the set of stable cutoffs is a lattice, there exists a largest (S^+) and smallest (S^-) stable cutoffs, where $S^+ \geq S^-$, with strict inequality for some $m \in \mathcal{M}$, $c \in \mathcal{C}$ as $S^+ \neq S^-$. But then, as there is full support of agent types and authority c fills the capacity under stable cutoffs S^+, it must exceed its capacity under S^-, which is a contradiction. Hence, we have shown that there exists a unique stable matching cutoff. □

The combination of Lemma 13 and Claim 11 completes the proof. □

B.1.7 Proof of Theorem 4

Proof. If ϕ is equivalent to A_c^*, Claim 9 implies that ϕ is consistent with stability.

We prove consistency with stability implies that ϕ is equivalent to A_c^* by the contrapositive. To this end, suppose that ϕ is not equivalent to A_c^*. It follows that there exists a full-support density $\{\tilde{f}(s_c, m)\}_{s_c \in [0,1], m \in \mathcal{M}}$ such that ϕ yields a different allocation than A_c^* under $\tilde{f}$. The rest of the proof constructs a full-support measure F with unique stable matching μ_F such that $\tilde{f}$ is the induced density of scores and groups of the agents who demand authority c at μ_F. Given such an F, we will have that ϕ cannot be consistent with stability as it yields a different allocation than A_c^*, which itself yields $\mu_F(c)$, the set of students c is matched to in the unique stable matching.

We first define some notation. Given a density f, for any set of types $\check{\Theta} \subseteq \Theta$, we define the marginal density of agents with score $s_c \in [0, 1]$ at authority c in group $m \in \mathcal{M}$ as:

$$f_{\text{marg}(\check{\Theta})}(s_c, m) = \int_{\check{\Theta}} \mathbb{I}[s_c(\theta) = s_c, m(\theta) = m] \, \mathrm{d}F(\theta) \tag{89}$$

To construct such an F, we proceed in three steps.

1. Take a full-support density f^0 that satisfies the following two conditions: i) Define $\hat{S}_c \in [0, 1]^{|\mathcal{M}|}$ as the cutoff vector that obtains by applying A_c^* to $\tilde{f}$.[2] We assume that f^0 is such that authority c's cutoff vector that is consistent with the unique stable matching, μ_{F_0}, coincides with $\hat{S}_c$; ii) for all $m \in \mathcal{M}$ and

[2]Which exists as any monotone APM admits a cutoff structure (Lemma 13) and the optimal APM is monotone (Theorem 1).

$s_c < \hat{S}_{m,c}$, $f^0_{\text{marg}(\Theta)}(s,m) < \tilde{f}(s,m)$; and iii) all authorities have strictly positive cutoffs for all groups at the unique stable matching.

2. We transform f^0 into a new density f^1 that differs from f^0 on the set of types that is matched with c under μ_{F_0}, which we call Θ_c.[3] We define the scaling factor $\iota^1(s_c, m)$ as:

$$\iota^1(s_c, m) = \frac{\tilde{f}(s_c, m)}{f^0_{\text{marg}(\Theta_c)}(s_c, m)} \tag{90}$$

Given this scaling factor, we define:

$$f^1(\theta) = \begin{cases} f^0(\theta)\iota^1(s_c(\theta), m(\theta)) & \text{if} \quad \theta \in \Theta_c, \\ f^0(\theta) & \text{otherwise.} \end{cases} \tag{91}$$

Observe that this changes the scores of the types who are allocated to c under μ_{F^0} but it does not change their total measure, their composition, or their scores at any other authority. Thus, the unique stable matching under f^1, μ_{F^1}, coincides with μ_{F^0}. Moreover, by assumption i) in the construction of f^0 in step 1, we have that $f^1_{\text{marg}(\Theta_c)}(s_c, m) = \tilde{f}(s_c, m)$ for all m and $s_c \geq \hat{S}_{m,c}$.

3. We transform f^1 into a new density f^2 that differs on the set of unmatched agents under f^0 (and also therefore f^1 by step 2), $\tilde{\Theta}$.[4] Define the set of types who strictly prefer c to their assignment under μ_{F^0} (and also therefore μ_{F^1} by step 2), $\hat{\Theta}_c$.[5] We define a new scaling factor $\iota^2(s_c, m)$ as:

$$\iota^2(s_c, m) = \frac{\tilde{f}(s_c, m) - f_{\text{marg}(\hat{\Theta}_c)}(s_c, m)}{f_{\text{marg}(\tilde{\Theta})}(s_c, m)} \tag{92}$$

which is strictly positive by assumption ii) of step 1. We then define f^2 as:

$$f^2(\theta) = \begin{cases} f^1(\theta)(1 + \iota^2(s_c(\theta), m(\theta))) & \text{if} \quad \theta \in \tilde{\Theta}, \\ f^1(\theta) & \text{otherwise.} \end{cases} \tag{93}$$

By construction, $f^2_{\text{marg}(\hat{\Theta}_c)}(s_c, m) = \tilde{f}(s_c, m)$ for all m and $s_c < \hat{S}_{m,c}$. Moreover, $\mu_{F^2} = \mu_{F^1} = \mu_{F^0}$ as all $\theta \in \tilde{\Theta}$ remain unmatched.

[3] Formally, $\Theta_c = \{\theta : \theta \in D_c(\mu_{F^0}), s_c(\theta) \geq \hat{S}_{m(\theta),c}\}$.

[4] Formally, $\tilde{\Theta} = \{\theta : \theta \in D_c(\mu_{F^1}), s_c(\theta) < \hat{S}_{m(\theta),c}, s_{c'}(\theta) < S^{\mu_{F^1}}_{m(\theta),c'}$ for all $c' \neq c\}$, where $S^{\mu_{F^1}}_{m,c'}$ denotes the group m cutoff at school c' at the stable matching μ_{F^1}, which is strictly positive by assumption iii) of step 2.

[5] Formally, $\hat{\Theta}_c = \{\theta : \theta \in D_c(\mu_{F^1}), s_c(\theta) < \hat{S}_{m(\theta),c}\}$.

We have now constructed a full-support density f^2 with unique stable matching μ_{F^2} (by Theorem 3) such that the density over $D_c(\mu_{F^2})$ coincides with $\tilde{f}$. Moreover, by Claim 9, A_c^* selects $\mu_{F^2}(c)$ from $D_c(\mu_{F^2})$. As ϕ selects a different allocation from $D_c(\mu_{F^2})$ (as it has density of types $\tilde{f}$), it is inconsistent with stability. □

B.1.8 Proof of Theorem 5

Proof. We prove that APM A_c^* implements a dominant strategy for all authorities in all stages by backward induction. Consider the terminal time $t = |\mathcal{C}| - 1$. Some measure of agents λ applies to the authority. Regardless of the measure λ, by Theorem 1 we have that the APM A_c^* is first-best optimal (to see this more concretely, simply index λ by an arbitrary $\omega \in \Omega$ and apply Theorem 1). Thus, A_c^* is dominant. Moreover, from Theorem 1, any strategy that differs from A_c^* on a strictly positive measure set cannot be optimal. Thus any dominant strategy implements essentially the same allocation as A_c^*. Consider now any time $t < |\mathcal{C}| - 1$, precisely the same argument applies and A_c^* is (essentially uniquely) dominant. □

B.1.9 Proof of Proposition 9

Proof. We first prove the following claim.

Claim 12. *μ_{Σ^*} is (almost surely) a deterministic allocation that corresponds to a cutoff matching μ^*.*

Proof. Since there is a continuum of agents, under any Σ^*, with probability 1, any authority c faces a given set of agents who apply Θ_c^{A,Σ^*} with induced measure $\lambda_c^{\Sigma^*}$. As c uses APM A_c^*, with probability 1, any agent θ is admitted to an authority if and only if $s_c(\theta) \geq S_{m,c}^{\Sigma^*}$, where $S_{m,c}^{\Sigma^*}$ denotes the cutoffs when APM A_c^* is applied to agent measure $\lambda_c^{\Sigma^*}$. Since the agents have strict preferences, in any equilibrium, each agent applies to the $\succeq_\theta$ −maximal authority in $\{c : s_c(\theta) \geq S_{m,c}^{\Sigma^*}\}$, and is admitted, which establishes that μ_{Σ^*} is (almost surely) deterministic allocation that corresponds to a cutoff matching with cutoffs $S_{m,c}^{\Sigma^*}$. □

We now establish that μ_{Σ^*} is the unique stable matching of the economy.

Claim 13. *μ^* is the unique stable matching of this economy.*

Proof. For a contradiction, assume μ_{Σ^*} is not stable. Let S denote the unique cutoffs associated with μ_{Σ^*}. Since μ_{Σ^*} is not stable, by Claim 9, S is not a fixed point of T. Let $t_c = T_c(S)$. Since S is not a fixed point of T, there exists $m \in \mathcal{M}$ and $c \in \mathcal{C}$ such that $t_{m,c} \neq S_{m,c}$. Moreover, let $\{x_{m,c}^t\}_{m \in \mathcal{M}}$ and $\{x_{m,c}^s\}_{m \in \mathcal{M}}$ denote the measure of agents in $\tilde{D}_c(S_{-c})$ who are above the admission thresholds for authority c under t_c and S_c. As in Claim 10, note that if there exists m, c such that $t_{m,c} > S_{m,c}$, then from full support, we have that $x^s_{m,c} > x^t_{m,c}$. Since the authority fills its capacity in both cases, there must exists m' such that $x^t_{m',c} > x^s_{m',c}$ which is only possible if $S_{m,c} > t_{m,c}$. By an identical argument, if there is m, c such that $t_{m,c} < S_{m,c}$, then there exists m' $S_{m',c} < t_{m',c}$. Therefore, whenever $t_{m,c} \neq S_{m,c}$, there exists c and m, m' such that $t_{m,c} > S_{m,c}$ and $S_{m',c} > t_{m',c}$. But now we have shown the following:

$$h_c(S_{c,m'})+u_{m'}(x^s_{c,m'}) > h_c(t_{c,m'})+u_{m'}(x^t_{c,m'}) \geq h_c(t_{c,m})+u_m(x^t_{c,m}) > h_c(S_{c,m})+u_m(x^s_{c,m})$$

where the first inequality follows by $t_{c,m'} < S_{c,m'}$, $x^s_{c,m'} < x^t_{c,m'}$, and concavity of u_m. The second inequality follows by optimality. This is because the facts that $t_{c,m} > 0$ and $t'_{c,m'} < 1$ imply that the Lagrange multipliers in the proof of Theorem 1 $\overline{\kappa}_m = \underline{\kappa}_{m'} = 0$. The final inequality follows since $t_{m,c} > S_{m,c}$ and $x^t_{m,c} < x^s_{m,c}$. However, this is a contradiction since $h_c(S_{c,m'}) + u_{m'}(x^s_{c,m'}) > h_c(S_{c,m}) + u_m(x^s_{c,m})$ implies that there exists $\varepsilon > 0$, an agent θ with score $s_c(\theta) = S_{c,m'} - \epsilon$ and type $m(\theta) = m$ has higher score under A^* than the agent θ' with score $s_c(\theta') = S_{c,m}$ and type $m(\theta') = m$. Since θ' is admitted to c, θ would be if it applied to c. Moreover, from full support, there is such θ whose top choice is c and the strategy of this agent is not a best response, which is a contradiction. □

The combination of Claims 12 and 13 completes the proof. □

B.1.10 Proof of Proposition 10

Proof. We prove the result by explicitly constructing an economy in which the optimal APMs lead to inefficiency. There are two authorities, c and c', both with capacity 1/2 and two groups of agents, m and m'. Both agent groups have a measure of 1 and their scores are uniformly distributed in $[1/2, 1]$. Authorities' utility functions are given by

$$\xi_c(\bar{s}_h, x) \equiv \bar{s}_h + \frac{1}{4}\sqrt{x_m} + \frac{1}{8}\sqrt{x_{m'}} \tag{94}$$

$$\xi_{c'}(\bar{s}_h, x) \equiv \bar{s}_h + \frac{1}{4}\sqrt{x_{m'}} + \frac{1}{8}\sqrt{x_m} \tag{95}$$

with $h(x) \equiv x$ while all agents of type m prefer authority c' to c while all agents of type m' prefer authority c to c'.[6]

[6]This assumption on the preferences and the distribution of scores violate our full support assumption, but adding an arbitrarily small full support density to all types makes arbitrarily small changes in the utility under the stable matching and optimal allocation but complicates the calcu-

We will now derive the stable outcome of this economy, which is (up to measure zero transformation) the unique outcome implemented when the authorities use the optimal APM. Let x^c_m denote the measure of type m agents at authority c. First, note that higher-scoring agents from the same group go to the more preferred authority. To see why this is true, note that if $m(\theta) = m(\theta') = m$, $s(\theta) > s(\theta')$ and $\mu(\theta) = c$ while $\mu(\theta') = c'$, c and θ would violate within group fairness since θ has higher priority at c than θ' regardless of the allocation. As a result, in any stable allocation μ, the highest-scoring $x^c_{m'}$ type m' agents are assigned to c and the next highest-scoring $x^{c'}_{m'}$ agents are assigned to authority c', while rest of the type m' agents are not assigned to any authority. The allocation for type m agents is analogous. Moreover, since $q = 1/2$ for both authorities, $x^{c'}_{m'}$=1/2-$x^c_{m'}$ and x^c_m=1/2-$x^{c'}_m$ and the allocation is completely determined by the measures $x^c_{m'}$ and $x^{c'}_m$.

Next, note that at μ, the adaptive priority of the lowest-scoring type m and m' agents must be equal at both authorities. To see why this is true, take authority c without loss of generality. Let $s^c_{m'} = 1 - x^c_{m'}$ and $s^c_m = 1 - x^c_{m'} - x^c_m$ denote the scores of the lowest-scoring type m and m' agents and A_m denote the optimal APM. For a contradiction, assume $A_m(x^c_m, s^c_m) > A_{m'}(x^c_{m'}, s^c_{m'})$. Since agents of type m' with scores lower than s^c_m are unassigned at μ, for small enough ϵ, a type m agent with score $s^c_m - \epsilon$ and authority c blocks the matching. Similarly, assume $A_m(x^c_m, s^c_m) < A_{m'}(x^c_{m'}, s^c_{m'})$. Since agents of type m' with scores lower than $s^{c'}_m$ are assigned to authority c or unmatched at μ, a type m' agent with score $s^c_{m'} - \epsilon$ and authority c blocks the matching μ. Thus, the following equations must be satisfied:

$$A_m(x^c_m, s^c_m) = A_{m'}(x^c_{m'}, s^c_{m'}) \text{ and } A_m(x^{c'}_m, s^{c'}_m) = A_{m'}(x^{c'}_{m'}, s^{c'}_{m'}) \tag{96}$$

As the optimal APM in this setting is given by:

$$A^*_{\hat{m},\hat{c}}(y_{\hat{m}}, s) \equiv s + u'_{\hat{m},\hat{c}}(y_{\hat{m}}) \tag{97}$$

for all $\hat{m} \in \{m, m'\}$ and $\hat{c} \in \{c, c'\}$, we have that:

$$1 - x^c_{m'} + \frac{1}{8}\frac{1}{\sqrt{x^c_{m'}}} = 1 - x^c_{m'} - x^c_m + \frac{1}{4}\frac{1}{\sqrt{1/2 - x^c_{m'}}} \tag{98}$$

and:

$$1 - x^{c'}_m + \frac{1}{8}\frac{1}{\sqrt{x^{c'}_m}} = 1 - x^{c'}_m - x^{c'}_{m'} + \frac{1}{4}\frac{1}{\sqrt{1/2 - x^{c'}_m}} \tag{99}$$

lation, so we omit it for expositional clarity.

These equations are identical up to relabelling and so $x^{c}_{m'} = x^{c'}_{m'} = x^*$ for some x^*. Thus, we need to find the solution to the following single equation to characterize the allocation:

$$1 - x^* + \frac{1}{8}\frac{1}{\sqrt{x^*}} = \frac{1}{2} + \frac{1}{4}\frac{1}{\sqrt{1/2 - x^*}} \tag{100}$$

Observe that this equation can be rewritten as the fixed point equation:

$$x^* = \frac{1}{2} + \frac{1}{8}\frac{1}{\sqrt{x^*}} - \frac{1}{4}\frac{1}{\sqrt{1/2 - x^*}} \tag{101}$$

We observe that the RHS satisfies the following properties: (i) $\lim_{x^* \to 0} \mathrm{RHS}(x^*) = \infty$, (ii) $\lim_{x^* \to \frac{1}{2}} \mathrm{RHS}(x^*) = -\infty$, and (iii) $\mathrm{RHS}'(x^*) < 0$ for all $x^* \in (0, \frac{1}{2})$. Thus, there exists a unique solution. Moreover, we can guess-and-verify that this solution is $x^* = \frac{1}{4}$.

In summary, if both authorities use the optimal APM, then the outcome is

$$\mu(\theta) = \begin{cases} c & \text{if } m(\theta) = m, s(\theta) \in [1/2, 3/4) \text{ or } m(\theta) = m', s(\theta) \in [3/4, 1] \\ c' & \text{if } m(\theta) = m', s(\theta) \in [1/2, 3/4) \text{ or } m(\theta) = m, s(\theta) \in [3/4, 1] \\ \theta & \text{otherwise} \end{cases} \tag{102}$$

In this outcome, both authorities have an average score of 3/4 and admit measure 1/4 agents from both groups, giving them a utility of 15/16. Thus, total utilitarian welfare is 15/8 under the decentralized outcome.

