REFUGEE SETTLEMENT AND OTHER MATCHING PROBLEMS WITH PRIORITY CLASSES AND RESERVES: A MARKET DESIGN PERSPECTIVE

Dilek Sayedahmed

A thesis
in
The Department
of
Economics

Presented in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy (Economics)
Concordia University
Montréal, Québec, Canada

April 2021

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Dilek Sayedahmed

Entitled: **Refugee Settlement and Other Matching Problems with Priority Classes and Reserves: A Market Design Perspective**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Economics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

Dr. Jorgen Hansen	Chair
Dr. Péter Biró	External Examiner
Dr. Ming Li	Examiner
Dr. Dipjyoti Majumdar	Examiner
	Examiner
Dr. Szilvia Pápai	Supervisor

Approved Dr. Christian Sigouin
Chair of Department or Graduate Program Director

January 29, 2021
Date of Defence

Dr. Pascale Sicotte
Dean of Faculty of Art & Science

Abstract

Refugee Settlement and Other Matching Problems with Priority Classes and Reserves:
A Market Design Perspective

Dilek Sayedahmed, Ph.D.

Concordia University, 2021

Recognizing the need to find a solution to the European refugee crisis, given the political deadlock preventing countries from participating in responsibility-sharing, this thesis addresses questions on international refugee placement from a market design perspective. Beyond refugee settlement, our theoretical findings also apply to immigration, centralized university admissions, public school choice systems, and other settings.

Following the Introduction in Chapter 1, in Chapter 2 we advocate a centralized global refugee matching system and design new matching algorithms with explicitly modelled UNHCR-mandated hierarchical priority classes of refugee families. Combined with the preferences of both sides, this set-up allows us to capture and analyse the impact of the mandated priority classes on the stability and fairness of the resulting refugee matching. We conduct an axiomatic analysis to further support the new matching designs.

Chapter 3 is motivated by the objective of designing a priority policy that effectively helps the prioritized agents. To benefit refugees in emergency zones, for example, a policy designed to create representation is not appealing, since strong candidates who would qualify without relying on their priority status may take up some or most of the reserved positions. Hence, we propose an alternative policy with targeted priority reserves, the DA-TPR mechanism, which targets those agents in the priority group who are in need of

a reserved position. We also study a general class of matching mechanisms with priority reserve policies, which includes both the DA-TPR and the primary representation rule, Hafalir et al.'s (2013) DA-MiR mechanism. The DA-TPR is the most targeted policy in this class of mechanism, which we characterize by a priority reserve stability axiom and a strong incentive property.

Chapter 4 introduces two new classes of matching mechanisms with explicitly specified partially targeted priority reserves. Both of these classes include the DA-TPR and DA-MiR mechanisms as extreme members. We identify one of these as a subclass of the characterized class of mechanisms from the previous chapter. Both studied classes of mechanisms are transparent and offer a range of policies between the DA-TPR and the DA-MiR policies, providing the designer with flexibility and clarity when choosing a priority reserve policy.

Acknowledgments

My deep gratitude goes first to my inspirational supervisor, Dr. Szilvia Pápai, who expertly and passionately guided me through my PhD and who shared the excitement of years of discovery. Her unwavering enthusiasm for market design inspired me, kept me constantly engaged with my research, and her personal generosity made my time at Concordia truly and deeply enjoyable.

I cannot begin to express how thankful I am for the FRQSC grant, which greatly supported my research and enabled me as a researcher to passionately pursue my scientific discoveries. For this, I am truly grateful.

I would like to express my deepest appreciation to my committee for spending their precious time on reviewing this work and attending the defence: Dr. Ming Li, Dr. Dipjyoti Majumdar, Dr. Péter Biró, and Dr. Anjali Awasthi. I would also like to thank my first Graduate Program Director, Dr. Effrosyni Diamantoudi, who supported me on many occasions.

I am also very thankful to my colleagues and classmates, who are now my life-long friends: Ribal Abi Raad, Rouzbeh Ghouchani, Pooya Ghasvareh, Asefeh Salarinezhad, Nickesha Ayoade, Noémie Cabau, Fadi Harb, Ehsan Fazel, Widad Achargui, Golnaz Davalloo, Liang Wang, and Olha Hnatyshyn. I am thankful to all of them for inspiring me with their hard work, for their joy and good humour.

I also want to extend my gratitude to the passionate staff who were helping me in the department and answering my countless questions. Among them: Mrs. Elise Melancon, who is truly inspirational for her commitment and hard work. Also, I want to thank Mrs. Melissa Faisal and Mrs. Sandra Topisirovic who were always helpful, passionate, and truly kind.

Above all, I am indebted to my family whose value to me only grows as the years go by. And finally, I acknowledge my partner, Dimitri Ayoub, who is always there for me with his wisdom and never-ending energy.

Contribution of Authors

Chapter 2 is not co-authored.

Chapter 3 is a joint work with my supervisor, Dr. Szilvia Pápai.

Chapter 4 is a joint work with my supervisor, Dr. Szilvia Pápai.

Table of Contents

1 Introduction **1**

1.1 Matching Theory and Market Design 1
1.2 Properties of Matching Mechanisms 4
1.3 A Prominent Matching Mechanism: Deferred Acceptance 5
1.4 Summary of the Results 7

2 Centralized Refugee Matching Mechanisms with Hierarchical Priority Classes **11**

2.1 Introduction 11
2.2 The Refugee Crisis: Background and Motivation 15
2.3 Centralized Refugee Matching: A Proposal 17
2.4 The Model 20
2.5 Weak Stability 26
2.6 Top Stability 34
2.7 Credible Stability 41
2.8 Conclusion 49

3 Targeted Priority Reserve Policies **52**

3.1 Introduction 52
3.2 Model with Priority Reserves 60
3.3 Stability Axioms 63
3.4 Deferred Acceptance with Targeted Priority Reserves 67
3.5 The Class of Sequential Priority Reserves Mechanisms 75
3.6 Selection Sequences and Consistent Selection 82
3.7 Comparing DA-SPR Mechanisms: Policy Implications 86
3.8 Policies for Compromising Between Protection and Representation 92
3.9 Conclusion 94

4 Partially Targeted Priority Reserve Policies **103**

4.1 Introduction 103
4.2 The Model 108

4.3 Deferred Acceptance Mechanisms 111
4.4 Reserved-PTPR Mechanisms 112
4.4.1 An Equivalence 114
4.4.2 Properties of Reserved-PTPR Mechanisms 119
4.5 Unreserved-PTPR Mechanisms 121
4.6 Conclusion 126

5 Bibliography **128**

Chapter 1

Introduction

1.1 Matching Theory and Market Design

Matching theory provides a new approach to study supply and demand. In commodity markets, where buyers and sellers can be anonymous, prices provide enough information to bring buyers and sellers together and determine the allocation. As opposed to these traditional markets, the new matching theory approach applies to markets where the economic agents' identities in the market matter. Who these agents are and whom they are matched to plays an important role, so prices which treat the agents anonymously cannot sort out who or what is matched to whom. Examples include internship and entry-level job markets or university admissions, where prices play little to no role in the selection process. A large portion of matching theory pertains to markets which are either centralized or should be centralized, even if they are currently decentralized. Stability and fairness of the matching mechanism, and truth-telling of the agents in the market are studied by matching theory, where participants' identities and preference rankings matter. Participants submit their preferences to the market designer whose goal is to achieve a desirable matching outcome.

Most traditional economics focuses on analysing the economy as it is. Recently economists, and in particular matching theorists, have started to design economic institutions successfully, such as student placement in school districts, labour market procedures to match medical students to hospitals, and organ donation networks. Matching theory and market design address problems in these markets and have improved real-life institutions in recent years. For example, economists have helped New York City and Boston design their school choice programs, helped medical communities reorganize their hiring procedures, and organized systematic kidney exchange programs to provide transplant kidneys to as many patients as possible.

In one-sided matching theory, one side of the market consists of participants with preferences over outcomes. Since the actual preferences are determined by the participants themselves, they have agency. Agents are typically individuals who have different characteristics, information, and tastes, which inform their preferences over different "objects." The other side of the market consists of the objects to be allocated to the agents. Objects do not have agency, but they often have mandated priorities that are imposed by law or follow norms or guidelines in the modeled environment. Objects may be schools, hospitals, countries, etc. In two-sided matching theory, both sides of the market have agents with preferences over agents on the other side of the market.

The two prominent theoretical matching models studied in the literature are the school choice model and the college admission model. Both of these models are many-to-one models that allow for a set of students to be matched to a single school or college. However, the key difference between the two theoretical models is that in school choice the schools are objects to be consumed by the students, hence it is one-sided, whereas in college admissions the colleges themselves are agents with preferences over students, hence it is two-sided.

This thesis focuses on the refugee relocation problem. In this context the school choice problem and the college admission problem are redefined as the host country choice problem and the country acceptance problem, respectively. The host country choice problem corresponds to the school choice problem. Instead of the matching of many students to a single school, the host country choice problem is the problem of matching many refugees to a single host country, where country residency permits are objects with mandated priorities and refugees are agents with preferences. Similarly, the country acceptance problem corresponds to the college admission problem. Instead of the matching of many students to a single college, the country acceptance problem is the problem of matching many refugees to a single host country, where both the countries and the refugees have preferences. As part of the overall introduction of this thesis, we provide the theoretical background on the school choice and college admissions models. Since this theoretical background directly applies to the host country choice and the country acceptance problems in the refugee allocation context, we use the refugee-country terminology in Chapter 2, while in Chapters 3 and 4 we use the more general agent-entity terminology.

It is important to highlight the widespread application of the theoretical school choice model. In many countries, children were automatically sent to a school in their neighbourhoods. Recently, more and more cities in the United States and in other countries have started to employ school choice programs: school authorities take into account the preferences of children and their parents, and there is flexibility to go to a school other than the child's district school. The typical goals of school authorities are efficient placement, fairness of the outcomes, and transparency, so participants can easily understand and use the system. Abdulkadiroğlu and Sönmez (2003)[2] showed that placement mechanisms used in several cities such as Boston are flawed: the mechanism is manipulable, i.e., students may benefit from reporting false preferences, and the result may be neither fair nor efficient. The authors proposed new mechanisms to improve upon existing placement mechanisms. Based on this and other studies, Boston and New York City changed their student placement mechanisms. For example, in New York City around 30,000 students were not matched to any of their preferred schools in the old system, but this number reduced to only 3,000 after a new mechanism was adopted.

1.2 Properties of Matching Mechanisms

The outcome of a matching market is a matching. In college admissions, a matching specifies which student attends which college. In school choice, a matching specifies which student attends which school. In refugee relocation, a matching specifies which refugee family is allocated which country's residency permit. A matching mechanism is a systematic procedure which determines a matching for every problem. In other words, a mechanism is a matching rule that produces a matching for any reported preference profile. For example, in the refugee relocation context a refugee assignment mechanism is a systematic procedure that selects a matching for each host country choice problem. Furthermore, a refugee matching mechanism requires that refugee families submit their preferences over countries, and selects a matching based on these submitted preferences and the countries' mandated refugee family priorities. A matching is Pareto-efficient if there is no other matching which assigns at least one refugee family a strictly better country while no refugee family is assigned a strictly worse country. A refugee family matching mechanism is Pareto-efficient if it always selects a Pareto-efficient matching.

Gale and Shapley (1962)[27] proposed the concept of a stable matching: a matching is stable if there is no student-school pair that are not matched to each other in the prescribed matching and want to match to each other, and there is no student that is acceptable to a school and the school has fewer students assigned to it than its quota. In the refugee relocation context, a matching is stable if there is no refugee family and a country that are not matched to each other in the prescribed matching and want to match to each other, and there is no refugee family that is acceptable to a country and the country has fewer refugee families assigned to it than its residency permit quota. More formally, a matching is stable if there is no individual player, or pairs of players, who can gain from matching outside the system. A stable matching is expected to be sustainable, and an unstable matching is expected to suffer from participants wanting to make their own matches. A mechanism is stable if it assigns a stable matching to each problem. Furthermore, we have the property of optimality which is based on stability and Pareto-efficiency in the following way. The refugee-optimal stable matching is the unique Pareto-efficient matching for refugees among the set of stable matchings. It is a stable matching that every refugee prefers at least weakly to any other stable matching. In addition, while the refugee-optimal stable matching may be Pareto-dominated for refugees, it is not dominated in the strong sense of making each refugee strictly better off.

We are also interested in incentive properties, which are properties of strategic behaviour. One of the prominent incentive properties is strategyproofness. A mechanism is strategyproof if telling the true preferences is a dominant strategy (that is, a best action no matter what others do) for everyone. Formally, if a mechanism is strategyproof then there is no agent who would be better off by reporting untruthful preferences when everyone else's preferences are fixed. For example, in the refugee allocation context, a mechanism is strategyproof if no refugee family can ever benefit by unilaterally misrepresenting its preferences.

1.3 A Prominent Matching Mechanism: Deferred Acceptance

Gale and Shapley (1962)[27] invented the Deferred Acceptance (DA) algorithm, which always results in a stable matching. The DA is an example of a mechanism.

Steps of the refugee-proposing DA rule (Gale and Shapley (1962) [27]):

In step 1:

- Each refugee family proposes to her first-choice country.
- Each country tentatively assigns its residency permits to its proposers up to its quota, following their preference order. Any remaining proposers are rejected.

In general, in any step k:

- Each refugee family who was rejected in the previous step proposes to her next-choice country.
- Each country considers the refugee families it has been holding together with its new proposers and tentatively assigns its residency permits to these refugee families up to its quota, following their preference order. Any remaining proposers are rejected.

The algorithm terminates when no refugee family's proposal is rejected, and each refugee family is assigned her final tentative assignment. We refer to the induced mechanism as the Gale-Shapley refugee optimal stable mechanism. The result of the refugee-proposing DA mechanism is the refugee-optimal stable matching. Hence, this mechanism leads to a matching for which there is no unmatched refugee family-country pair where a refugee family prefers that country to her assignment, and she has a higher priority rank than another refugee family that is assigned a residency permit at that country. Furthermore, the Gale-Shapley refugee optimal stable mechanism is strategy-proof, and it Pareto-dominates any other mechanism that is stable. Hence, telling the truth is a dominant strategy for every refugee family (Dubins and Freedman (1981)[18]; Roth (1982)[47]).

The real life applications of the DA mechanism are widespread, as the mechanism is used in many markets. Mainly, the DA is used in the American hospital-intern market. As part of the centralized DA matching mechanism in this market, known as the National Resident Matching Program (NRMP), medical students submit rank order lists over hospitals and hospitals submit rank order lists over students. At the end, the Program uses these lists to decide who works where (Roth (1984)[48]). The DA is also implemented as part of the British medical match in different regions of the UK, such as Edinburgh and Cardiff (Roth (1991)[49]). Variations of the DA are used in centralized university admissions in some countries. The Gale-Shapley agent optimal stable mechanism is used in Hong Kong to assign college seats to high school graduates (Roth and Peranson

(1997)[50]). Other markets in which the DA mechanism has been adopted for allocation are medical specialities, dental residencies, Canadian lawyer internships, and NYC high schools. Boston and other public school districts also use the DA (Abdulkadiroğlu and Sönmez (2003)[2]).

1.4 Summary of the Results

This thesis consists of three studies in matching theory and market design. Its main motivation and application is the refugee relocation problem. In order to bring the chaotic influx of refugees under control, one potential solution for countries is to participate in a centralized matching mechanism. However, not all countries can be persuaded to participate. Therefore, there is an urgent need to overcome the political deadlock preventing countries from participating in responsibility-sharing in the international refugee regime and, in particular, finding solutions to the European refugee crisis. With the goal of targeting, and therefore protecting, more refugee families, we design new matching algorithms with different underlying models to apply to the refugee relocation problem. Beyond refugee settlement, our theoretical findings also apply to immigration, centralized university admissions, public school choice systems, and other settings.

In the first study (Chapter 2), we conduct an axiomatic analysis with a focus on the stability and fairness of the refugee matching algorithm. We explicitly model and analyze countries' preferences and the prioritization of refugee families mandated for potential host countries. Having two kinds of ranking profiles for countries, mandated priority profile versus preference profile, allows us to realistically capture how the difference between these two profiles create blocking pairs of countries and refugee families due to enforced hierarchical priority classes. Therefore, we weaken the stability axiom, since a stable mechanism may no longer be stable with respect to countries' preferences leading to potential blocking pairs.

This first study contributes to the literature in multiple ways. Firstly, although countries may have their full preference rankings over refugee families, they do not need preference rankings over the entire set of refugee families to run and implement our proposed matching rules. This may make our model more efficient and thus more desirable in terms of implementation for countries, as they would only need preference rankings over refugees within the same priority class. Secondly, imposing type-specific quotas on countries in a refugee allocation setting may have its difficulties and may not be very palatable for countries. To persuade more countries to participate in a centralized refugee matching mechanism, a hierarchical priority class-based approach without type-specific set-aside quotas imposed on countries may be more palatable in terms of having more countries to willingly participate in solving the refugee matching problem. Lastly, beyond the refugee relocation context, our model has other applications, such as centralized college admissions, the design of public school choice systems, and immigration.

For the second study (Chapter 3), we recognize that affirmative action prioritizes minority students, often by using minority quotas. Since affirmative action policy is aimed at guaranteeing representation for minority students, it typically prevents a higher number of minority students from being admitted than the minority reserve quota. Even a non-wasteful matching mechanism that is more flexible than a rigid quota mechanism effectively places a cap on minority student admission. This policy serves the goal of minority representation well, but if the objective of the priority policy is to help the prioritized agents and make a difference, for example, to effectively benefit refugees in emergency zones, then it is not an appealing policy, since strong candidates who would qualify without relying on their priority status may take up most of the reserved positions.

Therefore, in the second study, motivated by the refugee relocation problem and the need for prioritizing refugee families coming from danger zones, we propose an alternative policy with targeted priority reserves, the DA-TPR mechanism, along with a corresponding new stability concept, which serve the goal of targeting those agents in the priority group who are in need of a reserved position, recognizing that not all agents in the priority group need such help. By giving an opportunity to priority agents to be matched without using up reserved positions, this matching mechanism allows for a higher number of priority agents to be selected by an entity than the reserve quota set aside for them, provided that there are strong priority agents who qualify regardless of their priority status. We show that the DA-TPR mechanism is agent-optimal among matching mechanisms that are stable with protective reserves, and prove that this rule is weakly group-strategyproof. Furthermore, we demonstrate that priority agents on the whole are better off under the DA-TPR mechanism than under the DA-MiR mechanism, while maintaining appropriate rights for non-priority agents.

We also study a general class of matching mechanisms with priority reserve policies which includes both the DA-TPR and the DA-MiR mechanisms: the DA mechanisms with sequential priority reserves (or DA-SPR mechanisms). The DA-TPR is the most targeted and the DA-MiR is the least targeted policy in this class of mechanisms, which we characterize by a weak priority reserve stability axiom and strategyproofness. Finally, we focus on a subclass of these mechanisms which stand out due to their consistent selection policies. These mechanisms are transparent and offer a range of policies between the DA-TPR and DA-MiR mechanisms, allowing the designer to have clarity and flexibility when choosing a priority reserve policy.

In the third study (Chapter 4), we continue to study Deferred Acceptance mechanisms with choice functions. This study introduces two new classes of matching mechanisms with explicitly specified partially targeted priority reserves. Both of these classes include the DA-TPR and DA-MiR mechanisms as extreme members.

In the first class, each member specifies directly for each entity the number of qualified priority group agents that are required to take up reserved positions, when choosing from the applicant pool of each entity in each step of the DA. In the second class that we study, each member specifies directly for each entity the maximum number of qualified priority group agents that are allowed to get unreserved positions, again when choosing from the applicant pool of each entity in each step of the DA. We show that the first class of mechanisms is a subclass of the characterized class of mechanisms from Chapter 3. Thus, these mechanisms satisfy the weak priority reserve-stability axiom and are weakly group-strategyproof. We also show that the second class of mechanism satisfies the weak priority reserve-stability axiom as well.

Chapter 2

Centralized Refugee Matching Mechanisms with Hierarchical Priority Classes

2.1 Introduction

There is an urgent need to overcome the political standstill preventing countries from participating in responsibility-sharing in the international refugee regime and, in particular, finding solutions to the European refugee crisis. In my prior study Sayedahmed (2017)[51] that inspired the work in this paper, I take a similar approach to Abdulkadiroğlu and Sönmez (2003)[2] in which they formulate the school choice problem as a mechanism design problem. They propose two competing mechanisms: the student optimal stable mechanism (Gale and Shapley (1962)[27]) and the top trading cycles mechanism, each providing a solution to critical school choice issues (for the literature on top trading cycles mechanism, see Shapley and Scarf (1974)[54], Pápai (2000)[43], and Abdulkadiroğlu and Sönmez (2003)[2]). I define two problems: the country acceptance problem, which is a two-sided problem with both refugee and country preferences, and the host country choice problem, which is a one-sided problem with refugee preferences and countries' predetermined priority rankings. I formulate the host country choice problem as a mechanism design problem and analyse the existing decentralized system of refugee relocation. I discuss the existing system's shortcomings and propose two competing matching rules from the mechanism design literature. I highlight each rule's strength in terms of refugee welfare and fairness of allocation, and I discuss each mechanism's contribution to solving the refugee allocation problem.

In this study I concentrate on a *country acceptance problem* and I design two new matching algorithms based on hierarchical priority classes, both of which are centralized refugee matching systems that match refugee families to countries. Throughout the paper I use the term "refugee", which refers to a "refugee family". In order to design a centralized matching system that ensures that refugee families who do not wish to be separated are not separated, the clearing-house is assumed to accept preference submissions from households, i.e, refugee families. All participating countries in the clearing-house treat refugee households as a single refugee family unit.

I conduct an axiomatic allocation analysis by focusing on the stability and fairness properties of the matching mechanisms. I explicitly model and analyse countries' preferences and the prioritization of refugee families imposed on host countries. Having two kinds of ranking profiles for countries, a *mandated priority profile* versus a preference profile, allows me to capture how the difference between the two profiles create blocking pairs of countries and refugees due to the enforced hierarchical priority classes.[1]

I recognize the importance of investigating the weakenings of stability axioms in this context. Since a stable mechanism may no longer satisfy the standard stability of the literature with respect to countries' preferences, leading to potential blocking pairs, I contribute to the literature by studying weaker stability and fairness axioms to see what type of stability and fairness properties succeed to hold in this setting. The motivation behind this is of two folds. First, I recognize that countries have their own preferences

[1]For models with distributional constraints that are studied independently from this study in the computer science and artificial intelligence literature, see Goto et al. (2016)[29].

and I take this reality into account. Second, the United Nations High Commissioner for Refugees' (the UNHCR) has strict humanitarian guidelines and principles Szobolits and Sunjic (2007)[58] for refugee settlement, which should be taken into account when designing a centralized refugee matching mechanism. Based on these guidelines, I impose priority classes on participating countries, which mandate the countries to change their preference rankings. This, however, does take us back from the goal of a fair refugee allocation result. Because of these imposed priorities for the countries which give certain refugees a priority at *each* country, I recognize the need and importance of taking into account the preferences of the refugees. A refugee in a lower category is always in that lower category. Since priority classes are mandated on all countries, every refugee in the first priority class is always prioritized over every refugee in the second priority class, and so on.

Considering the refugees in lower priority classes, I give these refugees an additional chance by looking at two different forms of priority profiles. In my first form, I give these refugees a better chance at their top ranked country and in my second form, I give them a better chance at their Deferred Acceptance (DA) matched country. Note that throughout the paper I refer to Gale and Shapley's (1962)[27] refugee-proposing DA mechanism applied to the refugee and country preferences as "the DA" in short. Moreover, in contrast to the first form, I find that the best matching the refugees in lower priority classes can get in a stable matching is their DA match, which shows the importance of prioritizing these refugees at their DA matched countries. In addition, prioritizing these refugees at their DA matched countries is more compelling for countries because each country has to modify their priorities with respect to their quotas less, compared to the first form of a priority profile that involves refugees' top ranked countries. Under the first form of a priority profile, it is possible that some countries do not have to modify the priority orders and others may have to modify much more. Consider Germany, which is a popular country for refugee settlement. Germany will be the top choice for many refugee families. Then, it will get to move many refugees up to its top priority class in which Germany ranks refugees according to its own point system, representing Germany's preferences. If all refugees rank Germany at the top, then all refugees will move up to top priority class of Germany leading to Germany ranking all refugees according to its own points-based system. When a popular country makes a lot of changes to its priority ordering, then it deviates more from the mandated priority classes.

Due to the imposed priorities, my intention is to give refugees in lower priority classes who otherwise may not have a good chance at getting matched to a country, an additional chance by lifting them using the refugee preferences. For this, I establish two new weak stability axioms: *top stability* and *credible stability*. These weak stability axioms rule out certain salient blocking pairs. For a matching to be *top stable* there can be no blocking pair such that a refugee family prefers her top ranked country to her current match and their top ranked country also prefers the corresponding refugee family to its current match. A pair blocking a matching is a *credible blocking pair* if it is matched under the DA. Blocking pairs should not be credible to achieve *credible stability* for a matching. I also define two new fairness axioms: *top priority class fairness* and *credible priority class fairness*.