We now show that this does not attain the efficient benchmark. A necessary condition for the (utilitarian) efficient outcome is that for c:

$$\frac{1}{4}\frac{1}{\sqrt{x^c_m}} = \frac{1}{8}\frac{1}{\sqrt{1/2 - x^c_m}} \tag{103}$$

and for c':

$$\frac{1}{4}\frac{1}{\sqrt{x^{c'}_{m'}}} = \frac{1}{8}\frac{1}{\sqrt{1/2 - x^{c'}_{m'}}} \tag{104}$$

This implies that $x^c_m = x^{c'}_{m'} = 4/10$ and $x^c_{m'} = x^{c'}_m = 1/10$. In this case, the same set of agents is admitted overall, so the score contribution to utility remains 3/4 on average across the authorities. Total utilitarian welfare is now:

$$3/2 + 1/2 \times \sqrt{4/10} + 1/4 \times \sqrt{1/10} \approx 1.895 > 1.875 = 15/16 \tag{105}$$

Completing the proof. □

B.1.11 Proof of Theorem 6

Proof. First, we define a fictitious *composite authority* with utility function defined over vectors of total scores $\bar{s}_h = \{\bar{s}_h^c\}_{c\in\mathcal{C}}$, and aggregate allocation to each group $x = \{x_m\}_{m\in\mathcal{M}}$. To do this, we define:

$$\begin{aligned}\tilde{u}(\{x_m\}_{m\in\mathcal{M}}) = &\max_{\{x_{m,c}\}_{c\in\mathcal{C}}} \sum_{c\in\mathcal{C}} \sum_{m\in\mathcal{M}} u_{m,c}(x_{m,c}) \\ &\text{s.t.} \sum_{c\in\mathcal{C}} x_{m,c} \leq x_m, \sum_{m\in\mathcal{M}} x_{m,c} \leq q_c, \forall m \in \mathcal{M}, c \in \mathcal{C}\end{aligned} \tag{106}$$

and $\tilde{\bar{s}}_h = \sum_{c\in\mathcal{C}} \bar{s}_h^c$. We write the utility function of this composite authority as

$$\tilde{\xi}\left(\tilde{\bar{s}}_h, x\right) = \tilde{\bar{s}}_h + \tilde{u}(x) \tag{107}$$

We first establish that $\tilde{u}$ satisfies the properties necessary to invoke Proposition 20, which establishes the optimality of the claimed APM for the fictitious authority.

Claim 14. *The function $\tilde{u}$ is concave and partially differentiable in each argument.*

Proof. First, we establish concavity. That is, for all $\lambda \in [0,1]$ and $x, x' \in \mathbb{R}_+^{|\mathcal{M}|}$, we have that $\tilde{u}(\lambda x' + (1-\lambda)x) \geq \lambda\tilde{u}(x') + (1-\lambda)\tilde{u}(x)$. Let $\{x_{m,c}^*\}_{m\in\mathcal{M},c\in\mathcal{C}}$ and $\{x_{m,c}^{*'}\}_{m\in\mathcal{M},c\in\mathcal{C}}$ correspond to optimal values under x and x'. Under $\tilde{x} = \lambda x' + (1-\lambda)x$, we have that $\lambda x_{m,c}^{*'} + (1-\lambda)x_{m,c}^*$ is feasible for all $m \in \mathcal{M}$ and $c \in \mathcal{C}$. Thus, we have that:

$$\begin{aligned}\tilde{u}(\tilde{x}) &\geq \sum_{m\in\mathcal{M}} \sum_{c\in\mathcal{C}} u_{m,c}(\lambda x_{m,c}^{*'} + (1-\lambda)x_{m,c}^*) \\ &\geq \sum_{m\in\mathcal{M}} \sum_{c\in\mathcal{C}} \lambda u_{m,c}(x_{m,c}^{*'}) + (1-\lambda)u_{m,c}(x_{m,c}^*) \\ &= \lambda\tilde{u}(x') + (1-\lambda)\tilde{u}(x)\end{aligned} \tag{108}$$

where the second inequality is by concavity of $u_{m,c}$ for all $m \in \mathcal{M}, c \in \mathcal{C}$.

Second, we establish partial differentiability in each argument. That is, for all $x \in \mathbb{R}_{++}^{|\mathcal{M}|}$, $\frac{\partial}{\partial x_m}\tilde{u}(x) = \tilde{u}^{(m)}(x)$ exists. This follows by Corollary 5 in Milgrom and Segal (2002). Concretely, the domain of optimization can be taken to be a compact and convex subset of a normed vector space – a sufficiently large cube in $\mathbb{R}_+^{|\mathcal{M}|\times|\mathcal{C}|}$ equipped with the Euclidean norm, for example. The objective function does not depend on x, and constraints are linear in x (and therefore both continuous and continuously differentiable). Moreover, as $x \gg 0$, there exists a $\{x_{m,c}\}$ that satisfies all constraints with strict inequality.

□

It follows that the objective function of the composite authority satisfies Assumption 3, and so Proposition 20 implies that the non-separable APM $\tilde{A}_m(y, s) = h^{-1}\left(h(s) + \tilde{u}^{(m)}(y)\right)$ uniquely implements the first-best optimal allocation for the composite authority.

It remains to establish that the quota functions implement the optimal allocation $\{x_{m,c}\}$ conditional on $\{x_m\}$. Let λ_m be the Lagrange multiplier on the x_m constraint, γ_c be the Lagrange multiplier on the q_c constraint and $\underline{\kappa}_{m,c}$ be the Lagrange multiplier on the positivity constraint. Under our maintained Inada condition, we have that $\underline{\kappa}_{m,c} = 0$. Moreover, by Corollary 5 in Milgrom and Segal (2002), we have that $\tilde{u}^{(m)}(x) = \lambda_m$, $\tilde{u}_{q_c}(x) = \gamma_c$, and $u'_{m,c}(x^*_{m,c}) = \lambda_m + \gamma_c - \underline{\kappa}_{m,c}$. Hence, we obtain that:

$$x^*_{m,c} = \left(u'_{m,c}\right)^{-1}\left(\tilde{u}^{(m)}(x) + \tilde{u}_{q_c}(x)\right) \tag{109}$$

Thus, the following profile of quota functions implements the optimal cross-sectional allocation:

$$Q_{m,c}(x) = \left(u'_{m,c}\right)^{-1}\left(\tilde{u}^{(m)}(x) + \tilde{u}_{q_c}(x)\right) \tag{110}$$

Completing the proof. □

B.2 Additional Results for the Example (Section 2.2)

B.2.1 Formal Equivalence Between Prices *vs.* Quantities and Priorities *vs.* Quotas

The structure of the comparative advantage of priorities over quotas from Section 2.2 hints at a more formal relationship between our analysis of affirmative action policies and Weitzman's analysis of price and quantity regulation. In Weitzman's model, there is a single firm producing a quantity of a single good $x \in \mathbb{R}$ with production costs $C(x, \zeta)$ and benefits $B(x, \zeta')$:

$$\begin{aligned} C(x,\zeta) &= a_0(\zeta) + (C' + a_1(\zeta))(x - \hat{x}) + \frac{C''}{2}(x-\hat{x})^2 \\ B(x,\zeta') &= b_0(\zeta') + (B' + b_1(\zeta'))(x - \hat{x}) + \frac{B''}{2}(x-\hat{x})^2 \end{aligned} \tag{111}$$

where $B', C', C'' > 0$, $B'' < 0$, and ζ and ζ' are random variables. The regulator can either set a price that the firm must charge (after which the firm chooses its optimal production quantity) or mandate the production of a given quantity. The comparative advantage of prices over quantities Δ^{Weitzman} is then defined as the difference between expected benefits net of costs under the optimal price regime minus the corresponding net benefits under the optimal quantity regime. This comparative advantage is given by:

$$\Delta^{\text{Weitzman}} = \frac{C''^{-1}}{2}\left(1 + C''^{-1}B''\right)\text{Var}[a_1(\zeta)] \tag{112}$$

The intuition for this formula is that when benefits are more curved than costs $|B''| > C''$, reducing variability in production is more valuable than the gain of having producers minimize costs. Thus, quantities are preferred. On the other hand, when costs are more curved than benefits, prices are preferred as there is greater production when producers have the lowest marginal costs of production.

These trade-offs are, in a certain sense, formally analogous to those that we have highlighted between priorities and quotas. In particular, under the mapping $C''^{-1} \mapsto \kappa$, $B'' \mapsto -\gamma\beta$, $\text{Var}[\omega] \mapsto \text{Var}[a_1(\zeta)]$, we have that $\Delta^{\text{Weitzman}} = \Delta$. The intuition for this is that C''^{-1} in the Weitzman framework determines how sensitive production is to changes in marginal cost, while κ in our framework determines how sensitive the admitted measure of minority students is to the relative scores. Moreover, B'' corresponds to curvature in the benefits of production while $\gamma\beta$ corresponds to curvature in the benefits of admitting more minority students. Finally, $\text{Var}[a_1(\zeta)]$

corresponds to the authority's uncertainty in the level of marginal costs of production while Var[ω] corresponds to the authority's uncertainty regarding the marginal cost of admitting more minority students in terms of lost total score. Thus, the positive selection effect whereby priorities admit more minority students in the states of the world where they score highest is directly analogous to the effect that price regulation gives rise to the greatest production in states where the firm's marginal cost is lowest. Moreover, the guarantee effect whereby quotas prevent variation in the measure of admitted minority students across states of the world is analogous to the ability of quantity regulation to stabilize the level of production. Importantly, our results therefore allow one to apply established price-theoretic intuition for the benefits of price *vs* quantity choice to matching markets without an explicit price mechanism.

B.2.2 Beyond Affirmative Action: Medical Resource Allocation

The lessons of this paper apply not only to affirmative action in academic admissions, but also more broadly to settings in which centralized authorities must allocate resources to various groups. One prominent such context is the allocation of medical resources during the Covid-19 pandemic. An important issue faced by hospitals is how to prioritize health workers (doctors, nurses and other staff) in the receipt of scarce medical resources: hospitals wish to both treat patients according to clinical need and ensure the health of the frontline workers needed to fight the pandemic. To map this setting to our example, suppose that the score s is an index of clinical need for a scarce medical resource available in amount q, the measure of frontline health workers is κ, and ω indexes the level of clinical need in the patients currently (or soon to be) treated by the hospital, which is unknown. The risk aversion of the authority $\gamma\beta$ corresponds to both a fear of not treating sufficiently many frontline workers and excluding too many clinically needy members of the general population.

In practice, both priority systems and quota policies have been used, as detailed extensively by Pathak, Sönmez, Ünver, and Yenmez (2021).[7] The primary concern that has been voiced is that if a priority system is used, some groups (or characteristics) may be completely shut out of allocation of the scarce resource and that this is unethical, so quotas should be preferred. Our framework can be used to understand this argument: if there is an unusually high draw of ω, a priority system would lead to the allocation of very few resources to frontline workers, and vice-versa. Our

[7] Some other papers that study the allocation of scarce medical resources are Akbarpour, Budish, Dworczak, and Kominers (2021), Grigoryan (2021) and Dur, Thayer, and Phan (2022).

Proposition 6 implies that if the authority is very averse to such outcomes ($\gamma\beta$ is high), quotas will be preferred and for exactly the reasons suggested. However, we also highlight a fundamental benefit of priority systems in inducing positive selection in allocation: when ω is high, it is beneficial that fewer resources go to the less sick medical workers and more to the relatively sicker general population. More generally, we argue that an adaptive priority mechanism that awards frontline workers a score subsidy that depends on the number of more clinically needy frontline workers could further improve outcomes.

An important additional consideration in this context arises if the hospital or authority must select a regime (priorities or quotas) before it understands the clinical need of its frontline workers κ, after which it can decide exactly how to prioritize these workers or set quotas, but before ultimate demand for medical resources ω is known. It follows from Proposition 6 that the comparative advantage of priorities over quotas is:

$$\mathbb{E}[\Delta] = \frac{1}{2}\left(\mathbb{E}[\kappa] - (\mathrm{Var}[\kappa] + \mathbb{E}^2[\kappa])\gamma\beta\right)\mathrm{Var}[\omega] \tag{113}$$

Thus, an increase in uncertainty $\mathrm{Var}[\kappa]$ regarding the need of frontline workers leads to a greater preference for quotas. This highlights a further advantage of quotas in settings where a clinical framework must be adopted in the face of uncertainty regarding the clinical needs of frontline workers, as was the case at the onset of the Covid-19 pandemic.

B.2.3 Optimal Precedence Orders

Thus far we have modelled quotas by first allocating minority students to quota slots and then allocating all remaining students according to the underlying score. However, we could have instead allocated $q - Q$ places to all agents according to the underlying score and then allocated the remaining Q places to minority students. The order in which quotas are processed is called the *precedence order* in the matching literature and their importance for driving outcomes has been the subject of a growing literature (see *e.g.,* Dur, Kominers, Pathak, and Sönmez, 2018; Dur, Pathak, and Sönmez, 2020; Pathak, Rees-Jones, and Sönmez, 2020). Our framework can be used to understand which precedence order is optimal, a question that has not yet been addressed.

In this example, the same factors that determine whether one should prefer priorities or quotas determine whether one should prefer processing quotas second or first. By virtue of uniformity of scores, it can be shown in the relevant parameter range that a priority subsidy of α is equivalent to a quota policy of $\kappa\alpha$ when quotas are processed second. Thus, the comparative advantage of priorities over quotas is exactly equal to the comparative advantage of processing quotas second over first. The intuition is analogous: processing quotas second allows for positive selection while processing quotas first fixes the number of admitted minority students. Thus, on the one hand, when the authority is more risk-averse, they should process quota slots first to reduce the variability in the admitted measure of minority students. On the other hand, when they are less risk-averse, they should process quotas second to take advantage of the positive selection effect such policies induce.

Corollary 1. *The optimal quota-second policy achieves the same value as the optimal priority policy; quota-second policies are preferred to quota-first policies if and only if* $\frac{1}{\kappa} \geq \gamma\beta$.

Proof. We show that a quota-second policy Q is equivalent to a priority subsidy of $\alpha(Q) = \frac{Q}{\kappa}$. A quota-second policy admits the highest-scoring $x = \kappa(1 - \omega) + Q$ minority students, floored by zero and capped by $\min\{\kappa, q\}$. A priority policy $\alpha(Q) = \frac{Q}{\kappa}$ admits the highest-scoring $x = \kappa(1 + \alpha(Q) - \omega) = \kappa(1 - \omega) + Q$ minority students, floored by zero and capped by $\min\{\kappa, q\}$. Thus, state-by-state, quota-second policy Q and priority subsidy $\alpha(Q) = \frac{Q}{\kappa}$ yield the same allocation. The claims then follow from Proposition 6. □

We emphasize that this equivalence is a result of the uniform distribution of scores and merely illustrates the similarity between priority policies and processing quotas second. This result does not hold in the more general model we study in the remainder of the paper. Indeed, in Theorem 2, we show that for any quota policy to be optimal in the presence of uncertainty, it must process quotas first.

B.3 Extension of the Main Results to Discrete Economies

In this Appendix, we extend the results in the main text to discrete economies and thereby establish that our analysis generalizes from the continuum framework. Concretely, we show that appropriate analogs of Theorems 1, 2, 5 and Proposition 8 carry over to discrete economies. As discrete economies do not necessarily admit a unique stable matching, Theorems 3 and 4 do not generalize. As Theorem 6 relies on convex optimization techniques, it also does not generalize as written to a discrete setting.

B.3.1 Primitives

An authority has q resources to allocate. At each state ω, the economy the authority faces corresponds to agents $\Theta^{\omega} = \{\theta_1, \ldots, \theta_{N(\omega)}\}$ where $q \leq |N(\omega)|$. As in the continuum case, $\theta \in [0,1] \times \mathcal{M}$ denotes the type of an agent who has score s and belongs to group m. We denote the score and group of any type θ by $s(\theta)$ and $m(\theta)$, respectively. For simplicity, we assume that no two agents have the same score at any ω, formally, if $\{\theta, \theta'\} \subseteq \Theta^{\omega}$, then $s(\theta) \neq s(\theta')$.

An allocation $\mu : \Theta \to \{0,1\}$ specifies for any type $\theta \in \Theta$ whether they are assigned to the resource. The set of possible allocations is $\mathcal{U}$ and Ω is the set of all possible economies. An allocation is feasible if it allocates no more than measure q of the resource. A mechanism is a function $\phi : \Omega \to \mathcal{U}$ that returns a feasible allocation for any possible Θ^{ω}.

The authority believes ω has distribution $\Lambda \in \Delta(\Omega)$. $x(\mu, \omega) = \{x_m(\mu, \omega)\}_{m \in \mathcal{M}}$ denotes the number of agents of each group allocated the resource at matching μ, while $\bar{s}_h(\mu, \omega) = \sum_{\theta \in \Theta^{\omega}} \mu(\theta) h(s(\theta))$ denotes the utility the authority derives from scores at μ. The preferences of the authority are given by $\xi : \mathbb{R}^{|\mathcal{M}|+1} \to \mathbb{R}$:

$$\xi\left(\bar{s}_h, x\right) \equiv \bar{s}_h + \sum_{m \in \mathcal{M}} u_m(x_m) \tag{114}$$

where h is continuous and strictly increasing and u_m is concave for all $m \in \mathcal{M}$.

B.3.2 Optimal Mechanisms in Discrete Economies

We adapt our definition of the Adaptive Priority Mechanisms to the discrete setting. An *adaptive priority policy* $A = \{A_m\}_{m \in \mathcal{M}}$, where $A_m : \mathbb{R} \times [0,1] \to \mathbb{R}$. The adaptive priority policy assigns priority $A_m(y_m, s)$ to an agent with score s in group m when y_m of agents of the same group is allocated the object. Given an adaptive priority policy, an APM implements allocations in the following way:

Definition 20 (Adaptive Priority Mechanism). *An adaptive priority mechanism, induced by an adaptive priority A, implements an allocation μ in state ω if the following are satisfied:*

1. *Allocations are in order of priorities: $\mu(\theta) = 1$ if and only if*

   (i) for all θ' with $m(\theta') \neq m(\theta)$ and $\mu(\theta') = 0$,

$$A_{m(\theta)}(x_{m(\theta)}(\mu, \omega), s(\theta)) \geq A_{m(\theta')}(x_{m(\theta')}(\mu, \omega) + 1, s(\theta')) \tag{115}$$

   (ii) for all θ' with $m(\theta') = m(\theta)$ and $\mu(\theta') = 0$, $s(\theta) > s(\theta')$

2. *The resource is fully allocated:*

$$\sum_{m \in \mathcal{M}} x_m(\mu, \omega) = q \tag{116}$$

Definition 20 makes two modifications relative to the continuum model. First, the measures of agents from each group are replaced by the number of agents from each group. Second, when $m(\theta) \neq m(\theta')$, the adaptive priority of θ' is now evaluated in the case where an extra agent from $m(\theta')$ is assigned the resource.[8] Unlike the continuum case, it is possible for a monotone APM to implement two different allocations, since it can assign the same priority to two different agents, which could happen only for a zero-measure set of agents in the continuum model.

Define $A^*_m(y_m, s) \equiv h(s) + u_m(y_m) - u_m(y_m - 1)$, which will turn out to be the optimal APM. We first show that A^* is monotone, and all allocations that A^* implements give the authority the same utility.