Since Gale and Shapley's DA mechanism is stable, the two new mechanisms I design and propose to combine countries' preferences and enforced priority rankings are based on the DA mechanism [27]. These two mechanisms are *the Top Prioritization* and *the DA-Match Prioritization Mechanism*. The Top Prioritization Mechanism prioritizes refugee families at their top ranked countries. I show that the Top Prioritization Mechanism is top stable and top priority class fair. The DA-Match Prioritization Mechanism prioritizes refugee families at their DA matched countries, which are matched to the refugees under the DA. I show that the DA-Match Prioritization Mechanism is credibly stable and credibly priority class fair. I find that for refugees, the DA-Match Prioritization Mechanism Pareto dominates the DA. Lastly, I discuss the incentive compatibility properties of these mechanisms using Abdulkadiroğlu et al.'s (2009) results on the trade-off between efficiency and strategyproofness [1].

This study contributes to the literature in multiple ways. First, although countries may know their preferences over refugees, they do not need to know all their preferences over the entire set of refugee families to run the mechanisms that are designed in this study. This may make both the Top Prioritization and DA-Match Mechanisms more efficient, and therefore more attractive, in terms of implementation for countries, as they would only need to submit preferences over refugees in the same priority class. Second, imposing type-specific (e.g., priority class (PC)-specific) set-aside reserve quotas on countries in a refugee allocation setting may have its difficulties and may not be very palatable for countries. In order to persuade more countries to participate in a centralized refugee matching mechanism, a hierarchical priority class-based approach without category-specific set-aside reserve quotas imposed on countries may be more acceptable and may induce more countries to willingly participate in solving the refugee relocation problem. Lastly, beyond the refugee relocation context, the results of this study have other important policy applications, such as centralized college admissions, the design of public school choice systems, and immigration, where a priority-based system may work better than a category-specific reserve quota system as well.

The paper is organized as follows. In section 2, I provide a background on the refugee relocation crisis and further elaborate on my motivation. In section 3, I outline the centralized system that I propose with the UNHCR-mandated priority classes. Section 4 contains the model. In section 5, I start to investigate how we can weaken stability in the setting of hierarchical UNHCR-mandated priority classes and I provide the basic definitions of the axioms and related theorems. In sections 6 and 7, I provide the two weak stability axioms that take into account refugees' top ranked and DA matched countries when modifying countries' priorities to give more chance to the refugees who face the risk of being stuck in lower priority classes. In section 8, I conclude.

2.2 The Refugee Crisis: Background and Motivation

As the number of refugees is reaching its highest level globally since the Second World War, and Europe faces the largest mass stream of refugees in history (Alfred, 2015)[5], the international refugee regime is looking increasingly vulnerable. Faced with refugees fleeing the collapse of Syria, the European Union response has been inconsistent, poorly coordinated, and to date inadequate, given the scale of the problem. Therefore, there are a growing number of studies providing potential solutions to both international and local refugee allocation problems. In the context of local refugee match within a country, there are many obstacles for successful integration. Andersson and Ehlers (2016)[7] focus on one of these obstacles, namely the problem of finding housing for refugees once they have been relocated to a European Union country. They propose an easy-to-implement algorithm for Sweden that finds a stable maximum matching. For more studies that independently focus on different aspects of the refugee allocation problem, see Bansak et al. (2018)[12], which studies optimization of refugee preferences, and Delacrétaz et al. (2019)[15], which focuses on family size.

In the context of the international refugee allocation problem, there have been repeated calls for the revision or replacement of the Dublin Regulation, particularly its requirement that the first country in the European Union that asylum seekers enter be responsible for processing their claim ((Giuffre et al., 2015)[28]; (Koser, 2011)[41]). The current decentralized system, where the first arrival country has responsibility for refugees, is unfair to border countries such as Greece, Italy, or Hungary, and creates chaos and tragedy as European "countries play pass-the-parcel with human lives" (Jones and Teytelboym, 2017)[35]. I agree with Jones and Teytelboym (2017) that "it has never been clearer that a new deal on responsibility-sharing within Europe is needed to replace the Dublin Regulation [35]." Hence, a centralized matching system could run either alongside or, ideally, instead of the current system of the Dublin Regulation, which allows a refugee family to apply for protection in only one country. Thus, refugee families are tied to a single application to a particular country. It advantages nobody that at present refugees must play the gamble in deciding where to apply to, and countries cannot evaluate and choose from a large pool of applicants.

2.3 Centralized Refugee Matching: A Proposal

Allowing any refugee to apply to a single system from any embassy is in the interests of both refugees and countries (Jones and Teytelboym, 2016 [35] & 2017 [34]). While it is obvious that refugees would be better off with such a centralized system, counties would also gain more control, through the centralization of refugee matching. Countries would gain more control than they currently possess over who is ultimately settled within their borders. Countries also would be empowered to give their preference rankings as to which refugees they wish to accept, just as refugees give their preference rankings over countries.

A centralised matching system for refugee families would allow refugees to apply for protection in several countries and allow countries to compete for refugees. Refugees could make one claim for asylum to a single centralised body, simultaneously specifying where, if successful, they would wish to be relocated. Countries come to the clearing-house with a quota of refugees they are willing to accept and a ranking of the refugees. The system would then match refugees to countries. It is important that once a country and a refugee have been matched, this match will actually happen, which is to say that a refugee will be granted refugee status and permit in whichever country they have been matched to. A centralized system allows refugees to submit an application for asylum to *every* participating country, and they could, in principle, submit it remotely. It could include, for example, regional processing centres in the Middle East and North Africa (Jones and Teytelboym, 2016 [35] & 2017 [34]). The core advantage is that, to the extent refugee families can have confidence that there is a fair and effective system that will grant them protection without attempting a dangerous crossing, they are less likely to take that risk.

Hopefully, it should become clear in time to countries staying out of the centralized matching system that they thereby receive less-preferred refugees in comparison to countries which do participate, thus encouraging wider participation in the centralized matching system. For this reason, the centralized refugee match may also persuade countries to opt into refugee burden-sharing, surrender control of eligibility determination, and ultimately protect more refugees (Schuck, 1997)[53]. If countries opt into a centralized matching system, rather than determining the status of those refugees that have applied to them specifically, they would submit their own ranked preferences over refugees to the clearing-house, which conducts the match.

The Mandated Priority Class Hierarchy

A matching system is to allocate refugees within that system to a country where they are most likely to flourish for the duration of their residency without affecting the rest of the countries, i.e., without causing further immigration spillover to other countries [34, 35]. In particular, Jones and Teytelboym (2016)[34] highlights that "the matching system should exclude discriminatory categories such as religion and race, which are incompatible with the UNHCR's humanitarian principles, as well as being morally unacceptable, and should focus on categories of vulnerability, suitability for integration, and presence of family [34]".

I propose *mandated hierarchical priority classes* of refugees for countries (Sayedahmed, 2017)[51] based on UNHCR's 1951 Convention and Protocol for refugees. Throughout the study I refer to this UNHCR-mandated priority class hierarchy of refugees as the **enforced priority classes**. An enforced priority class of refugees is a subset of the entire set of refugees that a country must consider as part of the priority class hierarchy, which is the same for all participating countries. As part of the centralized refugee matching system, I impose this exogenously mandated priority class hierarchy. Specifically, my proposed enforced priority profile for countries is the following:

- **Priority Class I:** Refugee Families in War Zones and with the Longest Waiting Period

- **Priority Class II:** Refugee Families in War Zones

- **Priority Class III:** Refugee Families with the Longest Waiting Period

- **Priority Class IV:** Other Refugee Families

Moreover, with initial theoretical approaches to complicated real-life problems, such as the refugee crisis, it is important to start the analysis with a static model. Therefore, for its theoretical tractability, I chose to take a static approach in this study. In addition, the Syrian refugee crisis and the associated relocation problems with the Syrian war were the key motivations for this study. Due to the Syrian crisis, we have an influx of refugees. We have a given refugee pool and we can do this exercise repeatedly. On the other hand, in school choice it is clear that we have natural time periods, as we have each academic year when we naturally go through and implement school choice. Yet, it is not clear that the same also applies to refugee settlement with a sudden influx of refugees. Nonetheless, a dynamic approach could also be used here since we can do this exercise, for instance, every three months when there is a crisis. However, a refugee crisis creating a sudden large pool of refugee families, which calls for an immediate allocation, also calls for a static approach.

Furthermore, typically dynamic models take away from the nice properties of a static allocation. In dynamic models agents arrive in different time periods, and sometimes continuously, matching happens in different time periods, matched agents leave the market, and new agents arrive. When this is the case, we almost certainly lose the nice properties that we can have in a static matching. Unlike the matching environment of dating applications, which is very difficult to turn into a static matching environment due to being a very fluid and dynamic market as agents enter and leave constantly, we can turn the matching environment of refugee allocation into a static environment and retain the nice properties of stability and efficiency. When possible, we prefer to pool our agents on both sides and do repeated static matchings. Especially in an environment with a crisis that is the optimal approach, instead of taking a dynamic nature into account and keep doing matchings in each time period as refugees enter and leave the market. Hence, when we have a sudden influx of refugee families, we conduct one matching with these participants, who are fixed at the moment of the crisis when the matching takes place.

In addition, I contribute to the literature by explicitly allowing in my model for the potential real world "discrepancy" between what countries actually desire to do (i.e., countries' preferences) and what countries are mandated to do (i.e., countries' enforced priorities). I have *two* profiles for the countries rather than only one. By working with two profiles I capture the "compromise" between how countries *really* want to rank refugees in a world with no special considerations and how they are *mandated* to follow a priority hierarchy based on UNHCR's humanitarian laws and principles.

2.4 The Model

Definition 2.4.1 (**Country Acceptance Problem**). A **problem** consists of the following:

1. A finite set of refugees $\mathcal{R} = \{r_1, r_2, ..., r_R\}$.
2. A finite set of countries $\mathcal{C} = \{c_1, c_2, ..., c_C\}$.
3. A quota vector $q = (q_{c_1}, ..., q_{c_C})$, where q_c is the capacity (the number of residency permits) of $c \in \mathcal{C}$.
4. Preference profile for refugees $P = (P_{r_1}, P_{r_2}, ..., P_{r_R}) = (P_r)_{r \in R}$, where P_r is the strict preference of refugee $r \in \mathcal{R}$ over $\mathcal{C}$.
5. Preference profile for countries $\succ = (\succ_{c_1}, \succ_{c_2}, ..., \succ_{c_C}) = (\succ_c)_{c \in C}$, where $\succ_c$ is the strict preference of country $c \in \mathcal{C}$ over $\mathcal{R}$ based on its country-specific point system.[2]

[2]Unacceptable countries are not allowed in the model, in which case each refugee family $r \in \mathcal{R}$ has strict preferences over $\mathcal{C}$, where a country c is never ranked below r. Similarly, unacceptable refugees are also not allowed as part of the model, hence each country $c \in \mathcal{C}$ has strict preferences over $\mathcal{R}$, where a refugee r is never ranked below c.

A finite set of refugees $\mathcal{R}$ is with a *fixed partition* such that $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup ... \cup \mathcal{R}_T$ and $\mathcal{R}_t \cap \mathcal{R}_{t'} = \varnothing$ for any $t, t' \in \{1, ..., T\}$). The T number of **enforced hierarchical priority classes** of the partition are class $\mathcal{R}_1$, class $\mathcal{R}_2$, and so on until class $\mathcal{R}_T$. These are the enforced priority classes introduced and discussed in the previous section.

As a primitive of my model, I also define the following.

Definition 2.4.2 (Enforced Priorities). Countries' enforced priority profile $\pi^E = (\pi^E_c)_{c \in \mathcal{C}}$ is based on the fixed partition $\mathcal{R}_1, ..., \mathcal{R}_T$ of the refugees into enforced priority classes.

Let $t, t' \in \{1, ..., T\}$ such that $t \leq t'$ and let $r_i \in \mathcal{R}_t$, $r_j \in \mathcal{R}_{t'}$. Then for all $c \in \mathcal{C}$,

(a) If $t = t'$, then $r_i \; \pi^E_c \; r_j$ if and only if $r_i \succ_c r_j$.

(b) If $t < t'$, then $r_i \; \pi^E_c \; r_j$.

For item (a) above, observe that each country's rankings within each priority class are country-specific preferences. For any $c \in \mathcal{C}$ with $r_i \succ_c r_j$, if r_i and r_j are prioritized within the *same* class, then *within* that priority class, country c's ranking will also be $r_i \; \pi^E_c \; r_j$ following country c's preference ranking. For item (b) above, observe that if a refugee is in a higher priority class than another refugee, then this implies that according to π^E, the refugee in the higher priority is strictly prioritized over the refugee who is in a lower priority class.

Remark. *In partitioned enforced priorities, each member of the partition corresponds to an exogenously imposed priority class. The priority classes are mandated across all participating countries and each mandated priority class consists of the same set of refugees for all participating countries.*

In summary, a many-to-one **country acceptance problem**, where each refugee family can be matched to at most one host country and each country can admit at most q_c refugee families, is defined as the tuple

$$\langle \mathcal{R} = \bigcup_{t=1}^{T} \mathcal{R}_t, \mathcal{C}, (q_c)_{c \in \mathcal{C}}, P, \succ \rangle.$$

When all other parameters, except the refugees' preference profile P, are fixed, I refer to P as a country acceptance problem. Let $\mathcal{P}$ denote the set of country acceptance problems.

In order to simplify the exposition, I assume that for all $c \in \mathcal{C}$, $|\mathcal{R}| > q_c$. This is a very mild assumption which is easily satisfied in any application and allows me to waive the case where there are few refugees. I will also need the following notation. Let R_r denote the weak preferences of refugee $r \in \mathcal{R}$ associated with P_r. Since preferences are strict, c R_r c' means that either c P_r c' or $c = c'$. The preferences of a coalition $L \subseteq \mathcal{R}$ in P are denoted by P_L. Finally, I denote the preference profile of all the refugees except for r by P_{-r}, and the preference profile of all refugees except the ones in coalition L by P_{-L}.

For simplicity, I assume that countries have *responsive preferences* over refugee families, which means, roughly speaking, that refugee families are not complements in the countries' preferences, and thus preferences over sets of refugee families can be interpreted as a natural extension of the preferences over individual refugees.

Definition 2.4.3 (Responsive Preferences). For any set of refugee families $\mathcal{Z} \subset \mathcal{R}$ with $|\mathcal{Z}| \leq q_c$, and any refugee families r and r' in $\mathcal{R} \setminus \mathcal{Z}$,

- $\mathcal{Z} \cup \{r\} \succ_c \mathcal{Z} \cup \{r'\}$ if and only if $r \succ_c r'$.

Note the slight abuse of notation in the above definition. To indicate preferences over sets $\succ_c$ is used which is normally used for showing preferences over singletons in $\mathcal{R}$. Responsive preferences are a natural extension of preferences over individuals to preferences over sets. This property does not give a complete ordering of all sets of size q_c for a country, as it does not pin down all preference rankings over sets. However, note that this does not affect the analysis in this study. The missing preference orderings over sets are not necessary to pin down here, and they can be in any order.

Definition 2.4.4 (**Matching**). A solution to a many-to-one refugee matching problem is μ: $\mathcal{R} \bigcup \mathcal{C} \rightarrow \mathcal{R} \bigcup \mathcal{C}$, a correspondence from $\mathcal{R} \bigcup \mathcal{C}$ to $\mathcal{R} \bigcup \mathcal{C}$ such that, for every refugee family $r \in \mathcal{R}$ and participating country $c \in \mathcal{C}$:

- $\mu(r) \in \mathcal{C}$
- $\mu(c) \subseteq \mathcal{R}$ and $|\mu(c)| \leqslant q_c$
- $\mu(r) = c \Leftrightarrow r \in \mu(c)$

Let μ and ν be two matchings. A matching μ **Pareto dominates** matching ν at preference profile P if for all refugees $r \in \mathcal{R}$, $\mu_r R_r \nu_r$, and there exists a refugee r such that $\mu_r P_r \nu_r$. A matching μ **weakly Pareto dominates** matching ν at P if either μ Pareto dominates ν or μ is the same as ν.

The set of problems is denoted with $\mathcal{P}$, which is $\langle \mathcal{R} = \bigcup_{t=1}^{T} \mathcal{R}_t, \mathcal{C}, (q_c)_{c \in \mathcal{C}}, P, \succ \rangle$. Let $\mathcal{M}$ be the set of matchings. Since everything else except P is fixed, I define a mechanism as follows.

Definition 2.4.5 (**Mechanism**). A mechanism is a mapping that assigns a matching to each country acceptance problem $P \in \mathcal{P}$. Formally, a mechanism is a mapping $f : P \rightarrow \mathcal{M}$.

Let f and g be two mechanisms. A **mechanism** f **Pareto dominates mechanism** g for $\mathcal{R}' \subseteq \mathcal{R}$ if for all profiles $P \in \mathcal{P}$, $f_{\mathcal{R}'}(P)$ weakly Pareto-dominates $g_{\mathcal{R}'}(P)$, and there exists $\bar{P} \in \mathcal{P}$ such that $f_{\mathcal{R}'}(\bar{P}) \neq g_{\mathcal{R}'}(\bar{P})$. If **mechanism** f **weakly Pareto dominates mechanism** g, then at every P, mechanism f produces a matching that weakly Pareto dominates the matching produced by g.

Since in this study I focus on the refugee-proposing DA and its modifications, matching results obtained differ with respect to the type of a country profile that is being used. A refugee assignment mechanism requires refugees to submit preferences over countries, and selects a matching based on these submitted preferences and refugee priorities. Note that the notation μ^{DA} is used for the matching result obtained from the refugee-proposing DA applied to P and $\succ$.

Definition 2.4.6 (**Blocking Pairs**)**.** A matching μ is blocked by a refugee-country pair $(r, c) \in \mathcal{R} \times \mathcal{C}$ if they prefer each other compared to μ:

1. the refugee family r prefers c to the country it is matched to in μ (i.e., $c \ P_r \ \mu(r)$), and

2. given $r \notin \mu(c)$,

   a. either the country prefers r to some refugee r' that the country is matched to in μ (i.e., $\{r\} \succ_c \{r'\}$ where $r' \in \mu(c)$)

   b. or refugee r is acceptable to country c and the country has fewer refugee families assigned to it than its quota (i.e., $\{r\} \succ_c c$ and $|\mu(c)| < q_c$).

For the definitions, recall that I use the term "refugee" in this study when referring to a "refugee family". Furthermore, following Abdulkadiroğlu and Sönmez (2003)[2], we can state a matching to be **fair** if there is no unmatched refugee–country pair where the refugee prefers the country to her current match and she has higher priority than some other refugee, who is assigned a residency permit at the country. In addition to this also requiring **non-wastefulness**, which requires not wasting an available permit at the country gives the **stability** notion in the two-sided matching context.

Next, I formally define stability using the blocking pair definition above.

Definition 2.4.7 (Stability). A matching is stable if it is not blocked by a refugee-country pair. A mechanism is stable if it assigns a stable matching to each country acceptance problem P.

Definition 2.4.8 (Refugee-Proposing Deferred Acceptance (DA) Mechanism (Gale and Shapley, 1962)**).**

Step 1: Each refugee proposes to her first choice country. Each country tentatively assigns its refugee residency permits to its proposers up to its quota following its priority order. Any remaining proposers are rejected.

In general, at

Step k: Each refugee who was rejected in the previous step proposes to her next choice country. Each country considers the refugees it has been holding together with its new proposers and tentatively assigns its refugee residency permits to these refugees up to its quota following its priority order. Any remaining proposers are rejected.

The mechanism terminates when no refugee proposal is rejected and each refugee is assigned her final tentative assignment. We refer to this mechanism as the *Gale-Shapley refugee-optimal stable mechanism*.

Note that Gale and Shapley (1962)[27] calls a stable mechanism *optimal* if every refugee is at least as well off under it as under any other stable matching. Furthermore, the DA procedure yields not only a stable but also an optimal matching [27]. Every refugee is at least as well off under the matching given by the DA mechanism as she would be under any other stable matching. This holds for both sides: refugees and countries. For every country acceptance problem, there exists a refugee-optimal stable matching, which is liked by each refugee at least as well as any other stable matching and also there exists a country-optimal stable matching, which is liked by each country at least as well as any other stable matching. The refugee-proposing DA mechanism leads to the refugee-optimal stable mechanism.

Definition 2.4.9 (The DA with Hierarchical Priority Classes). The DA with Hierarchical Priority Classes is the refugee-proposing DA applied to the mandated priority profile of countries π^E and refugee preference profile P.

Throughout the paper, I denote the DA's matching result with μ^{DA}. Similarly, I denote the matching result of the DA with Hierarchical Priority Classes with μ_E^{DA}.

2.5 Weak Stability

When mandated priority class profile π^E is used, there may exist blocking pairs with respect to country preferences $\succ$ and hence a stable mechanism with respect to π^E may not be stable with respect to $\succ$. Therefore, when a refugee family forms a blocking pair with a country, this blocking is always going to be with respect to another refugee family that is in a higher enforced priority class than the refugee family that is forming a blocking pair. The basis for blocking is the priority reversal that results due to the mandated (enforced) priority classes. This observation inspires the following axioms. Note that in my axioms the "PC" is an abbreviation for "priority class".

Next I define the first weak fairness axiom of this study.

Definition 2.5.1 (PC Fairness Axiom). A matching μ satisfies PC fairness at a particular profile P if, for every refugee-country pair (r, c) such that r prefers c to his assignment at μ and is preferred by c to a refugee $\hat{r}$ assigned to c at μ, $\hat{r}$ is in a higher priority class than r. A mechanism f satisfies PC fairness if, for every profile P, it assigns a matching μ that is PC fair.

According to PC fairness, if a refugee-country blocking pair (r, c) is such that r is preferred by c to a refugee $\hat{r}$ assigned to c then it must be the case that $\hat{r}$ is in a higher priority class than r.

Definition 2.5.2 (PC No-Envy Axiom). A matching μ satisfies PC no-envy at a particular profile P if $\hat{r}$ is in a higher enforced priority class than r, then $\hat{r}$ does not have envy for r, i.e., if $r \in \mu(c)$ then $\mu(\hat{r})\ R_{\hat{r}}\ c$. A mechanism f satisfies PC no-envy, if for every profile P, it assigns a matching μ that satisfies PC no-envy.

In order to demonstrate the independence of the PC no-envy and PC fairness axioms, first observe that given a fixed profile P, the matching result of the DA, specifically μ^{DA}, is fair, and thus it satisfies PC fairness. However, μ^{DA} does not satisfy PC no-envy, as it is not stable with respect to the enforced priority classes of π^E at each profile P. Hence, PC fairness does *not* imply PC no-envy. Consider the following example to intuitively visualize this.

Example 1. We are given the below problem P with refugees $\mathcal{R} = \{1, 2, 3, 4\}$, countries $\mathcal{C} = \{a, b, c, d\}$, and country quotas $q_c = 1$ for all $c \in \mathcal{C}$. Allocations of the matching result μ^{DA} at the given preference profile in this example are **bold.**

P_1	P_2	P_3	P_4
c	b	c	b
b	c	b	**c**
d	**d**	**a**	a
a	a	d	d

Refugee Preferences P

$\succ_a$	$\succ_b$	$\succ_c$	$\succ_d$
1	1	4	4
4	3	2	3
2	2	1	1
3	4	3	2

Country Preferences $\succ$

π_a^E	π_b^E	π_c^E	π_d^E
1	**1**	2	1
2	2	1	**2**
4	3	**4**	4
3	4	3	3

Enforced Profile π^E

Applying the refugee-proposing DA algorithm to $\succ$ and P gives us the following:

$$\mu^{DA} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ b & d & a & c \end{pmatrix}$$

The matching μ^{DA} satisfies PC fairness because it satisfies fairness. However, matching

μ^{DA} does not satisfy PC no-envy. This is because $(1, c)$ and $(2, c)$ are blocking pairs for μ^{DA} with respect to π^E. Refugees 1 and 2 are envious of refugee 4 who is in a lower priority class than both of the refugees and matched to c in μ^{DA}.

Furthermore, PC no-envy does *not* imply PC fairness. The matching result of the DA with Hierarchical Priority Classes, specifically μ_E^{DA} is stable with respect to the enforced priority classes of π^E at each profile P, and hence satisfies PC no-envy. Observe that PC no-envy is satisfied whenever a serial dictatorship procedure f with the permutation of agents based on the hierarchical priority classes is used. Consider an arbitrary priority ordering of agents and let this common ordering be denoted by π. Recall that given an ordering π of agents with any permutation of the entire set of agents, a serial dictatorship $f(\pi)$ (Hylland and Zeckhauser (1979)[33]) assigns the objects to agents as follows. The first agent is assigned her first choice among all the objects. The second agent is assigned her first choice among all the objects, excluding the choice of the first agent, and so on. Now consider an ordering π of refugees that follows the fixed hierarchy of priority classes and within priority classes any order of refugees is possible as long as it is the same for each country. Then the PC no-envy property is satisfied whenever $f(\pi)$ is used since $f(\pi)$ assigns the permits to refugees following the given common order. Consider the following example to see this visually.

Example 2. We are given the below problem P, where refugees $\mathcal{R} = \{1, 2, 3, 4\}$, countries $\mathcal{C} = \{a, b, c, d\}$, and country quotas $q_c = 1$ for all $c \in \mathcal{C}$. Let π be the permutation. Allocations of the result of the serial dictatorship f that uses the given common priority order π are in **bold.**

P_1	P_2	P_3	P_4
a	**a**	**c**	b
b	c	b	c
c	d	a	a
d	b	d	**d**

Refugee Preferences P

π_a	π_b	π_c	π_d
2	2	2	2
1	1	1	1
3	3	3	3
4	4	4	4

Common Priority Order π

π_a^E	π_b^E	π_c^E	π_d^E
1	2	2	1
2	**1**	1	2
4	4	**3**	**4**
3	3	4	3

Enforced Profile π^E

The result of the serial dictatorship $f(\pi)$ is the following:

$$f(\pi) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ b & a & c & d \end{pmatrix}$$

Observe that $f(\pi)$ satisfies PC no-envy as no refugee in a higher enforced priority class has envy for a refugee who is in a lower priority class. However, it does not satisfy PC fairness. This is because $(1, a)$ is a blocking pair for $f(\pi)$ with respect to π^E. Refugee 1 is envious of 2, who is preferred by a over 2 according to country a's preferences within the top priority class of a. Hence, country preferences are violated by $f(\pi)$ within the priority class. Therefore, whenever $f(\pi)$ is applied, PC no-envy is satisfied since it is not possible to have envy towards a refugee in a lower priority class. However, the result $f(\pi)$ fails PC fairness, as any order of refugees within a priority class is possible.