Lemma 14. *A^* is monotone. Moreover, if A^* implements μ and $\mu' \neq \mu$ in state ω, then $\xi(\mu, \omega) = \xi(\mu', \omega)$.*

Proof. Monotonicity is immediate from the definition of A^* and concavity of u_m. Assume that A^* implements two different allocations, μ and μ' at ω. Let x_l and x'_l denote the number of group $l \in \mathcal{M}$ agents assigned the resource at μ and μ'. Since A^* is monotone and $\mu \neq \mu'$, there are m and n such that $x_m > x'_m$ and $x'_n > x_n$. Let $\tilde{\theta}_l$ and $\tilde{\theta}'_l$ denote the lowest-scoring type l agent assigned the resource at μ and μ', respectively. Similarly, let $\hat{\theta}_l$ and $\hat{\theta}'_l$ denote the highest-scoring type l agents who is not assigned the resource at μ and μ', respectively. Let $\tilde{\mu}$ denote the matching given

[8] This was not the case in the continuum model since all types of agents have measure 0 and therefore replacing θ with θ' has no effect the evaluation of diversity.

by: $\tilde{\mu}(\theta) = \mu(\theta)$ if $\theta \notin \{\tilde{\theta}_m, \hat{\theta}'_n\}$, $\tilde{\mu}(\tilde{\theta}_m) = 0$ while $\mu(\hat{\theta}_n) = 1$. $\tilde{\mu}$ starts with μ, takes the resource away from the lowest-scoring group m agent who has it, $\tilde{\theta}_m$, and allocates it to the highest-scoring group n agent who does not have it, $\hat{\theta}_n$. Note that since A^* is monotone, from $x_m > x'_m$ and $x'_n > x_n$, under μ', $\hat{\theta}_n$ is already allocated the resource while $\tilde{\theta}_m$ is not.

Claim 15. *$\tilde{\mu}$ is implemented under A^* in state ω and $\xi(\mu, \omega) = \xi(\tilde{\mu}, \omega)$.*

Proof. Since A^* implements μ and $\mu(\hat{\theta}_n) = 0$, we have that $A^*_m(s(\tilde{\theta}_m), x_m) \geq A^*_n(s(\hat{\theta}_n), x_n + 1)$. Conversely, since A^* also implements μ' and $\mu'(\hat{\theta}'_m) = 0$, we have that $A^*_n(s(\tilde{\theta}'_n), x'_n) \geq A^*_m(s(\hat{\theta}'_m), x'_m + 1)$. Moreover, since $x_m > x'_m$ and $x'_n > x_n$, we have that $s(\hat{\theta}'_m) \geq s(\tilde{\theta}_m)$ and $s(\hat{\theta}_n) \geq s(\tilde{\theta}'_n)$. From this, it follows that:

$$\begin{aligned} A^*_n(s(\hat{\theta}_n), x_n + 1) &\geq A^*_n(s(\tilde{\theta}'_n), x_n + 1) \geq A^*_n(s(\tilde{\theta}'_n), x'_n) \\ &\geq A^*_m(s(\hat{\theta}'_m), x'_m + 1) \geq A^*_m(s(\hat{\theta}'_m), x_m) \geq A^*_m(s(\tilde{\theta}_m), x_m) \end{aligned} \tag{117}$$

where the first inequality holds as $s(\hat{\theta}_n) \geq s(\tilde{\theta}'_n)$, the second inequality holds as $x'_n > x_n$ (which implies $x'_n \geq x_n + 1$) and A^*_n is decreasing in its second argument, the third inequality holds as A^* also implements μ' (as stated above), the fourth inequality holds as $x'_m < x_m$ (which implies $x'_m + 1 \leq x_m$) and A^*_n is decreasing in its second argument, and the fifth inequality holds as $s(\hat{\theta}'_m) \geq s(\tilde{\theta}_m)$. Thus, $A^*_m(s(\tilde{\theta}_m), x_m) \leq A^*_n(s(\hat{\theta}_n), x_n + 1)$. This shows that $A^*_m(s(\tilde{\theta}_m), x_m) = A^*_n(s(\hat{\theta}_n), x_n + 1)$, which implies that $\tilde{\mu}$ is implemented under A^* and $\xi(\mu, \omega) = \xi(\tilde{\mu}, \omega)$. ☐

Note that Claim 15 shows that starting from a matching μ which is implemented by A^*, taking away the object from a particular agent who does not have it in μ' and allocating it to a particular agent who has it in μ', we arrive at another matching $\tilde{\mu}$ that is implemented under A^* and gives the authority the same payoff. Therefore, starting from any μ that is implemented by A^* and repeating this construction (by replacing μ at step i with $\tilde{\mu}$ at step $i-1$) where at each step we take the resource from an agent who is not allocated the resource at μ' and assign it to an agent who is, in finitely many steps we arrive at μ'. Since the payoff stays the same at each step, μ' gives the authority the same payoff as μ. ☐

Theorem 11. *If μ is implemented by A^*, then μ is an optimal matching.*

Proof. First, note that an optimal matching exists since the economy (and therefore the set of matchings) is finite. We first show the following lemma.

Lemma 15. *If μ is not implemented by A^*, then there exists μ' that gives the authority a strictly higher payoff.*

Proof. If μ is not implemented by A^*, then there exists θ and θ' such that $\mu(\theta) = 0$, $\mu(\theta') = 1$ and either $m(\theta) = m(\theta')$ and $s(\theta) > s(\theta')$ or $m(\theta) \neq m(\theta')$ and

$$\begin{aligned} &h(s(\theta)) + u_{m(\theta)}(x_{m(\theta)}(\mu) + 1) - u_{m(\theta)}(x_{m(\theta)}(\mu)) > \\ &h(s(\theta')) + u_{m(\theta')}(x_{m(\theta')}(\mu)) - u_{m(\theta')}(x_{m(\theta')}(\mu) - 1) \end{aligned} \tag{118}$$

However, in both cases, a μ' that allocates the resource to θ instead of θ' (while not changing any other agent's matching) strictly improves the utility of the authority. □

Lemma 15 proves that the optimal matching cannot be a matching that is not implemented by A^*. Since the optimal matching exists, then it is implemented by A^*. From Lemma 14, all matchings implemented by A^* give the authority the same payoff, proving the result. □

Note that Lemma 14 and Theorem 11 imply that any mechanism that is defined by an arbitrary singleton selection from the set of matchings that A^* implements would achieve the optimal matching under any ω and therefore would be first-best optimal.

B.3.3 Priorities *vs.* Quotas in Discrete Economies

Now, we define Priority and Quota Mechanisms in the discrete model and extend our (sub)optimality results to discrete economies.

A *priority policy* $P : \Theta \to [0, 1]$ awards an agent of type $\theta \in \Theta$ a priority $P(\theta)$.

Definition 21 (Priority Mechanisms)**.** *A priority mechanism, induced by a priority policy P, allocates the resource in order of priorities until measure q has been allocated, with ties broken uniformly and at random.*

A *quota policy* is given by (Q, D), where $Q = \{Q_m\}_{m\in\mathcal{M}}$ and $D : \mathcal{M} \cup \{R\} \to \{1, 2, \ldots, |\mathcal{M}| + 1\}$ is a bijection. The vector Q reserves Q_m objects for agents in group m, with residual capacity $Q_R = q - \sum_{m\in\mathcal{M}} Q_m$ open to agents of all types. The bijection D (often called the precedence order) determines the order in which the groups are processed.

Definition 22 (Quota Mechanisms)**.** *A quota mechanism, induced by a quota policy (Q, D), proceeds by allocating $Q_{D^{-1}(k)}$ objects to agents from group $D^{-1}(k)$ (if there are sufficient agents from this group) to the resource in ascending order of k, and in* *descending order of score within each k. If there are insufficiently many agents of any group to fill the quota, the residual capacity is allocated to a final round in which all agents are eligible.*

We also extend the definitions of risk-neutrality and high risk aversion to the discrete setting. Authority preferences are *non-trivial* if for all $m, n \in \mathcal{M}$:

$$h(1) + (u_n(1) - u_n(0)) > h(0) + (u_m(q) - u_m(q-1)) \tag{119}$$

The authority is *risk-neutral* if for all $m \in \mathcal{M}$, $u_m(x) = c_m x$ for some $c_m \geq 0$ and all $x \in \{0, 1, \ldots, q\}$. Define $\tilde{u}$ and $\tilde{h}$ as follows: there exists x_m^{tar} such that $\tilde{u}_m(x_m + 1) - \tilde{u}_m(x_m) = 0$ for all $x_m \geq x_m^{\text{tar}}$ and $\tilde{u}_m(x_m + 1) - \tilde{u}_m(x_m) \geq h(1) - h(0)$ for $x_m < x_m^{\text{tar}}$ and where $\sum_{m \in \mathcal{M}} x_m^{\text{tar}} \leq q$. Let $\tilde{\xi}$ denote the preferences of the authority under $\tilde{u}$ and $\tilde{h}$. The authority with preferences ξ is *extremely risk-averse* if the set of optimal allocations under ξ and $\tilde{\xi}$ coincide for all ω.

Theorem 12. *The following statements are true:*

1. *If there is no uncertainty, then there exist first-best priority and quota mechanisms.*

2. *Suppose that the authority has non-trivial preferences. There exists a first-best priority mechanism if and only if the authority is risk-neutral. This mechanism is given by $P(s, m) = s + u_m(1) - u_m(0)$.*

3. *Suppose that the authority has non-trivial preferences. There exists a first-best quota mechanism if and only if the authority is extremely risk-averse. This mechanism is given by $Q_m = x_m^{tar}$ and $D(R) = |\mathcal{M}| + 1$.*

Proof. Part (1):

Claim 16. *Let μ denote an optimal allocation at ω. Then μ is a cutoff matching.*

Proof. If μ is not a cutoff matching, then there exists (s, m) and (s', m) where $\mu(s, m) = 1$, $\mu(s', m) = 0$ and $s' > s$. Define μ' by setting: $\mu'(s, m) = 0$, $\mu(s', m) = 1$ and $\mu(\tilde{s}, \tilde{m}) = \mu(\tilde{s}, \tilde{m})$ for all $(\tilde{s}, \tilde{m})$ such that $(\tilde{s}, \tilde{m}) \notin \{(s, m), (s', m)\}$. Observe that, $\xi(\mu', \omega) - \xi(\mu, \omega) = s' - s > 0$. Therefore, μ is not an optimal allocation, which is a contradiction. □

Let μ denote an optimal allocation under ω, $\{\hat{s}_m(\mu, \omega)\}_{m \in \mathcal{M}}$ denote the cutoff scores at μ and s^* denote an arbitrary number. Any priority policy that assigns

$P(\hat{s}_m(\omega), m) = s^*$ for all $m \in \mathcal{M}$ and is strictly increasing in the first argument allocates the resource to any agent who has a higher score than the cutoff for their group and implements the optimal allocation.

Let x_m denote the number of group m agents who are allocated the resource at an optimal allocation under ω. Then a quota policy that sets $Q_m = x_m$ allocates the resource to any agent who has a higher score than the cutoff for their group and implements the optimal allocation.

Part (2): The if part of the result follows from observing the priority policy $P(s, m) = s + u_m(1) - u_m(0)$ is equivalent to the optimal APM A^* under risk neutrality since $u_m(1) - u_m(0) = u_m(y_m + 1) - u_m(y_m)$ for all m, y_m. Thus, by Theorem 11, $P(s, m) = s + u_m(1) - u_m(0)$ is first-best optimal.

To prove the only if part, assume risk neutrality does not hold and let m denote a group such that u_m does not satisfy risk neutrality. For a contradiction, assume that P is an optimal priority policy. First, we observe that $P(s, m)$ must be strictly increasing in s for all m. To see why, assume $P(s, m) = P(s', m)$ where $s > s'$ and just consider an ω where there are $q - 1$ group m agents with scores strictly higher than s, and no other agents. Clearly, the optimal allocation would be to allocate the resource to all agents but (s', m), while P allocates the resource to (s', m) with at least probability 1/2.

Second, let m denote a group such that u_m does not satisfy risk neutrality. Take another arbitrary group n. We have the following:

Claim 17. *Either (i) there exists $t < q$, s_m, s_n such that*

$$u_m(t + 1) - u_m(t) + h(s_m) = u_n(q - t) - u_n(q - t - 1) + h(s_n) \tag{120}$$

or (ii) there exists $t < q$ such that

$$u_m(t + 1) - u_m(t) + h(1) < u_n(q - t) - u_n(q - t - 1) + h(0) \tag{121}$$

$$u_m(t) - u_m(t - 1) + h(0) > u_n(q - t + 1) - u_n(q - t) + h(1) \tag{122}$$

Proof. From non-triviality, we know that $u_m(1) - u_m(0) + h(1) > u_n(q) - u_n(q - 1) + h(0)$ and $u_n(1) - u_n(0) + h(1) > u_m(q) - u_m(q - 1) + h(0)$. The result then follows from the fact that h is continuous and strictly increasing and u_m and u_n are concave. □

We first prove the result under case (ii). Fix two agents with scores $s_m \in (0, 1)$, who belong to group m and $s_n \in (0, 1)$, who belong to group n. Assume that there are

$t-1$ group m agents and $q-t$ group n agents with higher scores than $\max\{s_n, s_m\}$, so a total of t group m agents and $q-t+1$ group n agents. Note that in this case, only one agent will not be allocated the resource in the optimal allocation, and that would be either (s_m, m) or (s_n, n). From equation 122, (s_m, m) is more preferred than (s_n, n) and therefore it must be that $P(s_n, n) < P(s_m, m)$, as otherwise P would not be optimal. Next, assume that there are t group m agents and $q-t-1$ group n agents with higher scores than $\max\{s_n, s_m\}$. From equation 121, (s_n, n) is more preferred than (s_m, m) and therefore it must be that $P(s_m, m) < P(s_n, n)$, which is a contradiction.

We now prove the result under case (i).

Claim 18. *In case (i), any optimal priority policy P must satisfy $P(s_m + \epsilon, m) > P(s_n, n)$ for all $\epsilon > 0$ and $P(s_m - \epsilon, m) < P(s_n, n)$ for all $\epsilon > 0$*

Proof. From Equation 120, we see that when there are t group m agents and $q-t-1$ group n agents with higher scores, $(s_m + \epsilon, m)$ is strictly preferred to (s_n, n), which is strictly preferred to $(s_m - \epsilon, m)$. $\square$

Since u_m is not linear, there exists an l such that $u_m(l+1) - u_m(l) < u_m(l) - u_m(l-1)$. There are two possibilities: $l \leq t$ or $l > t$. First, suppose that $l \leq t$. We have that:

$$u_m(l) - u_m(l-1) + h(s_m) > u_m(l+1) - u_m(l) + h(s_m) \geq u_n(q-l) - u_n(q-l+1) + h(s_n) \tag{123}$$

where the first inequality follows from $u_m(l+1) - u_m(l) < u_m(l) - u_m(l-1)$ and the second inequality follows as $u_m(t+1) - u_m(t) + h(s_m) = u_n(q-t) - u_n(q-t-1) + h(s_n)$, u_m and u_n are concave, and $l \leq t$. Thus, for sufficiently small $\epsilon > 0$, we have that:

$$u_m(l) - u_m(l-1) + h(s_m - \epsilon) > u_n(q-l) - u_n(q-l+1) + h(s_n) \tag{124}$$

Given this inequality, we see that when there are $l-1$ group m agents and $q-l$ group n agents with higher scores, $(s_m - \epsilon, m)$ is strictly preferred to (s_n, n). Thus, to implement the optimal allocation, it must be that $P(s_m - \epsilon, m) \geq P(s_n, n)$, which is a contradiction to Claim 18.

Second, suppose that $l > t$. We know that:

$$u_m(t+1) - u_m(t) + h(s_m) = u_n(q-t) - u_n(q-t-1) + h(s_n) \tag{125}$$

As $l > t$, from concavity of u_m and u_n,

$$u_m(l) - u_m(l-1) + h(s_m) \leq u_n(q-l+1) - u_n(q-l) + h(s_n) \tag{126}$$

From concavity of u_n and u_m:

$$u_m(l+1) - u_m(l) + h(s_m) < u_n(q-l) - u_n(q-l-1) + h(s_n) \tag{127}$$

Thus, for sufficiently small $\epsilon > 0$, we have that:

$$u_m(l+1) - u_m(l) + h(s_m + \epsilon) < u_n(q-l) - u_n(q-l-1) + h(s_n) \tag{128}$$

Given this inequality, we see that when there are l group m agents and $q - l - 1$ group n agents with higher scores, (s_n, n) is strictly preferred to $(s_m + \epsilon, m)$. Thus, to implement the optimal allocation, it must be that $P(s_m + \epsilon, m) \leq P(s_n, n)$, which is a contradiction to Claim 18.

Part (3): To prove the if part, fix an ω and let μ^* denote the optimal allocation under ω. Let x^*_m denote the number of group m agents allocated the resource at μ^* and $x_m(\omega)$ denote the total number of group m agents under ω.

Claim 19. *If the authority is extremely risk-averse, then* $x^*_m \geq \min\{x_m(\omega), x^{tar}_m\}$

Proof. Assume for a contradiction this is not the case. Then $x^*_m < x_m(\omega)$ and $x^*_m < x^{tar}_m$. Since $\sum_{m \in \mathcal{M}} x^{tar}_m \leq q$ and $x^*_m < x^{tar}_m$, there exists $n \in \mathcal{M}$ such that $x^*_n > x^{tar}_n$. Let s_n denote the score of the lowest-scoring group n agent who is allocated the resource, and let s_m denote the score of any group m agent who is not allocated the resource, which exists as $x^*_m < x_m(\omega)$. Since the authority is extremely risk-averse, we have the following:

$$h(s_m) + u_m(x^*_m + 1) - u_m(x^*_m) > h(s_n) - u_n(x^*_n) + u_m(x^*_n - 1) \tag{129}$$

However, this contradicts the optimality of μ^* and proves the claim. □

Claim 20. *If the authority is extremely risk-averse,* $x^*_m > x^{tar}_m$ *and* $x^*_n > x^{tar}_n$, $\mu^*(s, m) = 0$ *and* $\mu^*(s', n) = 1$, *then* $s' > s$.

Proof. Assume for a contradiction that $s > s'$.[9] The difference in the utility of the

[9]Remember that $s' = s$ was ruled out by assumption.

authority when allocating the resource to (s, m) rather than (s', n) is given by

$$h(s)+u_m(x_m^*+1)-u_m(x_m^*)-(h(s')-u_n(x_n^*)+u_m(x_n^*-1)) = h(s)-h(s') > 0 \quad (130)$$

which is a contradiction to optimality of μ^*. □

The previous two claims show that under any ω, the optimal allocation admits (i) the highest-scoring x_m^{tar} agents from each group (provided that they exist) and (ii) highest-scoring agents who are not in (i), until the capacity is exhausted. Clearly, the quota policy $Q_m = x_m^{\text{tar}}$ and $D(R) = |\mathcal{M}| + 1$ implements this outcome at every ω.

To prove the only if part, assume that $\{Q_m\}_{m\in\mathcal{M}}$ is part of an optimal quota policy.

Claim 21. *For and each $m, n \in \mathcal{M}$ and any t, l such that $t \leq Q_m$, $Q_m > 0$ and $l \geq Q_n$, we have that:*

$$u_m(t) - u_m(t-1) + h(0) \geq u_n(l+1) - u_n(l) + h(1) \quad (131)$$

Proof. Assume that at ω, there are t group m agents, one of which one has score 0 and $l+1$ group n agents with scores higher than $1-\epsilon_1$ and q agents from other groups who have scores higher than $1-\epsilon_2$, where $\epsilon_1 > \epsilon_2 > 0$. As $t \leq Q_m$ and $Q_n < l+1$, t group m agents and $Q_n < l+1$ group n agents are admitted under Q. Since Q is optimal for all ϵ_1, we must have that:

$$u_m(t) - u_m(t-1) + h(0) \geq u_n(l+1) - u_n(l) + h(1-\epsilon_1) \quad (132)$$

The statement then follows from continuity of h by taking the limit $\epsilon_1 \to 0$. □

Claim 22. *Merit slots are processed last at the optimal quota policy.*

Proof. For a contradiction, assume there is a merit slot that is processed before a quota slot. Let l denote the last merit slot that precedes a quota slot. Let m denote a group that has a quota slot after l. We consider a state in which: (i) there are q group n agents with scores $\hat{s} - \epsilon_i$, where $\epsilon_i > 0$ for all $i \in \{1, \ldots, q\}$ (let $\underline{\hat{s}}$ denote the score of the highest-scoring agent from this group), (ii) there are Q_m group m agents with scores $\hat{s} + \epsilon_j$ for $j \in \{1, \ldots, Q_m\}$ (let $\overline{\hat{s}}$ denote the score of lowest-scoring agent from this group) and one with score $\hat{s}/2$, and (iii) q agents from other groups with scores in $(\underline{\hat{s}}, \overline{\hat{s}})$. A group m agent with score $\hat{s} + \epsilon_k$ for some k is matched to l, thus $(\hat{s}/2, m)$ is matched to a later quota slot, while some agents with type $(\hat{s} - \epsilon_j, n)$ are rejected for some j. Let $\hat{s} - \epsilon_{j'}$ be the score of the highest-scoring such agent. From the optimality of the quota policy we have that

$$u_m(Q_m + 1) - u_m(Q_m) + h(\hat{s}/2) \geq u_n(Q_n + 1) - u_n(Q_n) + h(\hat{s} - \epsilon_{j'}) \tag{133}$$

Let s^* be the score of the lowest-scoring group n agent (*i.e.*, $s^* = \min_{i \in \{1,\ldots,q\}} \hat{s} - \epsilon_i$). Next, consider the modified version of the above state, all group n agents are the same, but all of the other Q_m group m agents as well as q agents from other groups now have scores in $(s^* - \hat{\epsilon}, s^*)$ and the group m agent who had a score of $\hat{s}/2$ now has a score of $\hat{s}/2 + \hat{\epsilon}$ for $\hat{\epsilon} > 0$. Note that now the group n agent with score $\hat{s} - \epsilon_{j'}$ is allocated the slot l or an earlier slot, while the agent $(\hat{s}/2 + \hat{\epsilon}, m)$ is not allocated to any slot. Thus

$$u_m(Q_m + 1) - u_m(Q_m) + h(\hat{s}/2 + \hat{\epsilon}) \leq u_n(Q_n + 1) - u_n(Q_n) + h(\hat{s} - \epsilon_{j'}) \tag{134}$$

which, since h is strictly increasing, implies that $u_m(Q_m + 1) - u_m(Q_m) + h(\hat{s}/2) < u_n(Q_n+1) - u_n(Q_n) + h(\hat{s} - \epsilon_{j'})$. This contradicts Equation 133, proving the claim. □

Given the previous two claims, the following claim proves the result.

Claim 23. *If merit slots are processed last, then for all $l \geq Q_m$ and $j \geq Q_n$*

$$u_m(l + 1) - u_m(l) = u_n(j + 1) - u_n(j) \tag{135}$$

Proof. Assume for a contradiction this does not hold. Without loss of generality, assume $u_m(l + 1) - u_m(l) > u_n(j + 1) - u_n(j)$ and define δ as

$$\delta = (u_m(l + 1) - u_m(l)) - (u_n(j + 1) - u_n(j)) \tag{136}$$

Consider a state with $q - 1$ agents with scores higher than s^*, of which exactly Q_m are group m agents and Q_n are group n agents. Moreover, there is one more group m agent with score $s' < s^*$ (denote this agent by θ_m) and one more group n agent with score $s'' \in (s', s^*)$ where $h(s'') - h(s') < \delta$ (denote this agent by θ_n). Note that all agents apart from θ_m and θ_n are allocated the resource before the final merit slot. Moreover, since θ_n has a higher score, she obtains the final merit slot. However, this is a contradiction to the optimality of Q as $h(s'') - h(s') < \delta$ and allocating that resource to θ_m gives the authority higher utility. This proves the claim. □

Taken together, claims 21 and 23 prove that a fictitious authority that is extremely risk-averse with $x_m^{\text{tar}} = Q_m$ agrees with the authority on the optimal allocation, for all ω. To see this, observe that claim 21 implies that diversity preferences dominate any concern for scores when a group is allocated less than Q_m. Moreover, conditional on being allocated at least Q_m, it is as if there is no residual diversity preference, by claim 23. This proves the only if part of (3), which finishes the proof of the result. $\square$

B.3.4 Dominance of APM in Discrete Economies

In this section, we extend our discrete model to the multiple authority case and show that the dominance of the optimal APM in the decentralized admissions setting studied in Section 2.4.3 can be extended to this setting. Let Θ_0 denote the set of agents. $\mathcal{C} = \{c_0, c_1, \ldots, c_{|\mathcal{C}|-1}\}$ denote the set of authorities. q_c denotes the capacity of authority c and $q_{c_0} \geq |\Theta_0|$. $\theta = (s, m, \succ) \in [0,1]^{|\mathcal{C}|} \times \mathcal{M} \times \mathcal{R} = \Theta$, where $\mathcal{R}$ is set of all complete, transitive, and strict preference relations over $\mathcal{C}$ such that c_0 is less preferred than all $c \in \mathcal{C}$. For each type θ, $s_c(\theta)$ denotes the score of θ at authority c and $m(\theta)$ denotes the group of θ.

A matching in this environment is a function $\mu : \mathcal{C} \cup \Theta \to 2^\Theta \cup \mathcal{C}$ where $\mu(\theta) \in \mathcal{C}$ is the authority any type θ is assigned and $\mu(c) \subseteq \Theta$ is the set of agents assigned to authority c, which satisfies $|\mu(c)| \leq q_c$ for all c. $x_c(\mu) = \{x_{m,c}(\mu)\}_{m \in \mathcal{M}}$ denotes the number of agents of each group assigned to school c at μ while $\bar{s}_{h_c}(\mu) = \sum_{\theta \in \mu(c)} h(s(\theta))$ denotes the score utility the authority derives from μ. The preferences of the authority are given by:

$$\xi_c(\bar{s}_{h_c}, x_c) = \bar{s}_{h_c} + \sum_{m \in \mathcal{M}} u_{m,c}(x_{m,c}) \tag{137}$$

where h_c is continuous and strictly increasing and $u_{m,c} : \mathbb{R} \to \mathbb{R}$ is concave for all $m \in \mathcal{M}$ and $c \in \bar{\mathcal{C}}$.

Agents apply to the authorities sequentially, who decide which agents to admit. We index the stage of the game by $t \in \mathcal{T} = \{1, \ldots, |\mathcal{C}| - 1\}$. Each stage corresponds to an authority $I(t)$, where $I : \mathcal{T} \to \mathcal{T}$. At each stage t, any unmatched agents choose whether apply to authority $I(t)$. Given the set of applicants, authority $I(t)$ chooses to admit a subset of these agents. Given this, histories are indexed by the path of the remaining of agents who have not yet matched, $h^{t-1} = (\Theta_0, \Theta_1, \ldots, \Theta_{t-1}) \in \mathcal{H}^{t-1}$. Given each history h^{t-1} and set of applicants $\Theta_c^A \subseteq \Theta$, a strategy for an authority returns a set of agents $\Theta_c^G \subseteq \Theta$ whom they will admit such that $\Theta_c^G \subseteq \Theta_c^A$ and $|\Theta_c^G| \leq q_c$ for each time at which they could move $t \in \mathcal{T}$, $a_{c,t} : \mathcal{H}^{t-1} \times \mathcal{P}(\Theta) \to \mathcal{P}(\Theta)$, where $\mathcal{P}(\Theta)$ is the power set over Θ. A strategy for an agent returns a choice of whether to apply to authorities at each history and time for all agent types $\theta \in \Theta$,

$\sigma_{\theta,t} : \mathcal{H}^{t-1} \to [0,1]$. We moreover say that a strategy $a_{\tilde{c},t}$ for an authority $\tilde{c}$ at time t is *dominant* if it maximizes authority utility regardless of $\{\{a_{c,t}\}_{c\in\mathcal{C}/\{\tilde{c}\}}, \{\sigma_{\theta,t}\}_{\theta\in\Theta}\}_{t\in\mathcal{T}}$ and I.

Theorem 13. *The APM A_c^* is a dominant strategy for all authorities.*

Proof. We prove that APM A_c^* implements a dominant strategy for all authorities in all stages by backward induction. Consider the terminal time $t = |\mathcal{C}| - 1$. Some set of agents $\hat{\Theta} \subseteq \Theta$ applies to the authority. Regardless of $\hat{\Theta}$, by Theorem 11 we have that the set of agents chosen under any selection from APM A_c^* is first-best optimal. Thus, A_c^* is dominant. Consider now any time $t < |\mathcal{C}| - 1$, precisely the same argument applies and A_c^* is dominant. □

B.3.5 Discrete Model: Example under Imperfect Information

We develop a simple example to show how the qualitative trade-offs between priorities and quotas we have identified are those present in discrete matching markets. There are 4 students, $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and one authority c with capacity two. Students θ_3 and θ_4 belong to an underrepresented minority. The scores of minority students are distributed independently and uniformly on $[0,1]$, so that $s_3, s_4 \sim U[0,1]$. For simplicity, we assume there is no uncertainty over the scores of other students: $s_1 = s_2 = 1$.[10] We further specify that the authority has the following utility function:[11]

$$W(\beta,\mu) = \beta \mathbb{I}\{\mu(c) \cap \{\theta_3,\theta_4\} \neq \emptyset\} + \sum_{i:\mu(\theta_i)=c} s_i \tag{138}$$

This function embodies the main trade-off we have studied: the trade-off between scores and diversity. The first term indicates that whenever the authority admits at least one minority student, the utility of authority increases by β, which denotes the strength of affirmative action or diversity preferences. The second term simply indicates that the authority cares about scores and wants to admit the highest-scoring students they can. An alternative interpretation in this context, where allocating to agents with low scores is perceived as unfair, is that the authority wants to ensure outcomes that are fair in this sense.

The authority implements a stable matching and has two different policies at their disposal to influence the outcome of the matching mechanism. The first is a

[10] The qualitative result here does not change as long as non-minority students draw their scores from a distribution that FOSD $U[0,1]$.

[11] μ denotes a matching, where $\mu(c) \subset \Theta$ and $|\mu(c)| = 2$. We assume μ is stable, which uniquely determines the allocations.

priority subsidy, denoted by $\alpha \in [0,1]$. A subsidy α simply increases the scores of minority students by α and moves the distribution of scores of minority students to $U[\alpha, 1+\alpha]$. The second is a minority quota $Q \in \{0,1,2\}$ that reserves Q seats for minority students.

We start by characterizing the first-best where the authority can choose the matching they most prefer in each state of the world. Intuitively, if the score of the highest-scoring minority student is sufficiently high, then the designer prefers to admit her and one of the non-minority students. Otherwise, it is optimal for the designer to admit the two highest-scoring students, that is the two non-minority students. The first-best matching μ^* is therefore given by:

$$\mu^*(c) = \begin{cases} \{\theta_1, \theta_2\} & \text{, if} \quad \max_{i\in\{3,4\}} s_i < 1-\beta, \\ \{\theta_1, \theta_k\} & \text{, if} \quad \max_{i\in\{3,4\}} s_i > 1-\beta \text{ and } s_k = \max_{i\in\{3,4\}} s_i. \end{cases} \tag{139}$$

See that if the authority had perfect information and knew s_3 and s_4 that both a priority subsidy and a quota can implement the first-best.[12] In particular, if $\max_{i\in\{3,4\}} s_i > 1-\beta$, then both a quota $Q=1$ that reserves one seat for minority students and a priority subsidy for minority students of $\alpha \in (1-\min_{i\in\{3,4\}} s_i, 1-\max_{i\in\{3,4\}} s_i)$ implement the first-best. When $\max_{i\in\{3,4\}} s_i \leq 1-\beta$, then both a quota of $Q=0$ and a subsidy $\alpha = 0$ implement the first-best. Thus, with perfect information, both policies yield the first-best and there is no trade-off for the authority.

We now consider the second best where an authority is constrained to implement a priority or a quota before the realization of uncertainty. Note that implementing the first-best is impossible with both quotas and priorities as neither can be adapted to the underlying realized scores of the minority students. We now solve for the optimal quota and priority designs and compare their values. In order to characterize the optimal reserve policy, one first notes that reserving both seats for minority students is always strictly dominated by reserving only one. Thus the designer only needs to compare a policy with a quota of one against a policy with no quotas. With no quota, no minority student is admitted and the utility of the designer $W_{nr}(\beta) = 2$ with probability one. On the other hand, if a quota of one is used, only the highest-scoring minority student is admitted. Thus, the expected utility of the designer is:

$$\mathbb{E}[W(\beta)] = \beta + 1 + \mathbb{E}\left[\max\{s_3, s_4\}\right] = \frac{5}{3} + \beta \tag{140}$$

[12] Formally speaking, this is only true outside of the knife-edge case where $s_3 = s_4$, which is probability zero. In this case, there is no subsidy that can implement the first-best.

The optimal quota policy is therefore to reserve one seat for minority students if $\beta > 1/3$ and reserve no seats otherwise. Moreover, the utility of the designer under the optimal quota policy is:

$$V_Q(\beta) = \begin{cases} 2 & ,\text{if} \quad \beta \leq \frac{1}{3}, \\ \dfrac{5}{3} + \beta & ,\text{if} \quad \beta > \frac{1}{3}. \end{cases} \tag{141}$$

We now compute the optimal priority design and the authority's value thereof. To this end, we first calculate the utility of the designer under subsidy α. We start by calculating the matching conditional on the realized scores. There are three cases to consider. If both minority students have scores above $1-\alpha$, then both of them are admitted to the authority. If both minority students score below $1-\alpha$, then neither of them is admitted. Lastly, if one of them scores above $1-\alpha$ while the other scores below $1-\alpha$, then only one minority student is admitted. The following equation gives the utility of the designer as a function of β and α:

$$\begin{aligned} \mathbb{E}[W(\beta,\alpha)] =& \int_{1-\alpha}^{1}\int_{1-\alpha}^{1}(s_3+s_4+\beta)ds_3ds_4 + 2\int_{1-\alpha}^{1}\int_{0}^{1-\alpha}(1+s_4+\beta)ds_3ds_4 \\ &+ \int_{0}^{1-\alpha}\int_{0}^{1-\alpha} 2ds_3ds_4 \\ =& -(1+\beta)\alpha^2 + 2\beta\alpha + 2 \end{aligned} \tag{142}$$

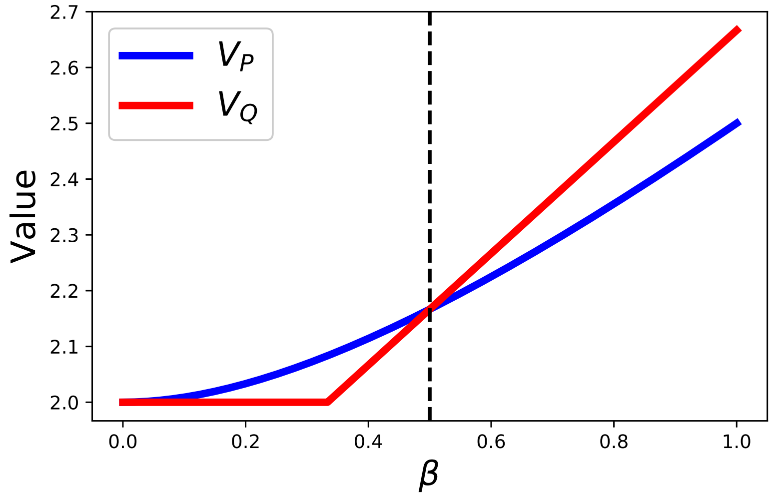
A quick calculation shows that the optimal subsidy is always interior and $\alpha^* = \frac{\beta}{1+\beta}$. Plugging the optimal subsidy policy into the authority's payoff function, we obtain:

$$V_P(\beta) = 1 + \beta + \frac{1}{1+\beta} \tag{143}$$

We now compare the value of the optimal quota and priority designs as the strength of the affirmative action motive changes. Comparing $V_P(\beta)$ and $V_Q(\beta)$ shows that the optimal policy depends on the strength of affirmative action preferences of the authority. Figure B-1 plots these two values as a function of the affirmative action motive with the dotted line giving the value $\beta = \frac{1}{2}$ at which the two value functions cross. Importantly, we see that a quota policy is optimal whenever $\beta > 1/2$ and a priority subsidy policy is optimal whenever $\beta < 1/2$.