Proposition 1. *Stability of DA with the enforced hierarchical priority classes is equivalent to PC no-envy and PC fairness.*

Proof. **"If" Part.** We show that stability of DA with the enforced priority classes implies PC no-envy and PC fairness. Let μ_E^{DA} be the matching result of the DA with Hierarchical Priority Classes at a given problem P. Given that the DA with Hierarchical Priority Classes is the DA applied to enforced priorities π^E, then μ_E^{DA} is stable with respect to π^E at any given preference profile. Then, there is no refugee-country pair (r, c) blocking μ_E^{DA} at any given preference profile. This implies two cases.

Case 1. There is no refugee-country pair (r, c) such that r prefers c to his assignment and is preferred by c according to c's own preference to a refugee $\hat{r}$ assigned to c, where r and $\hat{r}$ are in the same enforced priority class. Hence, there is not ranking violation within a priority class and PC fairness is satisfied.

Case 2. There is no refugee-country pair (r, c) such that r prefers c to his assignment and is ranked above by c according to c's priority ranking in π^E to a refugee $\hat{r}$ assigned to c, where r and $\hat{r}$ are *not* in the same enforced priority class. Then, there is no ranking violation across priority classes. Specifically, there is no r who is in a higher priority class than $\hat{r}$ and has envy for $\hat{r}$. Hence, PC no-envy is satisfied.

Therefore, stability of DA with the enforced priority classes of π^E implies PC no-envy and PC fairness.

"Only If" Part. We show that satisfying PC fairness and PC no-envy implies satisfying stability with respect to the enforced priority classes.

Case 1. Suppose PC fairness is satisfied. Then, if a refugee-country pair (r, c) is such that r prefers c to his assignment and is preferred by c according to c's own preference to a refugee $\hat{r}$ assigned to c then it must be that $\hat{r}$ is in a higher priority class than r. This implies $\hat{r}$ and r are not in the same priority class. Hence, there is no refugee-country pair (r, c) such that r prefers c to his assignment and is preferred by c according to c's own preference to a refugee $\hat{r}$ assigned to c. This satisfies stability within each enforced priority class.

Case 2. Suppose PC no-envy is satisfied. Then, for each preference profile P, if r is in a higher enforced priority class than $\hat{r}$, then r does not have envy for $\hat{r}$. Hence, for any given preference profile P, if $\hat{r} \in \mu_E^{DA}(c)$ then $\mu_E^{DA}(r)\ R_r\ c$. Therefore, there is no refugee-country pair (r, c) such that r prefers c to his assignment and is ranked above by c according to c's priority ranking in π^E to a refugee $\hat{r}$ assigned to c. This satisfies stability across enforced priority classes.

Therefore, PC fairness and PC no-envy imply stability with respect to π^E. In conclusion, satisfying the two axioms of PC fairness and PC no-envy is equivalent to satisfying stability with respect to enforced priority classes. $\square$

Next, I state the first corollary of Proposition 1.

Corollary 1. *The DA with Hierarchical Priority Classes satisfies*

1. *PC fairness and*

2. *PC no-envy.*

Proof. This follows from the equivalence established between stability with respect to π^E, and PC fairness and PC no-envy. This can also be shown by verifying for the two axioms.

1. PC fairness: Fix a profile P. Let μ_E^{DA} be the matching result of the DA with Hierarchical Priority Classes at a given preference profile P. If $\pi^E = \succ$ then the result holds. Assume π^E is different from $\succ$. Then, for contradiction, suppose μ_E^{DA} is not PC fair. Then there exists a blocking pair (r, c) at $\succ$ and $\hat{r} \in \mathcal{R}$ such that $\hat{r} \in \mu_E^{DA}(c)$ and $\hat{r}$ is not in a higher enforced PC than r and $r \succ_c \hat{r}$. If r and $\hat{r}$ are in the same PC, then $r \succ_c \hat{r}$ and $\hat{r}\ \pi_c^E\ r$ since $\hat{r} \in \mu_E^{DA}(c)$. Then PC fairness is violated. If $\hat{r}$ is in a lower PC than r, then the priority class hierarchy of π^E is violated by the DA with Hierarchical Priority Classes at P, which is a contradiction.

2. PC no-envy: Fix a preference profile P. For contradiction, suppose μ_E^{DA} does not satisfy PC no-envy. Then there exists at least one $\hat{r}$ who is in a higher enforced priority class than r and $\hat{r}$ has envy for r at P. This implies priority ranking violation of $\hat{r}$ at π^E. Then, μ_E^{DA} is not stable at π^E. Since μ_E^{DA} is the matching result of the DA with Hierarchical Priority Classes at P, this is a contradiction.

Hence, for all P, the DA with Hierarchical Priority Classes satisfies PC fairness and PC no-envy. $\square$

In the following I will call a matching *optimal with respect to a stability axiom* if every agent receives at least as good an assignment in this matching as in any other matching satisfying the stability axiom. Specifically, I will call a **matching** μ is **optimal with respect to PC fairness and PC no-envy** at profile P if, for each refugee $r \in \mathcal{R}$, $\mu(r) = c$ is the most preferred country among all countries that refugee r could be matched to at any matching that satisfies PC fairness and PC no-envy at P, when there is any such country. Given $P \in \mathcal{P}$, a matching **mechanism** is **optimal with respect to PC fairness and PC no-envy** if, for each preference profile $P \in \mathcal{P}$, it assigns a matching to profile P that is optimal with respect to PC fairness and PC no-envy. Next, we have the second corollary of Proposition 1.

Corollary 2. *The DA with Hierarchical Priority Classes is optimal with respect to PC fairness and PC no-envy.*

Proof. This is a direct corollary of Proposition 1. It follows from the Pareto efficiency of the refugee-optimal matching with the priorities π^E. □

Furthermore, there is a unique mechanism that is optimal with respect to PC fairness and PC no-envy, and it is the DA with Hierarchical Priority Classes.

Proposition 2. *If a mechanism is PC fair, PC no-envy, and is optimal with respect to PC fairness and PC no-envy, then it is the DA with Hierarchical Priority Classes.*

Proof. Since optimality implies uniqueness, this follows from Proposition 1, and the two corollaries of Proposition 1. □

It is important to discuss this result intuitively in terms of the two key weak stability and fairness axioms studied in this paper. PC no-envy of any arbitrary matching mechanism implies that a refugee $\hat{r}$ who is in a higher enforced priority class does not have envy for any of the allocations below $\hat{r}$ at any profile P. This implies that the allocation starts at the first priority class and follows a serial procedure of priority classes according to a given ordering of refugees that has the fixed hierarchy of priority classes, where within priority classes any order of refugees is possible. We allocate to the top priority class refugees, then allocate to the second priority class refugees, and so on, which is equivalent to the serial dictatorship procedure with the permutation of agents based on the hierarchical priority classes. Once the top priority class allocation is done, those allocations are removed as they are final. After, allocations are done for the second priority class refugees, and so on. Instead of individual agents picking their top choice, we have priority classes of refugees who get their top choice independent of the preferences of refugees in other priority classes.

In addition to PC no-envy, the requirements of PC fairness and optimality with respect to PC no-envy and PC fairness for a matching mechanism imply that for any problem P, within each priority class the match must be fair. Thus, for the allocation within each priority class to be fair, the DA must be applied to each priority class. Therefore, since the DA applied to P with a given π^E is the DA with Hierarchical Priority Classes, and π^E is a partition profile, it is equivalent to the DA applied to the top priority class than the DA applied to the second priority class, and so on, following the given hierarchical priority class ordering. Therefore, if a mechanism satisfies PC fairness, PC no-envy, and is optimal with respect to PC fairness and PC no-envy, then it is the DA with Hierarchical Priority Classes.

Theorem 1. *A mechanism satisfies PC fairness, PC no-envy, and is optimal with respect to PC fairness and PC no-envy if and only if it is the DA with Hierarchical Priority Classes.*

Proof. **"If" Part.** This follows from the Corollary 1 and Corollary 2 of the equivalence result in Proposition 1.

"Only If" Part. This is by Proposition 2. □

2.6 Top Stability

I recognize that the imposed priority classes force the countries to change their preference rankings and this takes us back from the goal of a fair refugee allocation result. Because of these imposed priorities for the countries which give certain refugees a priority at each country, I recognize the need to take into account the refugees' preferences. A refugee in a lower priority class is always in that priority class. Since priority classes are mandated on all countries, every refugee in the first priority class is always prioritized over every refugee in the second priority class. Considering the refugees in lower priority classes, I give these refugees a better chance at their top ranked country by lifting them up using their preferences. In order to pursue this I recognize the need to establish weak stability axioms that rule out certain salient blocking pairs. For interesting relevant notions, which were studied independently and in different contexts, see Pathak and Sönmez (2013)[45], Dur, Mennle, and Seuken (2018)[20], Afacan, Alioğulları, and Barlo (2015)[3], and Morrill (2013)[42].

Note that for the definitions of my weak stability axioms it is important to recall Definition 2.4.6 of blocking pair (r, c) in Section 2.4.

Definition 2.6.1 (**Top Stability Axiom**)**.** A matching is top stable if it cannot be blocked with a pair (r, c), where P_r ranks country c first at a fixed preference profile P. A mechanism is top stable if, for each preference profile P, whenever the matching assigned to P is blocked by (r, c), P_r does not rank c first.

Definition 2.6.2 (**Top PC Fairness Axiom**)**.** A matching μ satisfies top PC fairness if, for every refugee-country (r, c) such that r prefers c to his assignment at μ and is preferred by c to a refugee $\hat{r}$ assigned to c at μ,

- either $\hat{r}$ is in a higher enforced priority class than r,
- or $P_{\hat{r}}$ ranks c first and r is not in priority class $\mathcal{R}_1$ of country c.

A mechanism f satisfies top PC fairness if, for every profile P, it assigns a matching μ that is top PC fair. Note that whenever there is justified envy with respect to $\hat{r}$, either $\hat{r}$ is in a higher enforced class than r with justified envy, or $P_{\hat{r}}$ ranks c first and r is not in top priority class of country c.

In order to accommodate more refugees in lower priority classes by raising them to the first priority class at their top ranked country, I further weaken the PC fairness axiom. Therefore, PC fairness axiom implies top PC fairness. Since I make exceptions for these refugees at their top choice country and to account for their exceptions of gained rank, I declare blocking pairs that involve such refugees and their top choice country to be salient and therefore not allowed.

Countries' combined priority profile is obtained by combining π^E and P.

Definition 2.6.3 (Combined Priorities). The combined priority ranking profile for countries is obtained by combining $\succ$ and π^E, given a certain mechanism that specifies the combination procedure using refugee preferences P.

Remark. *The combined priority profile of countries is dependent on the refugee preference profile P. By contrast, the enforced priority profile of countries is independent of the refugee preference profile P.*

Definition 2.6.4 (The Top Prioritization Mechanism).

1. For every refugee r modify π_c^E by lifting refugee r up to priority class $\mathcal{R}_1$ of c, where c is top-ranked in P_r.

2. Position r within priority class $\mathcal{R}_1$ according to $\succ_c$ (i.e., leave all other priority rankings the same, except for r's). This leads to the new combined profile π^{E-top}.

3. Apply the DA to π^{E-top}. This leads to the matching result $f^{DA}_{E-top}(P)$ of the Top Prioritization Mechanism f^{DA}_{E-top} at problem P.

Similar to the other notations for the matching outcomes, I use μ^{DA}_{E-top} for the matching result of the Top Prioritization Mechanism f^{DA}_{E-top} at any given problem P.

Furthermore, given an ordering π_c, let $S^{\pi}_c(r)$ denote the *upper contour set* at r. The, the upper contour set of refugee r in country c's profile is the following:

$$S^{\pi}_c(r) = \{\hat{r} \in \mathcal{R} : \hat{r}\ \pi_c\ r\ \}.$$

Lemma 1. *If P_r ranks country c first, then $S^{\pi^{E-top}}_c(r) \subseteq S^{\succ}_c(r)$.*

Proof. Let r and c be such that P_r ranks country c first. Then, r is lifted to $\mathcal{R}_1$ in π^{E-top}_c. Fix $\hat{r} \in S^{\pi^{E-top}}_c(r)$. Then, we have $\hat{r} \in \mathcal{R}_1$. Moreover, given the construction of π^{E-top}, this means that $\hat{r} \succ_c r$. Thus, $\hat{r} \in S^{\succ}_c(r)$. □

Moreover, the Top Prioritization Mechanism guarantees top stability of μ^{DA}_{E-top} at any P and thus we have the following result.