This example highlights the main differences between priorities and quotas under uncertainty and suggests when we might expect to prefer one over the other. When the preference for diversity is low, the authority only wants to admit a minority

Figure B-1: Comparative Statics for the Preference Between Priorities and Quotas



Notes: Values of the optimal priority policy, V_P, and optimal quota policy, V_Q, as a function of the strength of the diversity preference β. The dashed black line corresponds to $\beta = 1/2$ and is the point at which both policies yield the same value.

student if her score is high enough. In this case, a subsidy is a better policy as its outcome can depend on the relative scores of the students. In particular, it only admits minority students if they obtain sufficiently high scores while a quota admits minority students equally across states of the world. Consequently, priority designs generate a desirable positive selection of minority students which tends to improve scores. However, the drawback of a subsidy policy is that it applies to all students and can therefore cause either the admission of a second minority student with a lower average score or fail to admit any minority students. On the other hand, if the preference for diversity is sufficiently high, then the authority wants to admit one minority student for sure, regardless of her score. In this case, the subsidy policy is undesirable as even under the optimal subsidy, there are many realizations where neither or both minority students are admitted, while the reserve policy ensures that one minority student is admitted in all states of the world.

B.4 Extension to More General Authority Preferences

In this Appendix, we relax Assumption 1 to allow for (i) non-separable diversity preferences, (ii) non-separable score and diversity preferences, (iii) non-differentiable preferences, and (iv) non-concave diversity preferences. We show how these changes in assumptions lead to certain modified APM mechanisms becoming first-best optimal.

B.4.1 Non-Separable Diversity Preferences

First, we relax Assumption 1 and instead suppose that the authority's preferences satisfy the following assumption:

Assumption 3. *The authority's utility function can be represented as:*

$$\xi\left(\bar{s}_h, x\right) \equiv g\left(\bar{s}_h + u(x)\right) \tag{144}$$

for some continuous, strictly increasing function $g : \mathbb{R} \to \mathbb{R}$ and a concave, partially differentiable u in each argument.

In this environment, we define a *non-separable APM* $\tilde{A} = \{\tilde{A}_m\}_{m\in\mathcal{M}}$ where $\tilde{A}_m : \mathbb{R}^{|\mathcal{M}|} \times [0,1] \to \mathbb{R}$. This implements allocation μ in state ω as per Definition 12 (under the modification of point 1 in Definition 12 to allow A_m to depend on x rather than just x_m).

We generalize Theorem 1 to show that the following non-separable APM uniquely implements the first-best optimal allocation:

Proposition 20. *The non-separable APM $\tilde{A}^*_m(y, s) \equiv h^{-1}(h(s)+u^{(m)}(y))$ and uniquely implements the first-best optimal allocation.*[13]

Proof. Follow every step in the proof of Theorem 1 with $\sum_{m\in\mathcal{M}} u_m(x_m)$ replaced by $u(x)$ and $u'_m(x_m)$ replaced by $u^{(m)}(x)$. □

Thus, allowing for non-separable diversity preferences does not substantially change the analysis of adaptive priority mechanisms. One must simply adapt the APM to be non-separable to allow cross-group diversity concerns to shape the marginal benefits of admitting agents from various groups. The main difference is that this a non-separable APM does not necessarily allow the greedy implementation of Algorithm 1. This is because, in the presence of cross-group adaptive priorities, it is no

[13] Where we define $u^{(m)}(y) = \frac{\partial}{\partial y_m} u(y)$.

longer enough to rank agents within their own group. A small adaptation to this algorithm that dynamically admits agents, starting from the highest-scoring agents in each group, would naturally implement the unique first-best optimal allocation.

B.4.2 Non-Separable Score and Diversity Preferences

Second, we relax Assumption 1 and instead suppose that the authority's preferences are represented by:

Assumption 4. *The authority's Bernoulli utility function can be represented as:*

$$\xi\left(\bar{s}_h, x\right) \tag{145}$$

where ξ is monotone, differentiable, and concave.

We define a state-dependent APM $\hat{A} = \{\hat{A}_m\}_{m\in\mathcal{M}}$ where $\hat{A}_m : \mathbb{R}^{|\mathcal{M}|} \times [0,1] \times \Omega \to \mathbb{R}$. This implements allocation μ in state ω as per Definition 12 (where point 1 in Definition 12 is modified to allow A_m to depend on both x and ω).

In this more general setting, we now find a state-dependent APM that implements the optimal allocation.

Proposition 21. *The following state-dependent APM implements a first-best optimal allocation:*

$$A_m(y, s, \omega) \equiv h^{-1}\left(h(s) + \frac{\xi_{x_m}\left(\bar{s}_h(y,\omega), y\right)}{\xi_{\bar{s}_h}\left(\bar{s}_h(y,\omega), y\right)}\right) \tag{146}$$

where $\bar{s}_h(y,\omega)$ is the score index in state ω when the highest-scoring $y = \{y_m\}_{m\in\mathcal{M}}$ agents of each attribute are allocated.

Proof. Follow every step in Theorem 1 with $\sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} + \sum_{m\in\mathcal{M}} u_m(x_m)$ replaced with $\xi\left(\bar{s}_h(y,\omega), x\right)$ where $\bar{s}_h(y,\omega) = \sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_{m,\omega}(\tilde{s})\mathrm{d}\tilde{s}$. □

There are two substantial differences in this optimal policy from our baseline APM. First, the policy depends on the joint distribution of agents in the population. Thus, specifying it *ex ante* is likely to be extremely challenging in any practical setting. This is necessary because the marginal rate of substitution between diversity and scores depends on the level of scores, which depends on the distribution of agents. Second, without assumptions on the shape of the distribution of agents, there is no guarantee that this policy is monotone and thus no guarantee that it implements a unique policy.

Thus, while Proposition 20 showed that cross-group separability is largely inessential for our main conclusions, separability between score and diversity preferences is key to the power of APM.

B.4.3 Non-Differentiable Preferences

In this section, we retain the majority of Assumption 1, where we instead suppose that the authority's diversity preferences $\{u_m\}_{m\in\mathcal{M}}$ are potentially non-differentiable at finitely many points.

As u_m is concave, the left and right derivatives of u_m, u_m^- and u_m^+, exist. The definition of our first-best APM is not applicable to this case since u_m' might not exist. Therefore, we define the following generalized optimal APM $A_m^*(y_m, s) \equiv h^{-1}(h(s)+u_m^-(y_m))$, which simply replaces u_m' with u_m^- in the definition. By concavity of u_m, u_m^- is monotone decreasing. Thus, this generalized optimal APM (as it is a montone APM) implements a unique allocation by Proposition 7. Moreover, the unique allocation that it implements is an optimal allocation:

Proposition 22. *Let μ^* denote the allocation implemented by the generalized optimal APM. μ^* is an optimal allocation.*

Proof. We first prove a claim. An allocation in this setting is a cutoff allocation if there exists cutoffs $\{s_m\}_{m\in\mathcal{M}}$ such that an agent θ is assigned the resource if and only if $s(\theta) \geq s_m$ and $m(\theta) = m$.

Claim 24. *There exists a unique optimal allocation μ' in the sense that all other allocations that attain the optimal payoff differ from μ' on at most a measure zero set of types. Moreover, there exists an optimal allocation that is a cutoff allocation.*

Proof. In the setting of Theorem 1, observe that $\underset{\sim}{s}_m(x_m)$ is strictly decreasing in x_m. This, together with the concavity of u implies that the objective is strictly concave and constraints are linear. Therefore an optimal allocation exists and is unique up to measure zero transformations. Given this allocation μ' (with measures x_m), an optimal cutoff allocation is obtained by the cutoff scores s_m^* that satisfy

$$s_m' = \sup\left\{s_m \in [0,1] : \int_{s_m}^1 \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} = x_m\right\} \tag{147}$$

□

Using this claim, toward a contradiction, assume there exists another allocation μ', which gives the authority a strictly higher utility. Moreover, take μ' to be an optimal cutoff allocation (which must exist by the claim). As μ' differs from μ^* and both are cutoff allocations, we have that there exist two groups $m, n \in \mathcal{M}$ such that: (i) $s'_m > s^*_m$ and $x'_m < x^*_m$ and (ii) $s'_n < s^*_n$ and $x'_n > x^*_m$. We have that:

$$A^*_m(x, s) \geq A^*_m(x^*_m, s) > A^*_m(x^*_m, s^*_m) \geq A^*_n(x^*_n, \hat{s}) \geq A^*_n(\hat{x}, \hat{s}) \tag{148}$$

for all $s \in (s^*_m, s'_m)$, $\hat{s} \in (s'_n, s^*_n)$, $x \leq x^*_m$, $\hat{x} \geq x^*_n$. The first inequality follows by concavity of u_m, the second follows by the fact that h is strictly increasing, the third follows by the definition of APM and the fact that $\mu^*(s^*_m, m) = 1$ and $\mu^*(\hat{s}, n) = 0$, and the fourth follows from concavity of u_n. Thus, we have that, for all $s \in (s^*_m, s'_m)$, $\hat{s} \in (s'_n, s^*_n)$, $x \leq x^*_m$, $\hat{x} \geq x^*_n$:

$$u^-_m(x) + h(s) > u^-_n(\hat{x}) + h(\hat{s}) \tag{149}$$

Thus, the total marginal utility obtained by replacing any positive measure type m students with scores $s \in (s^*_m, s'_m)$ with an identical measure of type n students with scores $\hat{s} \in (s'_n, s^*_n)$ is positive. But this contradicts the optimality of μ'. Thus, if $\tilde{\mu}$ is optimal, then $\tilde{\mu} = \mu^*$ (up to a measure zero set). □

B.4.4 Non-Concave Preferences

In this section, we relax the assumption that the u_m are concave.

Proposition 23. *If μ is an optimal allocation, then μ is implemented by A^*.*

Proof. Without concavity, the optimal allocation characterized in the proof of Theorem 1 is no longer unique. However, the Lagrangian conditions we have derived are still necessary for any optimal allocation $x = \{x_m\}_{m \in \mathcal{M}}$. Thus, any optimal allocation is implemented by A^*. □

This result shows that any optimal allocation is implemented by the optimal APM. However, when $\{u_m\}_{m \in \mathcal{M}}$ are not concave, A^* is not necessarily monotone. Therefore, A^* does not necessarily implement a unique allocation. Indeed, it is possible that A^* implements suboptimal allocations, as it will implement any locally optimal allocation. Therefore, a mechanism defined by an arbitrary selection from the allocations implemented by A^* would not be first-best optimal. However, A^* may still help decision-making in this setting as it implements any optimal allocation.

B.5 Implementation, Precedence Orders, and an Illustration from H1-B Visa Allocation

In this appendix, we show that (with no uncertainty) priority and quota policies can implement the same set of allocations. We apply this insight to study the effect of precedence orders in US H1-B visa allocation.

B.5.1 Equivalence of Priorities and Quotas for Implementation

In Proposition 8, we showed that if there is no uncertainty, both priorities and quotas can achieve the optimal allocation. We say that a priority policy P is monotone if $P(s, m)$ is strictly increasing in s. Note that since the authority prefers higher-scoring agents to lower-scoring ones, monotone policies perform better non-monotone policies. We will now show that, in the setting of Section 2.3, for a given ω (which we suppress for the rest of this section), these quota and monotone priority policies are equivalent in the sense that any allocation that is achieved by one can also be achieved by the other.

Proposition 24. *μ is implemented by a quota policy if and only if it is also implemented by a monotone priority policy.*

Proof. Assume that μ is implemented by a quota policy. Then μ is a cutoff allocation since the resource is allocated in descending order of score. Let s_m denote the lowest-scoring agent from group m who is allocated the object at μ for $m \in \mathcal{M}$. Let $\bar{s} = \max_{m \in \mathcal{M}} s_m$. Define the priority policy as $P(s, m) = s + (s_m - \bar{s})$. Note that if $\mu(s, m) = 1$ and $\mu(s', m') = 0$, then $s + s_m - \bar{s} > s' + s_{m'} - \bar{s}$ and therefore $P(s, m) > P(s', m')$. As P allocates the resource to measure q highest-scoring agents under P and measure q of agents who are allocated the resource under the quota policy has higher priorities than those who are not, P implements the same allocation as the quota policy.

Conversely, assume that μ is implemented by a monotone priority policy. Let x_m denote the measure of agents from group m allocated the object at μ for $m \in \mathcal{M}$. Let Q denote a quota policy where $Q_m = x_m$. Under any processing order, Q implements the same allocation and allocates the resource to the highest-scoring measure x_m agents from group m, for all m. This is the allocation under $P(s, m)$ since P is a monotone priority policy and allocates the resource to the highest-scoring measure x_m agents from group m, which proves the result. □

In the next section, we use this result to provide a diagnostic test for evaluating quota policies with different precedence orders by the strength of the equivalent priority policy.

B.5.2 Application of a Diagnostic Test for the Effect of Precedence Orders to US H1-B Visa Allocation

In this Appendix, we argue that in light of Proposition 24, priorities can be used as a diagnostic test for the effect of precedence orders in the context of US H1-B Visa allocation, which has had historical issues in implementation arising from the choice of precedence order. The American H1-B visa program enables American companies to temporarily employ educated foreign workers in high-skill occupations.[14] The statutory law enacted by the U.S. Congress mandates the total number of visas to be granted and The U.S. Customs and Immigration Service (USCIS) implements this mandate. The visa allocation is governed by the H-1B Visa Reform Act of 2004 that established an annual system in which 65,000 visas were made available for all eligible applicants and an additional 20,000 visas were reserved for applicants with advanced degrees. Until 2009, USCIS used the arrival time of the application to determine priorities. Since then, the priorities are determined according to a uniform lottery.

As we have emphasized, under quota policies, specifying the processing order is critical. Between 2009 and 2019, USCIS used a Reserve-Initiated processing rule. In 2020, in accordance with the 2017 *Buy American and Hire American Executive Order*, USCIS switched to a Unreserved-Initiated rule, in order to award visas to the most-skilled workers. Pathak, Rees-Jones, and Sönmez (2020) document this switch and give a detailed account of the consequences. In particular, they calculate the effect of this change on visa allocation between 2013-2017. They find that from (on average) 55,900 applicants with advanced degrees, 33,495 of them obtain a visa under the Reserve-Initiated rule, while 38,843 of them obtain a visa under the Unreserved-Initiated rule. This fact underscores how the complexity of quota policies can lead to issues in implementation, even in the simplest case with two groups.

We now use the structure of Proposition 24 to provide a diagnostic test that authorities can use to see the degree of effective affirmative action when employing a quota policy and apply it to the H1-B and Boston Public Schools settings.[15] Con-

[14] See Pathak, Rees-Jones, and Sönmez (2020) for a detailed account of H-1B policies and reforms.

[15] We note that the H-1B lottery is a setting where the perfect information assumption is justified. First, the only object that is allocated is the visa and all applicants prefer obtaining the visa to not obtaining it. This removes any uncertainty over the preferences of the individuals. Second, the priorities are determined according to a uniform random lottery and the market is large. In

Table B.1: Equivalent Priorities for Different Precedence Orders

	# Applicants		# Reserve-eligible Visas	
	General	Reserve-eligible	R-I Rule	NR-I Rule
5-yr Average (2013-2017)	137,017	55,900	33,495	38,834
Equivalent Subsidy (α)			23	35

Notes: Allocation of H-1B visas under Reserve-Initiated and Nonreserve-Initiated rules along with the equivalent subsidies that would induce these allocations.

cretely, when an authority is considering designing its precedence order, it can simply compute the implied priority subsidy being afforded to each group. In the context of H1-B allocation, we assume the uniform random lottery of USCIS is implemented by drawing a number uniformly from the interval $[0, 100]$. In the counterfactual priority mechanism, there are no quotas for reserve category applicants, but they get a score subsidy of α, i.e. their random numbers are distributed uniformly on $[\alpha, 100 + \alpha]$. Computing the implied α under both processing orders to compare the policies, we obtain Table B.1. Note that even though both quota policies correspond to 20,000 visas being reserved for applicants with advanced degrees, there is an important difference in the number of visas allocated to advanced degree applicants and therefore in the subsidy required to achieve that allocation. In particular, the Unreserved-Initiated order leads to a 12-point subsidy increase relative to the Reserve-Initiated benchmark.

particular, between 2013 and 2017, each year, 85,000 visas are allocated to an average of 137,017 reserve ineligible and 55,900 reserve eligible applicants. Thus, although there is uncertainty at the individual level, the distribution of lottery numbers conditional on reserve eligibility is essentially fixed.

B.6 Additional Quantitative Results

In this Appendix, we describe both the methodology and results of the two robustness exercises that are not discussed in full detail in the main text. First, we estimate the gains from APM when we assume that CPS sets one tier size for all tiers rather than separately optimizing the sizes of the four tiers. Second, we estimate the gains from APM under alternative utility functions that differentially penalize underrepresentation and overrepresentation.

B.6.1 Estimation with Homogeneous Reserves

As we have motivated, in this section we estimate an alternative model, where CPS chooses a single reserve size, r, instead of separate reserve sizes for all tiers. Formally, we replace the vector of reserve sizes of the four socioeconomic tiers, $r = (r_1, r_2, r_3, r_4)$ by $r = (r, r, r, r)$. In this setting, we define the marginal benefit of increasing reserve size as

$$G(r, \Lambda; \beta, \gamma) = \frac{\partial}{\partial r} \Xi(r, \Lambda; \beta, \gamma) \tag{150}$$

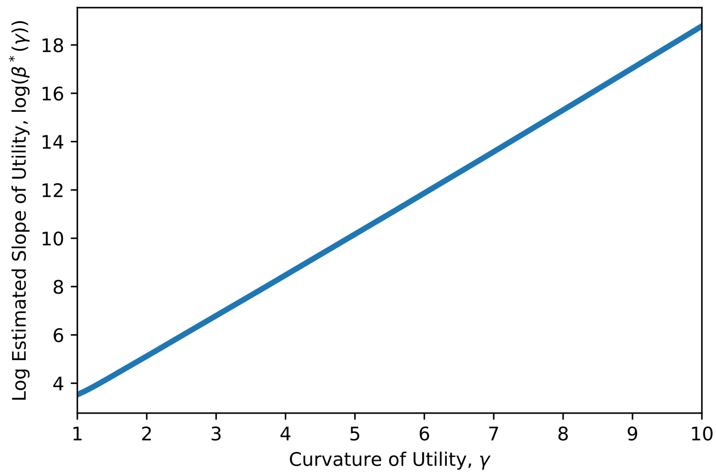
As in the general model, any (interior) reserve policy r^* must satisfy $G(\hat{r}^*, \hat{\Lambda}; \beta, \gamma) = 0$. This first-order condition yields one moment, and so we can estimate one parameter. To this end, we fix γ, and for each $\gamma \in [1, 10]$, and we estimate $\beta^*(\gamma)$ as the exact solution to the following empirical moment condition:

$$G(\hat{r}^*, \hat{\Lambda}; \beta^*(\gamma), \gamma) = 0 \tag{151}$$

Figure B-2 plots the logarithm of the estimated $\beta^*(\gamma)$. The estimated $\beta^*(\gamma)$ is increasing in γ. As the loss term $|x_t - 0.25|$ is in $(0, 1)$, $\beta^*(\gamma)$ is increasing and convex in γ, where $\beta^*(1) = 34$ and $\beta^*(10) = 1.436 \times 10^8$. In Figure B-3, we plot the gains as a function of γ, which shows that even though the estimated value for β moves quite a lot, the empirical gains range from 2 to 4 points. This also shows that the estimated gain from APM of 2.1 under our benchmark specification is close to the lower bound of the estimated gains under the alternative specification with homogeneous reserves.

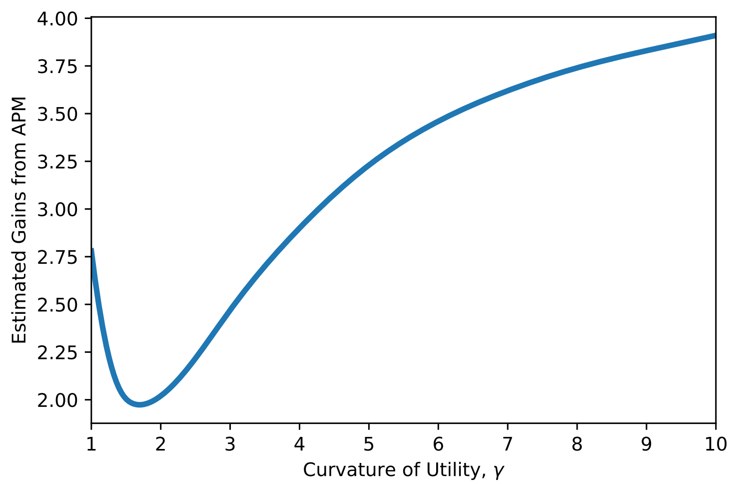
Finally, we benchmark these gains as a fraction of loss from underrepresentation under the CPS policy, where the loss of underrepresentation is calculated under the estimated parameter values. In Figure B-4, we plot the gains under APM as a percentage of diversity loss under the CPS policy. These range from 26% to 300%. Our baseline percentage gain estimate of 37.5% is again close to the lower bound that we

Figure B-2: Estimated Slope of Utility Under Homogeneous Reserves



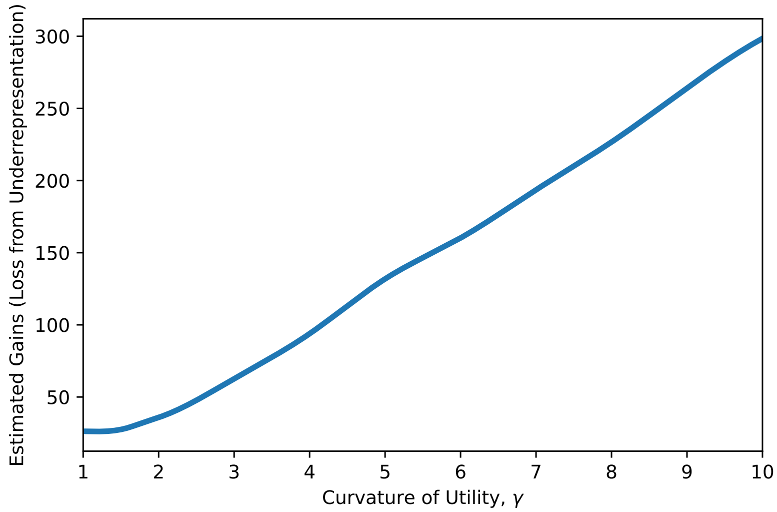
Notes: This graph plots the estimated logarithm of the slope of utility $\log \beta^*(\gamma)$ in the homogeneous reserve case as we vary the curvature of utility $\gamma \in [1, 10]$.

Figure B-3: Payoff Gains from APM Under Homogeneous Reserves



Notes: This graph plots the estimated difference in payoffs in the homogeneous reserves case between the optimal APM and the CPS policy as we vary the curvature of utility $\gamma \in [1, 10]$.

Figure B-4: The Gains from APM as a Fraction of the Loss From Underrepresentation Under Homogeneous Reserves



Notes: This graph plots the estimated difference between the payoffs under the optimal APM under homogeneous reserves as a fraction of loss from underrepresentation as we vary the curvature of utility $\gamma \in [1, 10]$.

estimate under the alternative specification with homogeneous reserves.

B.6.2 Gains from APM Under Different Utility Functions

In this section, as we have motivated, we estimate alternative objective functions to investigate the robustness of our findings.

First, we analyze a setting that includes a loss term only for underrepresented tiers (and does not penalize overrepresentation of any tier). To this end, we replace the term $|0.25 - x_t|$ with $\min\{0, (0.25 - x_t)\}$ and perform the same estimation with the following parametric utility function:

$$\xi(\bar{s}, x; \beta, \gamma) = \bar{s} + \beta \sum_{t=1}^{4} (\min\{0, (0.25 - x_t)\})^{\gamma} \tag{152}$$

The estimated parameter values are $\beta^* = 52058$ and $\gamma^* = 3.87467$. We compute the difference between the empirical payoffs under APM and the CPS reserve policy to be 0.262, which is significantly lower than our estimate of 2.1. However, the reason for this is that the diversity domain is estimated to be less important under this specification, and the diversity loss under the CPS policy is 2.71. Thus, improvements from APM correspond to 9.6% of the loss from underrepresentation, which is attenuated relative to our baseline specification, but remains non-negligible.

Second, we allow CPS to care differentially about underrepresentation and overrepresentation by considering a utility function with separate coefficients for underrepresented and overrepresented tiers. To this end, we define the following loss function:

$$f(x_t, \beta_l, \beta_h, \gamma) = \begin{cases} \beta_l (0.25 - x_t)^\gamma & \text{if } x_t \leq 0.25 \\ \beta_h (x_t - 0.25)^\gamma & \text{if } x_t > 0.25 \end{cases} \tag{153}$$

where β_l indexes the loss from underrepresentation of a tier, while β_h indexes the loss from overrepresentation. We then perform the same estimation with the following parametric utility function:

$$\xi(\bar{s}, x; \beta, \gamma) = \bar{s} + \sum_{t=1}^{4} f(x_t, \beta_l, \beta_h, \gamma) \tag{154}$$

This yields the following estimated values: $\beta_l^* = 1362270$, $\beta_h^* = 12278$, $\gamma^* = 5.28021$. We compute the difference between the empirical payoffs under APM and the CPS reserve policy to be 0.195 and the loss from underrepresentation under the CPS policy to be 2.24. Thus, we conclude that improvement from APM corresponds to 8.7% of loss from underrepresentation under the CPS policy, which is similar to what we obtain under the specification in which there is no loss from overrepresentation.

Appendix C

Appendix to Best Response Dynamics in the Boston Mechanism

C.1 Proofs for Discrete Markets

C.1.1 Proof of Lemma 1

First, a stable allocation exists and is implemented by DA by Gale and Shapley (1962). The proof follows the proof of stability of DA in Gale and Shapley (1962) with a minor modification for the possibility of skipping schools under MDA. A school c is *"possible"* for a student i if there exists a stable matching where i is matched to c. Assume that up to a given step of MDA, no student has been rejected by, or skips a school that is possible for him and let i denote the first student for which this happens. First, assume that i is rejected by or has skipped c. Then there exists set of students I_c who are accepted in that step and note that, all $i' \in I_c$ has higher scores than i. We will show that c is not a possible school for i. For a contradiction, assume that c is possible for i. Then there exists a stable matching μ such that $\mu(i) = c$. However, this means that there exists a student $i' \in I_c$ such that $c \succ_{i'} \mu(i')$, since i' cannot match any school it prefers to c as all schools i' is rejected by or skipped in the previous steps of MDA are impossible schools for i'. However, this contradicts that μ is a stable matching and therefore c is not possible for i. Since there are finitely many students and schools, and students never apply to same school twice after being rejected, both mechanisms terminate in finite time.

C.1.2 Proof of Lemma 2

If c' is not an achievable school at S, then in all σ that is compatible with S, there exists q_c students who have a higher score than $S^{c'}$ puts c' as their first choice. Therefore, under any σ_i, i is rejected by c' in any step it applied to it in Boston Mechanism. Moreover, since c is achievable, we have that $S^c \leq s(i,c)$, which means that there are at most $q_c - 1$ other students who rank c first and have higher scores than i. Therefore, i is accepted by c in step 1 of the Boston Mechanism. Since c is the most preferred school i can be matched under all σ_{-i} that is compatible with S, and any $\sigma_i \in FC_i(S)$ guarantees that i is matched to c, all such strategies are best responses.

C.1.3 Proof of Proposition 11

The proof is by induction. First, observe that in the first step of MDA and first step of BM in the first round of RBM, all students apply to the highest ranked school in their preferences and all schools admit q_c highest scoring applicants. Therefore, for all schools, the set of students who apply and the set of students who are accepted are exactly the same. Now, assume that for all $k \leq \hat{k}$, in the kth step of MDA and in first step of BM in the kth round of RBM, the set of students who apply (for MDA, this also includes the students who are tentatively assigned) and the set of students who are accepted are exactly the same. Let S_k^c denote the score of the lowest scoring student who is tentatively assigned to school c at the end of kth step of MDA. We now show a useful lemma.

Lemma 16. *S_k^c is increasing in k.*

Proof. Note that all students who are tentatively assigned to c at round k applies to c in round $k+1$. Therefore, either all these students are still tentatively assigned to c at round $k+1$ or lowest scoring l students are replaced with higher scoring l students for some $l \geq 1$. In the first case, $S_k^c = S_{k+1}^c$ while in the second case, $S_k^c < S_{k+1}^c$, which proves the result. □

We will show that if student i either (1) is tentatively assigned to school c in step k or (2) applies to school c in step $k+1$ of MDA, then i applies to c in the first step of $k+1$th round of RBM. Assume $c' \succ_i c$. Then in a previous step of MDA (say, step l) i either was rejected by school c or skipped school c. In both cases, we have that $S_l^c > s(i,c')$. Then as $k > l$, by lemma 16, $s(i,c') \geq S_k^{c'}$, which means that any such c' that is preferred to c has a cutoff above $s(i,c')$. Moreover, since i is assign to c in step k of MDA or applies to school c in step $k+1$ of MDA, we have that $s(i,c) \geq S_k^c$, which proves that c is the most preferred achievable school at S_k. Therefore, i applies to c in the first step of $k+1$th round of RBM. Since this is true for all $i \in I$, the set of applicants to each school is same under the kth step of MDA and in first step of BM in the kth round of RBM. As a result, if i is tentatively accepted to c in $k+1$th step of MDA, it is accepted to c in the first step of BM in the kth round of RBM, which finishes the induction argument.

C.1.4 Proof of Proposition 12

I first define another market, ω', where if $i \in I_U$, then $\succ_i'$ has only the first choice of i as an acceptable school. The rest of the market is the same. I will compare two cases, the outcome of RBM under ω' where all students are sophisticated and the outcome of RBM under ω where students in I_U always apply truthfully.

Claim 12. *The first round outcome of RBM under ω' (where all students are sophisticated) and the first round outcome of RBM (with unsophisticated students) under ω is same in every period.*

Proof. The proof is by induction. First, note that as in both cases first period applications are truthful for all students and first choice of all students are same in both cases, the applications in the first round of first period are the same. This implies that the set of admitted students, and first round cutoffs are the same.

To prove the induction step, suppose that first round cutoffs in period t in both settings are the same. Then all sophisticated students ($i \in I_S$) apply to the same school in both cases. Moreover, all unsophisticated students ($i \in I_U$) apply to the same school in both cases as the top school for i under $\succ_i$ is the only acceptable school for i under $\succ_i'$. Thus, the set of admitted students are the same in each school and the realized first round cutoffs are also same. Therefore, in each period, the outcome of both procedures are the same. □

As by Theorem 7, the first round outcome of RBM under ω' converges in finitely many periods, the first round outcome of RBM under ω with unsophisticated students also converges to the same outcome in the same period. Let $\hat{\mu}$ denote the matching that is realized in the first round after convergence, which is the student optimal stable matching under ω', by Theorem 7. Moreover, given this, in each period after convergence of the first round, the remaining rounds of the Boston Mechanism is the same and yields the same matching, which I denote by μ^*. Let $\hat{I}$ denote the set of all sophisticated students and unsophisticated students who are matched to a school in $\hat{\mu}$. Note that if $i \in \hat{I}$, then, $\hat{\mu}(i) = \mu^*(i)$. Let μ denote the student optimal stable matching under $\tilde{\omega}$.

Claim 13. *All students in $\hat{I}$ are weakly better off under $\hat{m}u$ compared to μ.*

Proof. If $i \in I_S$, then the result follows as under $\tilde{\omega}$, all students in I_U extend preferences. If $i \in I_U$, they are already matched to their top choice in $\hat{\mu}$, and therefore is weakly worse off under μ. □

Thus, if $i \in \hat{I}$, then i is better off under μ^* compared to μ.

Claim 14. *μ^* is stable under the market $\tilde{\omega}$.*

Proof. Assume for a contradiction i blocks μ^* at c under $\tilde{\omega}$. First, this means that c is acceptable to i. Then, c fills its quota as otherwise, i would either apply to c and get admitted or would be admitted to a more preferred school, contradicting that i blocks μ^* at c. Therefore, there exists $i' \in \mu^*(c)$ such that $i \succ_c i'$.

First, suppose that, $i \in \hat{I}$. Then there are two cases, (i) $i' \in \hat{I}$ or (ii) $i' \notin \hat{I}$. Under (i), as if $i \in \hat{I}$, then, $\hat{\mu}(i) = \mu^*(i)$, this blocking pair also blocks $\hat{\mu}$, which is a contradiction. Under (ii), since $i' \notin \hat{I}$, c does not fill its quota in round 1, which would contradict convergence to μ^* as $c \succ_i \mu^*(i)$.

Second, suppose that, $i \notin \hat{I}$. This means that $i' \notin \hat{I}$ and either (i) i' ranks c lower than i or (i) i' ranks c at the same spot as i and has higher score in c. In both cases, i would be accepted to c or a more preferred school if i' is accepted to c, which is a contradiction. □

As μ is the student optimal stable matching under $\tilde{\omega}$, from Claim 14, all students are weakly better off under μ compared to μ^*. Then Claim 13 imply that all students in $\hat{I}$ are matched to exactly same schools in μ and μ^*. The following claim completes the proof.

Claim 15. *If $i \notin \hat{I}$, then $\mu^*(i) = \mu(i)$.*

Proof. Let $\tilde{I}$ denote the set of students such that $\mu^*(i) \neq \mu(i)$. Note that any $i \in \tilde{I}$ applies to $\mu(i)$ and is rejected during the implementation of the Boston Mechanism in periods after convergence is achieved.

Let i denote a student that is rejected from $\mu(i)$ in the earliest round k during the implementation of the Boston Mechanism in periods after convergence is achieved (if there are multiple students rejected in this earliest round, i can be any of those students).

Since i is rejected, there must be another student, i' such that $\mu^*(i') \neq \mu(i')$ and i' is accepted to $\mu(i)$ before or at round k. Also note that $\mu^*(i') \neq \mu(i')$ implies that $i \notin \hat{I}$, and therefore is an unsophisticated student. Moreover, as $\mu^*(i') \neq \mu(i')$, and i' has not been rejected from $\mu(i')$ yet, which means that i' prefers $\mu^*(i')$ to $\mu(i')$, which contradicts that μ is the student optimal stable matching and completes the proof of the claim. ☐

C.1.5 Preliminaries for Discrete Markets

First, I will define some notation. Let $\tilde{S}_t$ denote the step t cutoffs under MDA. Let $B_i(S)$ denote the budget set of student i under cutoffs i. Formally, $B_i(S) = \{c \in \mathcal{C} : s_i(c) \geq S^c\}$. $D_i(S)$ denotes the demanded school of student i, which is the P_i-maximal school in $B_i(S)$.[1] $D_c(S)$ denotes the demand set of a school, which is the set of students who demand that school under S. Formally, $D_c(S) = \{i : D_i(S) = c\}$. Let $\mathcal{U}_c(\hat{s})$ denote the set of students who score higher than $\hat{s}$ in school c, $\mathcal{U}_c(\hat{s}) = \{i : s_i(c) \geq \hat{s}\}$. Moreover, let $T_c(S)$ denote the first step cutoffs under Boston Mechanism when students apply according to S. Formally,

$$T_c(S) = \begin{cases} 0 & \text{if} \quad |D_c(S)| < q_c \\ \min_{\{s:\exists i \in D_c(S) \text{ s.t. } s_c(i)=s\}} |D_c(S) \cap \mathcal{U}_c(s)| = q_c & \text{if} \quad |D_c(S)| \geq q_c \end{cases} \tag{155}$$

Moreover, define $T = \prod_{j=1}^{n} T_j$, where $T : [0,1] \to [0,1]$.

Claim 16. *$D_c(S^c, S^{-c})$ is increasing (in set inclusion sense) in S^{-c} and decreasing in S^c.*

Proof. To prove the first part, take $i \in D_c(S^c, S^{-c})$. Let $\hat{S}^{-c} \geq S^{-c}$. Then $B_i(S^c, \hat{S}^{-c}) \subseteq B_i(S^c, S^{-c})$. As $i \in D_c(S)$, c is maximal in $B_i(S)$, which implies that c is maximal in $B_i(S^c, \hat{S}^{-c})$. Thus, $i \in D_c(S^c, \hat{S}^{-c})$, proving the first part.