Theorem 2. *The Top Prioritization Mechanism is*

1. top stable and

2. top PC fair.

Proof.

1. Top stability: Fix a given preference profile P. For contradiction, suppose the matching result μ^{DA}_{E-top} is not top stable at the given profile P. Then there is a pair (r, c) blocking μ^{DA}_{E-top} with respect to $\succ$ at P, where refugee r ranks c first at preference profile

P. Given that the DA is applied to combined profile π^{E-top}, if country c rejects r in a round, then c is temporarily matched to at least one refugee other than r. Let this refugee be $\hat{r}$. Then, this implies $\hat{r}\ \pi_c^{E-top}\ r$.

Since r ranks c first at preference profile P then r was lifted to priority class $\mathcal{R}_1$ by country c in π_c^{E-top} by the Top Prioritization procedure. By Lemma 1, if r is in the top priority class of π_c^{E-top} and $\hat{r}\ \pi_c^{E-top}\ r$, then r and $\hat{r}$ are both in the top priority class. This contradicts the assumption that (r, c) is a blocking pair of μ_{E-top}^{DA} with respect to $\succ$. Hence, μ_{E-top}^{DA} is top stable.

2. Top PC fairness: Fix a given preference profile P. For contradiction, suppose μ_{E-top}^{DA} is not top PC fair at the given profile P. Then at the given P, there exists a salient pair (r, c) that blocks μ_{E-top}^{DA} with respect to $\succ$ that is *not* allowed under top PC fairness and there is $\hat{r} \in \mathcal{R}$ such that $\hat{r} \in \mu_{E-top}^{DA}(c)$. The violation of the top PC fairness of the matching result μ_{E-top}^{DA} at given P implies four non-trivial cases. We show contradiction under each case.

Observe first that when refugee $\hat{r}$ is not in a higher enforced priority class than r who is assumed to be blocking with c at $\succ$, then they are in the same priority class. Albeit this is trivial, it is still a possibility since the matching result is assumed to be not top PC fair at the given profile. However, this violates country preferences within same priority class, which are assumed to be preserved. Hence, this is a contradiction.

Case 1: Suppose $P_{\hat{r}}$ does not rank c first and r is in priority class $\mathcal{R}_1$ of country c. Then, either $\hat{r}$ is in a lower priority class than r, or $\hat{r}$ is already in $\mathcal{R}_1$ and hence both r and $\hat{r}$ are in $\mathcal{R}_1$ of country c. If $r \in \mathcal{R}_1$ of c and $\hat{r}$ is in a lower priority class, then this contradicts $\hat{r}\ \pi_c^{E-top}\ r$ since $\hat{r} \in \mu_{E-top}^{DA}(c)$. If both $r, \hat{r} \in \mathcal{R}_1$, then it must be that $r \succ_c \hat{r}$ since for contradiction we assumed the existence of a blocking pair (r, c) with respect to $\succ$. Country preferences are preserved within same priority classes. Hence, $r\ \pi_c^{E-top}\ \hat{r}$. However, we also have $\hat{r}\ \pi_c^{E-top}\ r$ since $\hat{r} \in \mu_{E-top}^{DA}(c)$. This is a contradiction.

Case 2: Suppose $P_{\hat{r}}$ ranks c first and r is in priority class $\mathcal{R}_1$ of country c. Then,

$r, \hat{r} \in \mathcal{R}_1$, then $r \succ_c \hat{r}$, which is preserved in $\mathcal{R}_1$ hence $r \ \pi_c^{E-top} \ \hat{r}$. However, this contradicts $\hat{r} \ \pi_c^{E-top} \ r$ given $\hat{r} \in \mu_{E-top}^{DA}(c)$.

Case 3: Suppose $P_{\hat{r}}$ does not rank c first and r is not in priority class $\mathcal{R}_1$ of country c. If $\hat{r}$ is already in mandated class $\mathcal{R}_1$, then $\hat{r} \ \pi_c^{E-top} \ r$, then there is nothing to prove as μ_{E-top}^{DA} is top PC-fair. However, consider otherwise. Then, neither of the two refugee families are in $\mathcal{R}_1$.

Hence, since we assumed μ_{E-top}^{DA} is not top PC fair, either r and $\hat{r}$ are in the same priority class, or $\hat{r}$ is in a lower priority class than r. If they are both in the same priority class, then (r, c) blocking μ_{E-top}^{DA} at $\succ$ implies that $r \succ_c \hat{r}$, which is preserved in the same priority class, then $r \ \pi_c^{E-top} \ \hat{r}$. This is a contradiction to $\hat{r} \ \pi_c^{E-top} \ r$ given $\hat{r} \in \mu_{E-top}^{DA}(c)$. If $\hat{r}$ is in a lower priority class than r, then r having envy for $\hat{r}$ who is in a lower priority class is a contradiction to PC no-envy property of the match. This is because if r is in a higher priority class than $\hat{r}$ then r does not have envy for $\hat{r}$ at preference profile P.

Case 4: Suppose $P_{\hat{r}}$ ranks c first and r is not in priority class $\mathcal{R}_1$ of country c. Then (r, c) is a salient blocking pair that is allowed under top PC fairness. Hence, this is a contradiction.

Therefore, the Top Prioritization Mechanism is top stable and top PC fair. □

Example 3. Consider the given refugee families $\mathcal{R} = \{1, 2, 3, 4, 5, 6\}$, countries $\mathcal{C} = \{a, b, c, d\}$, quota vector $q = (1, 1, 2, 2)$, and mandated priority classes $\mathcal{R}_1 = \{1, 2\}$ and $\mathcal{R}_2 = \{3, 4, 5, 6\}$. Since countries c and d have two residency quotas each, let there be two identical copies of country c each with preferences $\succ_c$ and capacity one, and similarly, let there be two identical copies of d each with $\succ_d$ and capacity one. Allocations that are underlined are for the matching result μ_{E-top}^{DA} of the Top Prioritization Mechanism.

P_1	P_2	P_3	P_4	P_5	P_6
d	d	$\underline{d}$	$\underline{d}$	a	a
$\underline{a}$	$\underline{b}$	c	c	d	d
b	c	a	a	$\underline{c}$	$\underline{c}$
c	a	b	b	b	b

Refugee Preferences P

$\succ_a$	$\succ_b$	$\succ_c$	$\succ_d$
$\underline{1}$	$\underline{2}$	3	5
5	1	4	6
6	5	$\underline{5}$	$\underline{3}$
2	6	$\underline{6}$	$\underline{4}$
3	3	2	1
4	4	1	2

Country Preferences $\succ$

π_a^E	π_b^E	π_c^E	π_d^E
1	2	2	1
2	1	1	2
5	5	3	5
6	6	4	6
3	3	5	3
4	4	6	4

Enforced Profile π^E

π_a^{E-top}	π_b^{E-top}	π_c^{E-top}	π_d^{E-top}
1	2	2	3
5	1	1	4
6	5	3	1
2	6	4	2
3	3	5	5
4	4	6	6

Combined Profile π^{E-top}

Observe that P_1, P_2, P_3, and P_4 rank country d first, and P_5 and P_6 rank country a first. Thus, country d lifts up $\{1,2,3,4\}$ and country a lifts up 5 and 6 to top priority class $\mathcal{R}_1$. This gives us the combined priority profile π^{E-top} above.

After applying the DA to π^{E-top} and P, we obtain the matching below:

$$\mu_{E-top}^{DA} = \begin{pmatrix} a & b & c & d \\ 1 & 2 & \{5,6\} & \{3,4\} \end{pmatrix}$$

Observe that the blocking pair set of μ_{E-top}^{DA} with respect to $\succ$ consists of only $(5,d)$ and $(6,d)$. Since d is not the top choice of either of the refugees 5 and 6, top stability of μ_{E-top}^{DA} is satisfied at the given problem P. Moreover, the reason for refugees 5 and 6 blocking μ_{E-top}^{DA} with country d with respect to $\succ$ is the loss of country d to refugees 3 and 4 due to the priority reversal between $\{5,6\}$ and $\{3,4\}$. This priority reversal stems from P_3 and P_4 ranking d as their top choice, and this leads to $\{3,4\}$ getting lifted up to the top priority class of country d and gaining rank over refugees 5 and 6. Therefore, top PC fairness of μ_{E-top}^{DA} is satisfied.

A mechanism f is **strategyproof** if for all $r \in \mathcal{R}$, for all $P \in \mathcal{P}$ and all P'_r, $f_r(P)\ R_r\ f_r(P'_r, P_{-r})$. A coalition $L \subseteq \mathcal{R}$ can **manipulate** matching $f(P)$ at P if there exists P'_L such that for all $r \in L$, $f_r(P'_L, P_{-L})\ P_r\ f_r(P)$.[3]

Remark. *The Top Prioritization Mechanism is not strategyproof for refugee families.*

Proof. Let P be the original refugee preference profile and P' as the misreported refugee preference profile. Let $P_4 = aP_4dP_4cP_4b$ such that the only difference between these two profiles is that refugee 4 misreports her top choice country as d instead of the truthful a. For simplicity suppose $q_c = 1$ for each $c \in \mathcal{C}$. Note that refugee $f^{DA}_{E-top}(P)(4) = c$, where $f^{DA}_{E-top}(P)(4)$ is the outcome assigned to refugee 4 by the Top Prioritization Mechanism f^{DA}_{E-top} at P. Suppose the given priority classes $\mathcal{R}_1 = \{1, 2\}$ and $\mathcal{R}_2 = \{3, 4\}$.

P_1	P_2	P_3	P'_4
d	d	d	d
a	b	c	a
b	c	a	c
c	a	b	b

Refugee Preferences P'

$\succ_a$	$\succ_b$	$\succ_c$	$\succ_d$
1	2	3	4
4	1	4	3
2	4	2	1
3	3	1	2

Country Preferences $\succ$

π^E_a	π^E_b	π^E_c	π^E_d
1	2	2	1
2	1	1	2
4	4	3	4
3	3	4	3

Enforced Profile π^E

$\pi^{E-top'}_a$	$\pi^{E-top'}_b$	$\pi^{E-top'}_c$	$\pi^{E-top'}_d$
1	2	2	4
2	1	1	3
4	4	3	1
3	3	4	2

Combined Priority $\pi^{E-top'}$

After applying the DA to $\pi^{E-top'}$ and P', we have the matching below:

$$\mu^{DA'}_{E-top} = \begin{pmatrix} a & b & c & d \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

Therefore, refugee 4 manipulates the Top Prioritization Mechanism f^{DA}_{E-top} at P since there exists P'_4 such that $f^{DA}_{E-top}(P'_4, P_{-4})(4)\ P_4\ f^{DA}_{E-top}(P)(4)$. Hence, refugee 4 manipulates

[3] Alternative notations to $f_r(P'_L, P_{-L})$ and $f_r(P)$ are $f(P'_L, P_{-L})(r)$ and $f(P)(r)$, respectively.

by getting her untruthful top choice. □

2.7 Credible Stability

Drawing attention to the refugees who are mandated to be in lower priority classes, I recognize the need and importance of investigating the ways of giving these refugees an additional chance of increased rank. As I have mentioned, I pursue this by looking at two different forms of priority profiles. My first form is obtained by the Top Prioritization Mechanism, which gives these refugees a better chance at their top ranked country. In this section I study my second form, under which I give these refugees a better chance at their Deferred Acceptance (DA) matched country. For a relevant mechanism, which was studied independently and in a different context, see Biró and Gudmundsson (2020)[13].

In contrast to the Top Prioritization Mechanism, I find that the best matching these refugees can get in a stable matching is their DA match, which shows the importance of prioritizing these refugees at their DA matched countries. In addition, prioritizing these refugees at their DA matched countries is more compelling for countries because *each* country has to modify their priorities with respect to their quotas *less* compared to the Top Prioritization Mechanism. This is because under the Top Prioritization Mechanism it is possible that some countries do not have to modify the priority orders as much as the popular countries. For example, unpopular countries for settlement, such as Poland, are less likely to be top ranked by refugees. These countries will have fewer refugees to move up to their top priority class. Whereas, popular countries, such as Germany, may have to modify their priority rankings much more. If all refugees rank Germany as their top choice, then all refugees will be moved up to top priority class of Germany. Within class, Germany will rank according to its own preferences based on its points system. If a country is popular, then that country needs to make a lot of changes to its priority ordering, which then deviates more from the mandated priority classes.

I start by providing the definition of a *credible blocking pair* to build the new weak stability axiom. Here, I would like to remind the readers again that the matching result of the DA ran with $\succ$ and P is denoted by μ^{DA}.

Definition 2.7.1 (Credible Blocking Pair). A pair blocking a matching μ is credible if it is matched under μ^{DA}.