To prove the second part, take $i \in D_c(S^c, S^{-c})$. Let $\hat{S}^c \leq S^c$. $B_i(\hat{S}^c, S^{-c}) \setminus \{c\} = B_i(S^c, S^{-c}) \setminus \{c\}$ as the cutoffs of all other schools are the same. As $i \in D_c(S^c, S^{-c})$ and $\hat{S}^c \leq S^c$, we have that $c \in B_i(\hat{S}^c, S^{-c})$, proving the second part. ☐

Claim 17. *$T_c(S)$ is increasing in S^{-c}, decreasing in S^c.*

Proof. Suppose that $\hat{S}^{-c} \geq S^{-c}$. From Claim 16, $D_c(S^c, S^{-c}) \subseteq D_c(S^c, \hat{S}^{-c})$, which implies that $D_c(S^c, S^{-c}) \cap \mathcal{U}(S^c) \subseteq D_c(S^c, \hat{S}^{-c}) \cap \mathcal{U}(S^c)$. The first part of the result then follows from the definition of T_c. Similarly, suppose that $\hat{S}^c \leq S^c$. From

[1] If $B_i(S) = \emptyset$, then $D_i(S)$ is the most preferred school of i.

Claim 16, $D_c(S^c, S^{-c}) \subseteq D_c(\hat{S}^c, S^{-c})$, which implies that $D_c(\hat{S}^c, S^{-c}) \cap \mathcal{U}(\hat{S}^c) \subseteq D_c(S^c, S^{-c}) \cap \mathcal{U}(S^c)$. The second part of the result then follows from the definition of T_c. □

Lemma 17. *S_{DA} is a fixed point of T.*

Proof. Let μ denote the student optimal stable matching. Note that as S_{DA} is the stable matching cutoff, if $\mu(i) = c$, then $i \in D_c(S_{DA})$. To see why, suppose that, for a contradiction, $i \in D_{c'}(S_{DA})$. If $s_{c'}(i) \geq S^{c'}_{DA}$, then i and c' block μ, which is a contradiction. If $s_{c'}(i) < S^{c'}_{DA}$, this means that c is not acceptable to i, which contradicts the stability of μ.

Next, suppose that c fills its quota in the student optimal stable matching and let j denote the lowest scoring student who is matched to c at μ. Note that $D_c(S_{DA} \cap \mathcal{U}_c(S_{DA})) = \mu(c)$. If $T_c(S_{DA}) < S^c_{DA}$, then from definition of T_c, there exists a student $j' \in D_c(S_{DA})$ such that $s_c(j) = T_c(S_{DA})$. However, this is a contradiction as $|D_{S_{DA} \cap \mathcal{U}_c(T_c(S_{DA})}| > q_c$, as it includes j' and all q_c students who are matched to c at μ. Conversely, if $T_c(S_{DA}) > S^c_{DA}$, then $|D_c(S_{DA} \cap \mathcal{U}_c(T_c(S_{DA}))| < q_c$ as $j \notin U_c(S_{DA})$, which is a contradiction.

Finally suppose that c does not fill its quota in the student optimal stable matching, which means that $S^c_{DA} = 0$. Moreover, from stability of μ, $|D_c(S_{DA})| < q_c$, which means that $T_c(S_{DA}) = 0$, proving the result. □

C.1.6 Proof of Proposition 13

Fix an $\hat{S}_t$ such that $\hat{S}_t \leq S_{DA}$. Define S_t by setting $S^c_t = S^c_{DA}$ and $S^{-c}_t = \hat{S}^{-c}_t$. Let $\mathcal{D}(S) = \{i : i \in D_c(S_t)\, s_c(i) \geq S^c_{DA}\}$. From Lemma 17, $T_c(S_{DA}) = S^c_{DA}$. Then by Claim 17, we have $T_c(S) \leq S^c_{DA}$. Therefore, there are at most q_c students who has scores above S^c_{DA} demand c, that is, $|\mathcal{D}(S_t)| \leq q_c$. Moreover, if $i \in \mathcal{D}(\hat{S}_t)$ but $i \notin D(S_t)$, then $s_c(i) \leq S^c_t \leq S^c_{DA}$. Thus, $|\mathcal{D}(\hat{S}_t)| \leq q_c$, which implies $T_c(\hat{S}_t) \leq S_{DA}$, proving the result.

C.1.7 Proof of Proposition 14

Suppose that μ is in a cycle and is Pareto dominated by the SOSM, denoted by μ^*. Suppose that μ appears in round $t+1$, that is, following the cutoffs S_t. First, note that $S_t \leq S_{DA}$. Therefore, in the first step of BM in period t of RBM, all students apply to a school that they weakly prefer to their match in the μ^*. Thus, at the end of first step, all students who are matched to a school are weakly better off compared to μ^*. As μ^* Pareto dominates μ, all such students are matched to the school they match at μ^*.

Suppose that k is the first step in the period t of RBM such that a student is matched to a school which is less preferred than her match under μ^*, or does not have any school to apply even though she is matched to a school at μ^*. Denote this student by i. Note that this is only possible if $\mu^*(i)$ has exhausted its capacity in the previous step. This means that there exists a student i' such that $\mu^*(i') \neq \mu^*(i)$, but i' is matched to $\mu^*(i)$ in a previous step. However, as no student has received a match that is less preferred to their match under μ^* at any previous step, i' must be strictly better off under μ compare to μ^*, which is a contradiction.

C.1.8 Proof of Proposition 15

I first prove a few useful results.

Claim 18. $T_c(\tilde{S}_t) = \tilde{S}^c_{t+1}$

Proof. Immediate from the fact that the set of students who apply to c in step $t+1$ of MDA is $D_c(S)$. □

Claim 19. *If $T_c(S^c, S^{-c}) = \hat{S}^c$ and $\hat{S}^c \geq S^c$, then $T_c(s, S^{-c}) = \hat{S}^c$ for all $s \leq \hat{S}^c$.*

Proof. If $\hat{S}^c = 0$, then the result is immediate as $s = \hat{S}^c$. Suppose that $\hat{S}^c > 0$. Then $|D_c(S^c, S^{-c}) \cap \mathcal{U}(\hat{S}^c)| = q_c$. Then, $T_c(\hat{S}^c, S^{-c}) = \hat{S}^c$. Thus, from Claim 17, $T_c(s, S^{-c}) \leq \hat{S}^c$ for all $s \leq \hat{S}^c$. Moreover, observe that for all $s \leq \hat{S}^c$, $D_c(\hat{S}^c, S^{-c}) \cap \mathcal{U}(\hat{S}^c) = D_c(s, S^{-c}) \cap \mathcal{U}(\hat{S}^c)$, which means that there are q_c students who demand c at cutoffs s, S^{-c} with scores over $\hat{S}^c$. Therefore, $T_c(s, S^{-c}) = T_c(\hat{s}, S^{-c})$. □

Now, assume that $S_k \in [\tilde{S}_t, \tilde{S}_{t+1}]$. We have the following claim.

Claim 20. $T_c(S_k) \leq \tilde{S}^c_{t+2}$.

Proof. From Claim 18, $T_c(\tilde{S}^c_{t+1}, \tilde{S}^{-c}_{t+1}) = \tilde{S}^c_{t+2}$. From Claim 19, as $\tilde{S}^c_{t+2} \geq \tilde{S}^c_{t+1}$, we have that $T_c(S^c_k, S^{-c}_k) = T_c(\tilde{S}^c_{t+1}, \tilde{S}^{-c}_{t+1})$. Moreover, as $S^{-c}_k \leq \tilde{S}^{-c}_{t+1}$, from Claim 17, we have that

$$T_c(S^c_k, S^{-c}_k) \leq T_c(S^c_k, \tilde{S}^{-c}_{t+1}) = T_c(\tilde{S}_{t+1}) = \tilde{S}^c_{t+2} \tag{156}$$

□

Claim 21. $T_c(S_k) \geq \tilde{S}^c_{t+1}$.

Proof. From Claim 18, $T_c(\tilde{S}^c_t, \tilde{S}^{-c}_t) = \tilde{S}^c_{t+1}$. From Claim 19, as $\tilde{S}^c_{t+1} \geq S^c_k$, $T_c(S^c_k, \tilde{S}^{-c}_t) = \tilde{S}^c_{t+1}$. Moreover, as $S^{-c}_k \geq \tilde{S}^{-c}_t$, from Claim 17, we have that

$$T_c(S^c_k, S^{-c}_k) \geq T_c(S^c_k, \tilde{S}^{-c}_t) = \tilde{S}^c_{t+1} \tag{157}$$

The following Lemma is immediate from Claims 20 and 21.

Lemma 18. *If S_k is compatible with step t cutoffs of MDA, then S_{k+1} is compatible with step $t+1$ cutoffs of MDA*

Given the lemma, the proposition follows from the fact that MDA converges to the student optimal stable matching in finitely many steps (Proposition 11 and Theorem 7). □

C.1.9 Proof of Proposition 16

The proof is by induction. First, note that in the first period, regardless of S_0, the highest scoring student applies to her most preferred school and is accepted. Note that this school is the student's assignment under the serial dictatorship (and therefore, the deferred acceptance) mechanism. Let μ_S denote the serial dictatorship outcome.

Let I_t denote highest scoring t students. Now suppose that at period t, all students in I_t are assigned to their match under the serial dictatorship mechanism at the first round of the Boston Mechanism. I will show that in period $t+1$, all students in I_{t+1} are assigned to their match under the serial dictatorship mechanism at the first round of the Boston Mechanism. First, note that in the first round of period $t+1$, if $i \in I_t$, then i applies to $\mu_S(i)$ as $\mu_S(i)$ is attainable under S_t and all schools more preferred to $\mu_S(i)$ are not attainable under S_t. Moreover, all such i are admitted.

Next, let j denote the student with $t+1$th highest score. As $c \equiv \mu_S(j)$ is the school of j under serial dictatorship and all higher scoring students are assigned to their serial dictatorship outcome at round t, $S_t^c \leq s(i)$. Moreover, for any school c' that j prefers to c, $S_t^{c'} > s(i)$ as otherwise, i would be matched to c' under serial dictatorship. Therefore, j applies to c. Moreover, as all students in I_t are assigned to their match under serial dictatorship in the first round of period $t+1$, j is admitted to c in the first round of period $t+1$ as otherwise, j must be rejected in favor of a lower scoring student, which is a contradiction. This proves the result as the set of students are finite.

C.2 Proofs for Large Markets

C.2.1 Preliminaries for Continuum Markets

The Boston Mechanism (BM) - Continuum Market

Step 1: Students apply to their first choice school. Each school c admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. All other offers are immediately accepted and become permanent matches. School capacities are adjusted accordingly.

Step t: If a student is rejected in Step $t-1$, he applies to the next school on his rank-order list. If he has no more schools on his list, he applies nowhere. Each school c admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. All other offers become permanent matches. School capacities are adjusted accordingly. **Stop**: The algorithm stops when no rejections are issued.

Note that Boston Mechanism terminates in at most $|\mathcal{C}|$ rounds, as all students are either permanently matched or have run out of schools to apply.

The Modified Deferred Acceptance Mechanism (MDA) - Continuum Market

Step 1: Students apply to their first choice school. Each school c tentatively admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. Schools reject all students who are not tentatively accepted.

Step t: If a student is rejected in Step $t-1$, he *applies to the highest school on his rank-order list within the schools that either (i) did not fill its capacity in the last round or (ii) temporarily hold the offer of a student who has lower ranking in that school.* If there are no schools in his list that satisfies either (i) or (ii), or that school is ranked below the outside option, he applies nowhere. Schools consider both new offers and the offers held from previous rounds and tentatively admits all students over some cutoff S_c, where S_c is the infimum over all cutoffs where the school c does not exceed its capacity. Schools reject all students who are not tentatively accepted.

Stop: The algorithm stops when no rejections are issued. Each school is matched to the students it is holding at the end.

Let $\tilde{S}_n$ denote the cutoffs in step n of the MDA mechanism. I now define a useful function that maps a distribution of types, capacities and cutoffs to a new set of cutoffs after one round of admissions. As in the finite market model, $B_\theta(S)$ denotes the budget set of student θ while $D_\theta(S)$ denote the demand of student θ.

$D_c(S) = \{\theta : D_\theta(S) = c\}$ is the demand set of school c and $\mathcal{U}_c(s)$ return the students with scores higher than s at school c.

$$H_c(S) = \min_{\hat{s}\in[0,1]} |D_c(S) \cap \mathcal{U}_c(\hat{s})| \leq q_c \tag{158}$$

where the minimum exists as η has full support. Let $H(S) = \{H_{c_1}(S), \ldots, H_{c_n}(S)\}$. The following claim is immediate from the definition of MDA.

Claim 22. $H^n(0, \ldots, 0) = \tilde{S}_n$. *That is, starting with zero cutoffs in all schools and applying T n times gives the cutoffs in step n of MDA.*

Let $\tilde{\Theta}^\epsilon_S = \{\theta : s_c(\theta) \in \mathcal{B}_\epsilon(S_c)$ for some $c\}$, where $\mathcal{B}_\epsilon(S_c)$ is the ϵ neighborhood around S_c. Let d denote the euclidean distance. The following lemma is useful in showing the continuity of T.

Claim 23. *For each $\theta \notin \tilde{\Theta}^0_S$, if $d(S, \hat{S}) < \min_c |s_c(\theta) - S_c|$, then $D_\theta(\hat{S}) = D_\theta(S)$.*

Proof. As $\theta \notin \tilde{\Theta}^0_S$, $\hat{\delta} \equiv \min_c |s_c(\theta)-S_c| > 0$. Moreover, for all $\hat{S}$ such that $d(S, \hat{S}) < \hat{\delta}$, $B_\theta(\hat{S}) = B_\theta(S)$. As students demand the same school when their budget set is the the same, $D_\theta(\hat{S}) = D_\theta(S)$, proving the claim. □

Claim 24. *For each $\epsilon > 0$, there exists α_ϵ such that $\eta(\tilde{\Theta}^\alpha_S) < \epsilon$ for all $\alpha \leq \alpha_\epsilon$.*

Proof. Immediate from the fact that η has no mass points and admits a density. □

Lemma 19. *$H_c(S)$ is continuous in S.*

Proof. Let $DD(S, \hat{S}) = \{\theta : D_\theta(S) \neq D_\theta(\hat{S})\}$ denote the set of students whose demanded school is different under S and $\hat{S}$. Note that by Claim 23, $DD(S, \hat{S}) \subseteq \tilde{\Theta}^\alpha_S$ for $\alpha = d(S, \hat{S})$.

Claim 25. *Fix S and let $\bar{\eta} > 0$. There exists $\delta_{\bar{\eta}}$ such that if $d(S, \hat{S}) < \delta_{\bar{\eta}}$, then $\eta(DD(S, \hat{S})) < \bar{\eta}$.*

Proof. From Claim 24, there exists $\alpha_{\bar{\eta}}$ such that $\eta(\tilde{\Theta}^{\alpha_{\bar{\eta}}}_s) < \bar{\eta}$ for all $\alpha < \alpha_{\bar{\eta}}$. From Claim 23, for all $\theta \notin \tilde{\Theta}^{\alpha_{\bar{\eta}}}_S$, $D_\theta(S) = D_\theta(\hat{S})$ for all $\hat{S}$ with $d(S, \hat{S}) < \alpha_{\bar{\eta}}$. As $DD(S, \hat{S}) \subseteq \tilde{\Theta}^{\alpha_{\bar{\eta}}}_S$ for all $\hat{S}$ with $d(S, \hat{S}) < \alpha_{\bar{\eta}}$ and $\eta(\tilde{\Theta}^{\alpha_{\bar{\eta}}}_S) < \bar{\eta}$, we have that $\eta(DD(S, \hat{S})) < \bar{\eta}$ for all $\hat{S}$ with $d(S, \hat{S}) < \alpha_{\bar{\eta}}$, proving the claim. □

To prove the lemma, let $\epsilon > 0$ be given and fix S. Define

$$\Theta^{\epsilon\downarrow} = \{\theta : s_c(\theta) \in (H_c(S) - \epsilon, H_c(S)), c \succ_\theta c_0 \succ_\theta c' \text{ for all } c' \notin \{c, c_0\}\} \tag{159}$$

$$\Theta^{\epsilon\uparrow} = \{\theta : s_c(\theta) \in (H_c(S) + \epsilon/2, H_c(S) + \epsilon), c \succ_\theta c' \text{ for all } c' \neq c\} \tag{160}$$

Let $\bar{\eta} = \min\{\eta(\Theta^{\epsilon\uparrow}), \eta(\Theta^{\epsilon\downarrow})\}$. By Claim 25, there exists $\delta_{\bar{\eta}}$ such that if $d(S, \hat{S}) < \delta_{\bar{\eta}}$, then $\eta(DD(S, \hat{S})) < \bar{\eta}$. Moreover, if $d(S, \hat{S}) < \min\{\epsilon/2, \delta_{\bar{\eta}}\}$, for $\theta \in \Theta^{\epsilon\uparrow} \cup \Theta^{\epsilon\downarrow}$, $D_\theta(S) = D_\theta(\hat{S}) = c$.

I will now show that if $d(S, \hat{S}) < \min\{\epsilon/2, \delta_{\bar{n}}\}$, then $|H_c(\hat{S}) - H_c(S)| < \epsilon$. There are three cases, $H_c(\hat{S}) = H_c(S)$, $H_c(\hat{S}) > H_c(S)$ and $H_c(\hat{S}) < H_c(S)$. If $H_c(\hat{S}) = H_c(S)$, then we are done.

Suppose that $H_c(\hat{S}) > H_c(S)$. Let $\Theta_N = \{\theta : D_\theta(\hat{S}) = c, D_\theta(S) \neq c\}$ denote the set of students who demand c under $\hat{S}$ but not under S. As $d(S, \hat{S}) < \delta_{\bar{\eta}}$, $DD(S, \hat{S}) < \bar{\eta}$. Therefore, $\eta(\Theta_N) < \bar{\eta}$. But this means that, there can be at most measure $\bar{\eta}$ new students who demand c under $\hat{S}$. As all $\theta \in \Theta^{\epsilon\uparrow}$ still demand c under $\hat{S}$ and $\eta(\Theta^{\epsilon\uparrow}) = \bar{\eta}$, $H_c(\hat{S}) - H_c(S) < \epsilon$.

Next, suppose that $H_c(\hat{S}) < H_c(S)$. Note that $\Theta^{\epsilon\downarrow}$ has positive measure. Let $\Theta_O = \{\theta : D_\theta(S) = c, D_\theta(\hat{S}) \neq c\}$ denote the set of students who demand c under S but not under $\hat{S}$. As $d(S, \hat{S}) < \delta_{\bar{\eta}}$, $DD(S, \hat{S}) < \bar{\eta}$. Therefore, $\eta(\Theta_O) < \bar{\eta}$. But this means that, there can be at most measure $\bar{\eta}$ students who demand c under S but not under $\hat{S}$. As all $\theta \in \Theta^{\epsilon\downarrow}$ still demand c under $\hat{S}$ and $\eta(\Theta^{\epsilon\uparrow}) = \bar{\eta}$, $H_c(S) - H_c(\hat{S}) < \epsilon$.