A pair blocking a matching μ is *non-credible* if it is *not* matched under μ^{DA}.

Definition 2.7.2 (Credible Stability Axiom). A matching is credibly stable if it cannot be blocked with a credible blocking pair at a given P. A mechanism is credibly stable if, for each preference profile P, whenever the matching assigned to P is blocked by (r, c), (r, c) is not a credible blocking pair at P.

Definition 2.7.3 (Credible PC Fairness Axiom). A matching μ satisfies credible PC fairness at a particular profile P if, for every refugee-country pair (r, c) such that r prefers c to his assignment at μ and is preferred by c to a refugee $\hat{r}$ assigned to c at μ,

- either $\hat{r}$ is in a higher enforced priority class than r,
- or $\hat{r} \in \mu^{DA}(c)$ at P and r is not in priority class $\mathcal{R}_1$ of country c.

A mechanism f satisfies credible PC fairness if, for every profile P, it assigns a matching μ that is credibly PC fair. Note that whenever there is justified envy with respect to $\hat{r}$, either $\hat{r}$ is in a higher enforced class than r with justified envy, or $\hat{r} \in \mu^{DA}(c)$ at P and r is not in top priority class of country c.

In order to accommodate more refugees in lower priority classes by raising them to the first priority class at their DA matched country, I further weaken the PC fairness axiom. Therefore, PC fairness axiom implies credible PC fairness. Since I make exceptions for these refugees at their DA matched country and to account for their exceptions of gained rank, I declare blocking pairs that involve such refugees and their DA matched country to be salient and not allowed. In addition, credible stability implies credible PC fairness. We can see this in Example 4.

Definition 2.7.4 (**The DA-Match Prioritization Mechanism**).

1. For every refugee r, modify π_c^E by lifting refugee r up to priority class $\mathcal{R}_1$ of c, where $r \in \mu^{DA}(c)$.

2. Position r within priority class $\mathcal{R}_1$ according to $\succ_c$ (i.e., leave all other priority rankings the same, except for r's). This leads to the new combined profile π^{E-DA}.

3. Apply the DA to π^{E-DA}. This leads to the matching result $f_{E-DA}^{DA}(P)$ of the DA-Match Prioritization Mechanism f_{E-DA}^{DA} at problem P.

Similar to my other notations for the matching outcomes, I use μ_{E-DA}^{DA} for the matching result of the DA-Match Prioritization Mechanism f_{E-DA}^{DA} at any given problem P.

Unlike the Top Prioritization Mechanism, each country's DA matched refugees get lifted up to priority class $\mathcal{R}_1$. Under the Top Prioritization Mechanism, only the top-ranked countries lift refugees up to their priority class $\mathcal{R}_1$. However, now I lift DA matched refugees up to their DA matched country's priority class $\mathcal{R}_1$. The key distinction is that it is the DA-match that is used to move refugees up, which is an assignment, unlike the top choices. Thus, each country gets to lift up her DA matched refugees.

Lemma 2. *If $r \in \mu^{DA}(c)$, then $S_c^{\pi^{E-DA}}(r) \subseteq S_c^{\succ}(r)$.*

Proof. Let r and c be such that $r \in \mu^{DA}(c)$. Then, r is lifted to $\mathcal{R}_1$ in π_c^{E-DA}. Fix $\hat{r} \in S_c^{\pi^{E-DA}}(r)$. Then, we have $\hat{r} \in \mathcal{R}_1$. Moreover, given the construction of π^{E-DA}, this means that $\hat{r} \succ_c r$. Thus, $\hat{r} \in S_c^{\succ}(r)$. □

Lemma 3. *μ^{DA} is stable with respect to the priority profile π^{E-DA}.*

Proof. For contradiction, suppose there is a pair (r, c) blocking μ^{DA} with respect to π^{E-DA}. Let r be one of the $|\mu^{DA}(\bar{c})|$ refugees matched to $\bar{c}$. That is, let $r \in \mu^{DA}(\bar{c})$. Then, $c \; P_r \; \bar{c}$. Let $\hat{r}$ be one of the $|\mu^{DA}(c)|$ refugees matched to c. That is, let $\hat{r} \in \mu^{DA}(c)$. Then, $r \; \pi_c^{E-DA} \; \hat{r}$. Thus, $r \in S_c^{\pi^{E-DA}}(\hat{r})$, and Lemma 2 implies that $r \in S_c^{\succ}(\hat{r})$. Then (r, c) is also blocking μ^{DA} with respect to $\succ$, which is a contradiction, since μ^{DA} is stable at $\succ$. □

Lemma 3 can occur under two cases: μ^{DA} is stable with respect π^{E-DA} either by simply being equal to μ^{DA}_{E-DA} or when these two matches differ. Through the DA-Match Prioritization Mechanism, we have either $\mu^{DA}_{E-DA} = \mu^{DA}$ implying μ^{DA} is still stable with respect to π^{E-DA} and even μ^{DA}_{E-DA} is stable with respect to $\succ$. Or, $\mu^{DA}_{E-DA} \neq \mu^{DA}$, where μ^{DA} is still stable with respect to π^{E-DA}, and μ^{DA}_{E-DA} Pareto dominates μ^{DA} for refugees, which I prove in the next theorem.

Remark. *When μ^{DA}_{E-DA} coincides with μ^{DA}, then both μ^{DA} and μ^{DA}_{E-DA} are stable with respect to both $\succ$ and π^{E-DA}.*

Definition 2.7.5 (Weak Pareto Domination). A mechanism f weakly Pareto dominates another mechanism g if for every P, the matching assigned to f weakly Pareto dominates the matching assigned to g.

Theorem 3. *The DA-Match Prioritization Mechanism weakly Pareto dominates the DA at P for the refugees.*

Whenever the matching results do not coincide at a particular profile P, then the DA-Match Prioritization Mechanism leads to a matching μ^{DA}_{E-DA} that Pareto dominates μ^{DA} for refugees.

Proof. From Gale and Shapley's (1962) optimality result, we know the refugee-proposing

DA-Match Prioritization leads to the refugee-optimal stable matching with respect to π^{E-DA}. Since μ^{DA}_{E-DA} is the refugee-optimal stable matching at π^{E-DA} and matching μ^{DA} is also stable with respect to π^{E-DA} by Lemma 3, if $\mu^{DA}_{E-DA} \neq \mu^{DA}$, then μ^{DA}_{E-DA} Pareto dominates μ^{DA} with respect to P. □

Furthermore, by Knuth's (1976)[38] polarity result, we know that the refugees' preferences P and the combined profile for the countries π^{E-DA} are opposed to each other on the set of stable matchings. Thus, if μ and μ' are stable matchings, then all refugees like μ at least as well as μ' if and only if all countries like μ' at least as well as μ. Intuitively, the best stable matching for one side is the worst stable matching for the other side. Thus, the refugee-optimal stable matching is the worst stable matching for countries (country-pessimal), and the country-optimal stable matching is the worst stable matching for refugees (refugee-pessimal). Therefore, using Knuth's (1976)[38] polarity result and by Lemma 3, I observe that when μ^{DA}_{E-DA} is different than μ^{DA}, we have the matching μ^{DA}_{E-DA} that is the country-pessimal stable matching with respect to π^{E-DA}. Thus, from the point of view of countries, μ^{DA}, which is stable at π^{E-DA}, Pareto dominates μ^{DA}_{E-DA} at π^{E-DA}.

Moreover, the DA-Match Prioritization Mechanism leads to a matching μ^{DA}_{E-DA} that satisfies credible stability. Recall that a credible blocking pair is a pair that is matched under μ^{DA}.

Theorem 4. *The DA-Match Prioritization Mechanism is*

1. *credibly stable and*
2. *credibly PC fair.*

Proof.

1. *Credible stability:* Fix a given preference profile P. Suppose $\mu^{DA}_{E-DA} \neq \mu^{DA}$. For

contradiction, suppose the matching result μ^{DA}_{E-DA} is not credibly stable at the given profile P. Then there is a pair (r, c) blocking μ^{DA}_{E-DA} with respect to $\succ$ that is matched under μ^{DA}. Let r be one of the $|\mu^{DA}_{E-DA}(\bar{c})|$ refugees matched to $\bar{c}$ under the DA-Match Prioritization Mechanism. That is, $r \in \mu^{DA}_{E-DA}(\bar{c})$. We know $\mu^{DA}_{E-DA} \neq \mu^{DA}$ and $c \neq \bar{c}$. Thus, by Theorem 3, $\bar{c}\ P_r\ c$. This contradicts the assumption that (r, c) is a blocking pair of the matching μ^{DA}_{E-DA} with respect to $\succ$, given the profile P.

2. Credible PC fairness: Fix a given preference profile P. For contradiction, suppose μ^{DA}_{E-DA} does not satisfy credible PC fairness at P. Then at the given P, there exists a pair (r, c) blocking μ^{DA}_{E-DA} with respect to $\succ$ that is *not* allowed under credible PC fairness and there exists $\hat{r} \in \mathcal{R}$ such that $\hat{r} \in \mu^{DA}_{E-DA}(c)$. The violation of the credible PC fairness of the matching result μ^{DA}_{E-DA} at given P implies four non-trivial cases. We show contradiction under each case when there is a salient blocking pair that is not allowed under the weak fairness axiom of credible PC fairness.

Observe first that when refugee $\hat{r}$ is not in a higher enforced priority class than r who is assumed to be blocking with c at $\succ$, then they are in the same priority class. Albeit this is trivial, it is still a possibility since the matching result is assumed to be not top PC fair at the given profile. However, this violates country preferences within same priority class, which are assumed to be preserved. Hence, this is a contradiction.

Case 1: Suppose $\hat{r} \notin \mu^{DA}(c)$ at given P and r is in top priority class $\mathcal{R}_1$ of country c. Then, either $\hat{r}$ is in a lower priority class than r, or $\hat{r}$ is already in top priority class $\mathcal{R}_1$ and hence they are both in $\mathcal{R}_1$ of country c. If $r \in \mathcal{R}_1$ of c and $\hat{r}$ is in a lower priority class, then this contradicts $\hat{r}\ \pi^{E-DA}_c\ r$ since $\hat{r} \in \mu^{DA}_{E-DA}(c)$. If both $r, \hat{r} \in \mathcal{R}_1$, then $r \succ_c \hat{r}$ and thus should be preserved in $\mathcal{R}_1$. Then $r\ \pi^{E-DA}_c\ \hat{r}$. However, this contradicts $\hat{r}\ \pi^{E-DA}_c$ r given $\hat{r} \in \mu^{DA}_{E-DA}(c)$.

Case 2: Suppose $\hat{r} \in \mu^{DA}(c)$ at P and r is in priority class $\mathcal{R}_1$ of country c. Then, $r, \hat{r} \in \mathcal{R}_1$, and we know $r \succ_c \hat{r}$, but this contradicts $\hat{r}\ \pi^{E-DA}_c\ r$.

Case 3: Suppose $\hat{r} \notin \mu^{DA}(c)$ at P and r is not in priority class $\mathcal{R}_1$ of country c. If $\hat{r}$

is already in mandated class $\mathcal{R}_1$, then $\hat{r}\ \pi_c^{E-DA}\ r$ and thus there is nothing to prove as μ_{E-DA}^{DA} is credibly PC-fair. However, consider otherwise. Then, neither of the two refugee families are in $\mathcal{R}_1$.

Hence, since we assumed μ_{E-DA}^{DA} is not credible PC fair, either r and $\hat{r}$ are in the same priority class or $\hat{r}$ is in a lower priority class than r. If they are both in the same priority class, then (r, c) blocking μ_{E-DA}^{DA} at $\succ$ implies that $r \succ_c \hat{r}$, which is preserved in the same priority class, hence $r\ \pi_c^{E-DA}\ \hat{r}$. This is a contradiction to $\hat{r}\ \pi_c^{E-DA}\ r$ given $\hat{r} \in \mu_{E-DA}^{DA}(c)$. If $\hat{r}$ is in a lower priority class than r, then r having envy for $\hat{r}$ who is in a lower priority class is a contradiction to the PC no-envy property of the match. This is because if r is in a higher priority class than $\hat{r}$ then r does not have envy for $\hat{r}$ at P.

Case 4: $\hat{r} \in \mu^{DA}(c)$ at P and r is not in priority class $\mathcal{R}_1$ of country c. Then, (r, c) is a salient blocking pair that is allowed under credible PC fairness. Hence, this is a contradiction.