This shows that for each ϵ, we can find $\alpha = \min\{\bar{\eta}, \epsilon/2\}$ such that $|H_c(S) - H_c(\hat{S})| < \epsilon$. Thus, for each ϵ, we can find a δ such that $d(T(S), T(\hat{S})) < \epsilon$ whenever $d(S, \hat{S}) < \delta$.

□

Lemma 20. *S_{DA} is the unique fixed point of H.*

Proof. I start by showing that, if S is a fixed point of H, then $\nu = \mathcal{M}(S)$ is a stable matching. First, I show that ν is a matching. To see why, from definition of H_c, for each c there is a measure of q_c students with $D_\theta(S) = c$ and $s_c(\theta) \geq S_c$. Therefore, there is a measure of q_c students with $\tilde{D}_\theta(S) = c$ and $\eta(\nu(c)) \leq q_c$ for all c. Assume for a contradiction (θ, c) blocks ν. From definition of ν, $\tilde{D}_\theta(S) = \nu(\theta)$. Moreover as (θ, c) blocks ν, we have $c \succ_\theta \nu(\theta)$ and $s_c(\theta) \geq S_c$. However, this implies that $\tilde{D}_\theta(S) \neq \nu(\theta)$, which is a contradiction.

As full support assumption implies that the market has a unique stable matching, there is exactly one fixed point of H, which corresponds to the stable matching. □

C.2.2 Proof of Proposition 17

First, note that $\tilde{S}_1 \geq \tilde{S}_0 = (0, \ldots, 0)$.

Lemma 21. *$\tilde{S}_n$ is increasing. That is, $\tilde{S}^c_{n+1} \geq \tilde{S}^c_n$ for all c, n.*

Proof. Proof is by induction. The base case holds as $\tilde{S}_1 \geq \tilde{S}_0$. Suppose that $\tilde{S}_n \geq \tilde{S}_{n-1}$. Fix a $c \in \mathcal{C}$. If $\tilde{S}^c_n = 0$, then $\tilde{S}^c_{n+1} \geq \tilde{S}^c_n = 0$ and we are done. Therefore, suppose that $\tilde{S}^c_n > 0$. Define

$$\Theta^c_n = \{\theta : D_\theta(\tilde{S}_{n-1}) = c, s_c(\theta) \geq \tilde{S}^c_n\} \tag{161}$$

As $\tilde{S}^c_n > 0$, $\eta(\Theta^c_n) = q_c$. Moreover, as for all c $\tilde{S}^c_n \geq \tilde{S}^c_{n-1}$ and $s_c(\theta) \geq \tilde{S}^c_n$, for all $\theta \in \Theta^c_n$, we have that $D_\theta(\tilde{S}_n) = c$ for all $\theta \in \Theta^c_n$. As $\eta(\Theta^c_n) = q_c$, $\tilde{S}^c_{n+1} = H_c(\tilde{S}_n) \geq \tilde{S}_n$. Repeating this for all c proves the lemma. □

Given Lemma 21, as $\tilde{S}_n$ is bounded, $\lim_{n\to\infty} \tilde{S}_n = S^*$ for some S^*. From continuity of H (Lemma 19) and the fact that $H^n(0, \ldots, 0) = \tilde{S}_n$, S^* is a fixed point of H. As H has a unique fixed point, which is S_{DA} (Lemma 20), the result follows.

C.2.3 Proof of Proposition 18

First, Claims 16, 17, 19, 20 and 21 as well as Lemma 18 hold in the continuum model, with essentially the same proofs, replacing the function T with H. The result then follows from Proposition 17.

C.2.4 Preliminaries for Finite Markets Sampled From a Continuum Market

A sequence of finite markets $\{F^k\}_{k\in\mathbb{N}}$ where $F^k = [\eta^k, Q^k]$ *converges* to a continuum market $F = [\eta, Q]$ if the empirical distribution of types η^k converges to η in the weak sense and if capacity per student Q^k converges to Q.

C.2.5 Proof of Proposition 19

The proof is by induction. Suppose that $\lim_{k\to\infty} \tilde{R}_t(k) \xrightarrow{p} R_t$. I will show that $\lim_{k\to\infty} \tilde{R}_{t+1}(k) \xrightarrow{p} R_{t+1}$, which amounts to showing, for all ϵ,

$$\lim_{k\to\infty} Pr\left(|\tilde{R}_{t+1}(k) - R_{t+1}| > \epsilon\right) = 0 \tag{162}$$

Let F^k_t denote the distribution of $\tilde{R}_t(k)$, while F_t denotes the (degenerate) distribution of R_t. Then the following must be shown:

$$\lim_{k\to\infty} Pr\left(\left|\int T(r,k)dF_t^k(r) - H(R_t)\right| > \epsilon\right) = 0 \tag{163}$$

Claim 26. *The following is true*

$$\lim_{k\to\infty} Pr\left(|H(\tilde{R}_t(k)) - H(R_t)| > \epsilon/2\right) = 0 \tag{164}$$

Proof. Follows from continous mapping theorem given the continuity of H and the assumption that $\lim_{k\to\infty} \tilde{R}_t(k) \xrightarrow{p} R_t$. □

Claim 27. *The following is true*

$$\lim_{k\to\infty} Pr\left(\left|\int T(r,k)dF_t^k(r) - H(\tilde{R}_t(k))\right| > \epsilon/2\right) = 0 \tag{165}$$

Proof. We can rewrite $H(\tilde{R}_t(k))$ as

$$H(\tilde{R}_t(k)) = \int H(r)dF_t^k(r) \tag{166}$$

Thus, Equation 165 becomes

$$\lim_{k\to\infty} Pr\left(\left|\int T(r,k) - H(r)dF_t^k(r)\right| > \epsilon/2\right) = 0 \tag{167}$$

As $T(r,k)$ converges to $H(r)$ pointwise, F_t^k converges to F_t and both $T(r,k)$ and F_t^k are bounded for all k, by dominated convergence theorem,

$$\lim_{k\to\infty} \int |T(r,k) - H(r)|\, dF_t^k(r) = 0 \tag{168}$$

which implies equation 167 and proves the claim. □

Taken together, claims 26 and 27 imply equation 163 and finishes the proof of the inductive step. To prove the base case of induction, we prove the following claim, which finishes the proof of the proposition.

Claim 28. $\lim_{k\to\infty} \tilde{R}_1(k) \xrightarrow{p} R_1$.

Proof. Since $\tilde{R}_0(k) = R_0$, we have that $\lim_{k\to\infty} \tilde{R}_0(k) \xrightarrow{p} R_0$. The result then follows from the proof we had for the inductive step. □

List of Figures

List of Tables

Bibliography

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

ABDULKADIROĞLU, A., Y.-K. CHE, AND Y. YASUDA (2011): "Resolving conflicting preferences in school choice: The" boston mechanism" reconsidered," *American Economic Review*, 101(1), 399–410.

——— (2015): "Expanding 'choice' in school choice," *American Economic Journal: Microeconomics*, 7(1), 1–42.

ABDULKADIROĞLU, A. (2005): "College Admission with Affirmative Action," *International Journal of Game Theory*, 33, 535–549.

ABDULKADIROĞLU, A., AND T. SÖNMEZ (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

ADAMS, E. W. (1965): "Elements of a theory of inexact measurement," *Philosophy of science*, 32(3/4), 205–228.

AGARWAL, N., AND P. SOMAINI (2018): "Demand analysis using strategic reports: An application to a school choice mechanism," *Econometrica*, 86(2), 391–444.

AKBARPOUR, M., E. B. BUDISH, P. DWORCZAK, AND S. D. KOMINERS (2021): "An economic framework for vaccine prioritization," *Available at SSRN 3846931*.

AKBARPOUR, M., A. KAPOR, C. NEILSON, W. VAN DIJK, AND S. ZIMMERMAN (2022): "Centralized School choice with unequal outside options," *Journal of Public Economics*, 210, 104644.

ANGRIST, J. D., P. A. PATHAK, AND R. A. ZÁRATE (2019): "Choice and consequence: Assessing mismatch at chicago exam schools," Discussion paper, National Bureau of Economic Research.

ATHEY, S., C. AVERY, AND P. ZEMSKY (2000): "Mentoring and diversity," *American Economic Review*, 90(4), 765–786.

AYGUN, O., AND I. BÓ (2021): "College admission with multidimensional privileges: The Brazilian affirmative action case," *American Economic Journal: Microeconomics*, 13(3), 1–28.

AYGÜN, O., AND T. SÖNMEZ (2013): "Matching with contracts: Comment," *American Economic Review*, 103(5), 2050–51.

AYGÜN, O., AND B. TURHAN (2017): "Large-scale affirmative action in school choice: Admissions to IITs in India," *American Economic Review*, 107(5), 210–13.

——— (2020): "Dynamic reserves in matching markets," *Journal of Economic Theory*, 188, 105069.

Azevedo, E. M., and J. D. Leshno (2016): "A supply and demand framework for two-sided matching markets," *Journal of Political Economy*, 124(5), 1235–1268.

Babaioff, M., Y. A. Gonczarowski, and A. Romm (2018): "Playing on a level field: sincere and sophisticated players in the Boston mechanism with a coarse priority structure," *arXiv preprint arXiv:1810.06229*.

Balinski, M., and T. Sönmez (1999): "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84, 73–94.

Boston Herald (2022): "New Boston point system dings students for good grades," *Boston Herald*, January 18, 2022.

Calsamiglia, C., C. Fu, and M. Güell (2020): "Structural estimation of a model of school choices: The boston mechanism versus its alternatives," *Journal of Political Economy*, 128(2), 642–680.

Carvalho, J.-P., B. Pradelski, and C. Williams (2022): "Affirmative Action with Multidimensional Identities," *Available at SSRN 4070930*.

Çelebi, O., and J. P. Flynn (2021): "Priorities *vs.* Quotas," *Available at SSRN 3562665*.

Çelebi, O. (2022): "Diversity Preferences, Affirmative Action and Choice Rules," *Working Paper*.

Çelebi, O., and J. P. Flynn (2022a): "Adaptive Priority Mechanisms," *Working Paper*.

——— (2022b): "Priority design in centralized matching markets," *The Review of Economic Studies*, 89(3), 1245–1277.

Chan, J., and E. Eyster (2003): "Does banning affirmative action lower college student quality?," *American Economic Review*, 93(3), 858–872.

Chen, Y., and O. Kesten (2017): "Chinese college admissions and school choice reforms: A theoretical analysis," *Journal of Political Economy*, 125(1), 99–139.

Chicago Public Schools (2022): "Selective Enrollment Admissions Process: Community Feedback," *Chicago Public Schools*, March 9, 2022.

Crenshaw, K. (2013): "Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics," in *Feminist Legal Theories*, pp. 23–51. Routledge.

Doğan, B. (2016): "Responsive affirmative action in school choice," *Journal of Economic Theory*, 165, 69–105.

Dubins, L. E., and D. A. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *The American Mathematical Monthly*, 88(7), 485–494.

Dur, U., R. G. Hammond, and O. Kesten (2021): "Sequential school choice: Theory and evidence from the field and lab," *Journal of Economic Theory*, 198, 105344.

Dur, U., S. D. Kominers, P. A. Pathak, and T. Sönmez (2018): "Reserve design: Unintended consequences and the demise of Boston's walk zones," *Journal of Political Economy*, 126(6), 2457–2479.

Dur, U., P. A. Pathak, F. Song, and T. Sönmez (2022): "Deduction dilemmas: The Taiwan assignment mechanism," *American Economic Journal: Microeconomics*, 14(1), 164–85.

Dur, U., P. A. Pathak, and T. Sönmez (2020): "Explicit vs. statistical targeting in affirmative action: theory and evidence from Chicago's exam schools," *Journal of Economic Theory*, 187, 104996.

Dur, U., M. Thayer, and W. Phan (2022): "Allocating Medical Resources During a Pandemic," *Working Paper*.

Echenique, F., and M. B. Yenmez (2015): "How to control controlled school choice," *American Economic Review*, 105(8), 2679–94.

Ehlers, L., I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim (2014): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," *Journal of Economic Theory*, 153, 648–683.

Ellison, G., and P. A. Pathak (2021): "The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools," *American Economic Review*, 111(3), 943–75.

Epple, D., R. Romano, and H. Sieg (2008): "Diversity and affirmative action in higher education," *Journal of Public Economic Theory*, 10(4), 475–501.

Erdil, A., and T. Kumano (2019): "Efficiency and stability under substitutable priorities with ties," *Journal of Economic Theory*, 184, 104950.

Ergin, H., and T. Sönmez (2006): "Games of school choice under the Boston mechanism," *Journal of public Economics*, 90(1-2), 215–237.

Gale, D., and L. S. Shapley (1962): "College admissions and the stability of marriage," *The American Mathematical Monthly*, 69(1), 9–15.

Goto, M., F. Kojima, R. Kurata, A. Tamura, and M. Yokoo (2017): "Designing matching mechanisms under general distributional constraints," *American Economic Journal: Microeconomics*, 9(2), 226–62.

Grigoryan, A. (2021): "Effective, fair and equitable pandemic rationing," *Available at SSRN 3646539*.

Gul, F., and E. Stacchetti (1999): "Walrasian equilibrium with gross substitutes," *Journal of Economic theory*, 87(1), 95–124.

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

Hastings, J., T. J. Kane, and D. O. Staiger (2009): "Heterogeneous preferences and the efficacy of public school choice," *NBER Working Paper*, 2145, 1–46.

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

Hatfield, J. W., and F. Kojima (2008): "Matching with Contracts: Comment," *American Economic Review*, 98, 1189–1194.

——— (2010): "Substitutes and stability for matching with contracts," *Journal of Economic theory*, 145(5), 1704–1723.

He, Y. (2015): "Gaming the Boston school choice mechanism in Beijing," *Working Paper*.

Hughes, M. M. (2018): "Ethnic Quotas in Electoral Politics," *Gender Parity and Multicultural Feminism: Towards a New Synthesis*, p. 97.

Kamada, Y., and F. Kojima (2017): "Stability concepts in matching under distributional constraints," *Journal of Economic Theory*, 168, 107–142.

——— (2018): “Stability and strategy-proofness for matching with constraints: A necessary and sufficient condition,” *Theoretical Economics*, 13(2), 761–793.

Kapor, A. J., C. A. Neilson, and S. D. Zimmerman (2020): “Heterogeneous beliefs and school choice mechanisms,” *American Economic Review*, 110(5), 1274–1315.

Kelso, A. S., and V. P. Crawford (1982): “Job matching, coalition formation, and gross substitutes,” *Econometrica: Journal of the Econometric Society*, pp. 1483–1504.

Kojima, F. (2012): “School Choice: Impossibilities for Affirmative Action,” *Games and Economic Behavior*, 75, 685–693.

Kojima, F., N. Sun, and N. N. Yu (2020a): “Job matching under constraints,” *American Economic Review*, 110(9), 2935–47.

——— (2020b): “Job matching with subsidy and taxation,” *Available at SSRN 3624343*.

Kominers, S. D., and T. Sönmez (2016): “Matching with slot-specific priorities: Theory,” *Theoretical Economics*, 11(2), 683–710.

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

Milgrom, P. (2000): “Putting auction theory to work: The simultaneous ascending auction,” *Journal of political economy*, 108(2), 245–272.

Milgrom, P., and I. Segal (2002): “Envelope theorems for arbitrary choice sets,” *Econometrica*, 70(2), 583–601.

Murota, K. (1998): “Discrete convex analysis,” *Mathematical Programming*, 83(1), 313–371.

Murota, K., et al. (2016): “Discrete convex analysis: A tool for economics and game theory,” *Journal of Mechanism and Institution Design*, 1(1), 151–273.

Nishimura, H., E. A. Ok, and J. K.-H. Quah (2016): “A Comprehensive Approach to Revealed Preference Theory,” Working Papers 201614, University of California at Riverside, Department of Economics.

Pathak, P. A., A. Rees-Jones, and T. Sönmez (2020): “Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms,” Discussion paper, National Bureau of Economic Research.

Pathak, P. A., and T. Sönmez (2008): “Leveling the playing field: Sincere and sophisticated players in the Boston mechanism,” *American Economic Review*, 98(4), 1636–52.

Pathak, P. A., T. Sönmez, M. U. Ünver, and M. B. Yenmez (2021): “Fair allocation of vaccines, ventilators and antiviral treatments: leaving no ethical value behind in health care rationing,” in *Working Paper*.

Richter, M. K. (1966): “Revealed preference theory,” *Econometrica: Journal of the Econometric Society*, pp. 635–645.

Roth, A. E. (1982): “The economics of matching: Stability and incentives,” *Mathematics of operations research*, 7(4), 617–628.

——— (1984): “Stability and polarization of interests in job matching,” *Econometrica: Journal of the Econometric Society*, pp. 47–57.

Scott, D. (1964): “Measurement structures and linear inequalities,” *Journal of mathematical psychology*, 1(2), 233–247.

Song, Y., K. Tomoeda, and X. Xia (2020): “Sophistication and Cautiousness in College Applications,” *Available at SSRN 3711118*.

Sönmez, T., and M. B. Yenmez (2022a): “Affirmative action in India via vertical, horizontal, and overlapping reservations,” *Econometrica*, 90(3), 1143–1176.

——— (2022b): “Constitutional Implementation of Affirmative Action Policies in India,” *arXiv preprint arXiv:2203.01483*.

Temnyalov, E. (2021): “An information theory of efficient differential treatment,” *Available at SSRN 3302150*.

Tuoi Tre News (2017): “Vietnamese college hopefuls miss dream school despite acing national exam,” *Tuoi Tre News*, August 4, 2017.

Tversky, A. (1964): “Finite additive structures,” *Michigan Mathematical Psychology Program*.

Weitzman, M. L. (1974): “Prices vs. quantities,” *The Review of Economic Studies*, 41(4), 477–491.

Westkamp, A. (2013): “An analysis of the German university admissions system,” *Economic Theory*, 53(3), 561–589.

Zhang, J. (2021): “Level-k reasoning in school choice,” *Games and Economic Behavior*, 128, 1–17.