Therefore, the DA-Match Prioritization Mechanism satisfies credible stability and credible PC fairness. □

Example 4. Consider the given refugees $\mathcal{R} = \{1, 2, 3, 4, 5, 6, 7, 8\}$, countries $\mathcal{C} = \{a, b, c, d\}$, and mandated priority classes $\mathcal{R}_1 = \{1, 2, 5, 6\}$ and $\mathcal{R}_2 = \{3, 4, 7, 8\}$. This is an example to demonstrate the case where the DA-Match Prioritization Mechanism leads to a matching μ_{E-DA}^{DA} that does not coincide with μ^{DA}. Suppose $q = (2, 2, 2, 2)$. Allocations that are double-underlined are for μ_{E-DA}^{DA}, which is obtained from applying the DA to P and π^{E-DA}. Allocations that are underlined are for μ^{DA}. Allocations that are marked in **bold** are when μ_{E-DA}^{DA} coincides with μ^{DA}. Note that . indicates any arbitrary ranking.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
d	d	d	a	d	d	d	a
a	**b**	**c**	d	**a**	**b**	**c**	**d**
b	c	a	c	b	c	a	c
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

Refugee Preferences P

$\succ_a$	$\succ_b$	$\succ_c$	$\succ_d$
1	**2**	**3**	4
5	**6**	**7**	**8**
4	1	2	3
8	5	6	7
2	3	1	1
6	7	5	2
3	4	4	5
7	8	8	6

Country Preferences $\succ$

π^E_a	π^E_b	π^E_c	π^E_d
1	2	2	1
5	6	6	2
2	1	1	5
6	5	5	6
4	3	3	4
8	7	7	8
3	4	4	3
7	8	8	7

Enforced Profile π^E

π^{E-DA}_a	π^{E-DA}_b	π^{E-DA}_c	π^{E-DA}_d
1	2	3	4
5	6	7	8
2	1	2	1
6	5	6	2
4	3	1	5
8	7	5	6
3	4	4	3
7	8	8	7

Combined Profile π^{E-DA}

After applying the DA to π^{E-DA} and P, we get:

$$\mu^{DA}_{E-DA} = \begin{pmatrix} a & b & c & d \\ \{4,5\} & \{2,6\} & \{3,7\} & \{1,8\} \end{pmatrix}$$

$$\mu^{DA} = \begin{pmatrix} a & b & c & d \\ \{1,5\} & \{2,6\} & \{3,7\} & \{4,8\} \end{pmatrix}$$

Applying the DA to P and π^{E-DA} gives the refugee-optimal stable matching μ^{DA}_{E-DA} at π^{E-DA}. The two matching results above are different from each other. Looking at P, observe that μ^{DA}_{E-DA} Pareto dominates μ^{DA} for refugees. At P, refugees 1 and 4 are better off under the matching μ^{DA}_{E-DA}. Note that $(3,d)$ and $(7,d)$ are the only pairs blocking μ^{DA}_{E-DA} with respect to $\succ$. Since $(3,d)$ and $(7,d)$ are not matched in μ^{DA}, they are not credible. Then, these blocking pairs are allowed to exist under the weak stability of credible stability. Thus, the matching μ^{DA}_{E-DA} is credibly stable. In addition, the reason for refugees 3 and 7 blocking μ^{DA}_{E-DA} with respect to $\succ$ is the loss of country d to refugee 1 due to the priority class reversals between 1 and 3, and 1 and 7. These priority reversals stem from the discrepancy between π^E and $\succ$. That is, 1 is in a higher enforced priority class than 3 and 7 in π^E, which leads to the following reversal: $3 \succ_d 1$ versus $1\ \pi_d^E\ 3$ and $7 \succ_d 1$ versus $1\ \pi_d^E\ 7$. Therefore, μ_{E-DA}^{DA} *is credibly PC fair because it is PC fair.*

In this example, we can also see that *credible stability implies credible PC fairness.* Notice that *none* of the blocking pairs $(3, d)$ and $(7, d)$ are matched under the DA. In the combined profile π^{E-DA} refugees are already moved up to top priority class of their DA-matched countries. Therefore, the priority reversal between refugees $\{3, 7\}$ and 1 is due to the mandated priority classes, i.e., difference between the mandated profile of countries π^E and country preferences $\succ$. Refugees 3 and 7 have justified envy for 1 since 1 is enforced above these other two refugees in π^E. These blocking pairs are then allowed under PC fairness, and PC fairness implies credible PC fairness.

Remark. *The DA-Match Prioritization Mechanism is not strategyproof for refugee families.*

Proof. The matching assigned by the DA-Match Prioritization Mechanism Pareto-dominates the matching made by the DA at each preference profile. Thus, by Kesten (2006)[36] and Ergin (2002)[25], or Kesten and Kurino (2019)[37], the DA-Match Prioritization Mechanism cannot be strategyproof.[4]

□

2.8 Conclusion

In order to bring the chaotic stream of refugees under control, one potential solution

[4] Kesten (2006)[36] or Kesten and Kurino (2019)[37] are more relevant here, since there are no outside options for refugees in the present setting. Note that this holds also by Abdulkadiroğlu, Pathak, Roth (2009)[1] (also by Alva and Manjunath (2019)[6]), however they require refugees have an outside option.

for countries is to participate in a centralized matching mechanism. However, not all countries can be persuaded to participate in a centralized computerized matching mechanism. There is an urgent need to overcome the political impasse preventing countries from participating in accountability-sharing in the international refugee crisis and finding solutions to the European refugee crisis. In this study, by defining and investigating a country acceptance problem, I design two new matching mechanisms: the Top Prioritization Mechanism and the DA-Match Prioritization Mechanism. I base these new mechanisms on the priority class hierarchy that I propose using the UNHCR humanitarian principles and guidelines. Both of these mechanisms are to be implemented as a centralized and computerized refugee matching system to match refugee families to countries. I contribute to the literature by having the country preferences alongside the UNHCR-mandated hierarchical priority classes of refugees. That is, with the two new mechanisms in this study, I contribute to the literature by proposing methodologies to reconcile country preferences and the UNHCR-mandated hierarchical priority classes while maintaining the good stability and fairness properties. I recognize the importance of explicitly analysing countries' preferences simultaneously with the prioritization of refugees enforced on countries. I conduct an axiomatic and fair allocation analysis with a focus on the stability and fairness of the two new matching mechanisms.

Having two kinds of ranking profiles for countries, the mandated priority profile versus the preference profile, allows me to study the real world predicaments associated with the refugee relocation problem. I capture how the difference between the two ranking profiles creates blocking pairs of countries and refugees due to the mandated hierarchical priority classes. I weaken the stability axiom, since a mechanism that is stable with respect to a mandated profile may no longer be stable with respect to countries' preferences leading to potential blocking pairs of countries and refugees. With the Top Prioritization Mechanism, I provide an additional chance for refugees who are mandated in lower priority classes and therefore face the risk of being stuck in lower priority classes, by prioritizing these refugees at their top-choice country. I show that the Top Prioritization Mechanism is top stable and top PC fair. With the DA-Match Prioritization Mechanism, I provide an additional chance for refugees who are mandated in lower priority classes, by lifting these refugees up to the first priority class of their DA matched country. I show that the DA-Match Prioritization Mechanism is credibly stable and credibly PC fair. I also find that for refugees, the DA-Match Prioritization Mechanism weakly Pareto dominates the DA. Furthermore, I recognize the importance of persuading countries to participate in a centralized refugee matching mechanism. I believe that a priority class-based approach with no imposed category-specific (e.g., PC-specific) set-aside reserve quotas may be more attractive in terms of increasing countries' willingness and incentive to participate in solving the refugee matching problem. Beyond the refugee relocation context, the results of this study have other policy applications, such as centralized college admissions, the design of public school choice systems, and immigration.

Chapter 3

Targeted Priority Reserve Policies

3.1 Introduction

Affirmative action policy limits the number of admitted majority students to give minority students a better chance to be represented at universities. The literature on matching with an affirmative action policy demonstrates that simply restricting the number of majority students who can be admitted by a university, as first introduced by Abdulkadiroğlu and Sönmez (2003)[2] and later examined by Kojima (2012)[39], can easily cause inefficiency, since some university seats may remain unfilled if there are not enough minority applicants to fill all the reserved minority seats, given that majority students who desire empty seats are not allowed to get them. Hafalir et al. (2013)[30] propose an affirmative action policy with minority reserves, the DA-MiR (Deferred Acceptance with Minority Reserves) mechanism, which is a Deferred Acceptance mechanism with a quota set aside at each university to prioritize minority students.

Hafalir et al. (2013)[30] demonstrate that the DA-MiR mechanism does not strictly limit the number of majority students who can be admitted and show that their mechanism leads to a Pareto-improvement over the previously studied majority quota policies, remedying the inefficiency. Further theoretical studies of affirmative action include Ehlers et al. (2014)[23] and Doğan (2016)[16]. Many of the studied mechanisms are Deferred Acceptance mechanisms with choice functions, which are explored in general by Ehlers and Klaus (2016[24]). Related studies of choice functions that are useful for understanding matching with a priority treatment are Echenique and Yenmez (2015)[22], Chambers and Yenmez (2017[14]) and Doğan et al. (2019)[17].[1]

The DA-MiR mechanism of Hafalir et al. (2013)[30] gives higher priority to minority students up to the reserved number of seats, but this quota can only be exceeded by the minority students if they no longer enjoy a preferential treatment at the university and can all qualify without the reserve policy, effectively putting a cap on minority student admission numbers. This matching mechanism serves the goal of minority representation well, given its attractive stability and incentive properties. However, if the primary goal is not representation but to help prioritize a group of typically disadvantaged agents, such as children with disabilities or refugee families in war zones, who may not be matched otherwise to highly-ranked entities or may not be matched at all, then the DA-MiR mechanism is less appealing.

To serve the goal of effective priority treatment (what we refer to as "protection" in short), we propose an alternative matching mechanism with *targeted* priority reserves, the DA-TPR mechanism (Deferred Acceptance with Targeted Priority Reserves), along with a corresponding new stability concept, which capture the objective of targeting agents in the priority group (the "priority agents") who are in need, while recognizing that not all priority agents may need help. This allows for a higher number of priority agents to be matched to an entity than the quota set aside for them, which is desirable when the objective is not representation or diversity, as in affirmative action, and "overrepresentation" is not an issue. The DA-TPR mechanism fills the unreserved positions of the entities (e.g., countries or schools) first, based on a general competition of all agents regardless of their priority status. This gives an opportunity to priority agents to get accepted by entities on their own right, without using up reserved positions. Only after

[1]For additional literature, see Fleiner (2003)[26], Hatfield and Milgrom (2005)[32], Hatfield and Kominers (2016)[31], and Aygün and Turhan (2020)[11].

the unreserved positions are filled do entities target priority group agents who are indeed in need of preferential treatment and fill the reserved positions. This is the exact reverse of the DA-MiR mechanism which fills the reserved positions first, which may lead to a situation where already qualified priority agents take up most of the reserved positions, making the priority treatment policy unhelpful.

For example, when refugees from war zones (such as Syria) are to be matched to countries which set aside a quota of reserved slots for such refugees to help them, if many Syrian refugee families qualify for a particular country, say Germany, without relying on their priority status, then they may take up most (or even all) the set-aside slots, rendering the priority treatment policy ineffective. If the qualified Syrian refugee families were allowed to be matched to Germany without using up the set-aside slots, then other Syrian refugee families that would not qualify otherwise would be able to use the reserved asylum slots. Thus, the number of refugees from war zones would exceed the quota set by a country by the number of qualified war-zone refugees who did not need take up reserved positions, while all reserved positions would go to (or "target") those war-zone refugees who need prioritizing to be accepted by this country. A targeted reserve policy may be applicable to any situation where a disadvantaged group is granted preferential treatment but some members of this group do not require this treatment since they qualify without their priority status, provided that the main objective is not representation or diversity, such as affirmative action for visible or ethnic minorities, but rather to help those priority agents who are in need of preferential treatment.

Our proposed targeted priority reserve policy, the DA-TPR mechanism, may promote diversity, but does not aim for representation, and when applying this mechanism the number of set-aside positions at each entity should reflect how many priority agents are to be actually prioritized by the entity, not how many priority agents are to be matched to the entity. In other words, when using the DA-TPR mechanism appropriately, the reserve quota should be set according to the number of priority agents who are to be matched to the entity *who would not be matched to it without the priority reserve policy.* Let us emphasize that our proposal is not equivalent to simply increasing the reserve quotas, rather, it advocates using a mechanism that corresponds to the objective of the priority policy. The DA-MiR mechanism is to be used when we want to make sure that a certain number of priority agents is matched to an entity and set the reserve quota accordingly, serving the purpose of representation. By contrast, the DA-TPR mechanism is to be used when we want to make sure that a certain number of priority agents is effectively helped by the priority policy, regardless of whether more priority agents are matched to an entity than its reserve quota, and we set the number of reserved positions accordingly. Indeed, if the DA-MiR or a similar representation policy is used inappropriately in a situation where the main goal is protection, were the designer to switch to the correct DA-TPR policy the reserve quotas may often be reduced, and this policy would still result in a more effective priority treatment than the inappropriate DA-MiR mechanism if there are significant numbers of highly-qualified priority agents. In a broader view, the DA-MiR and DA-TPR mechanisms can be seen as the two extreme members of a class of matching mechanisms that allow for all sequences of reserved and unreserved positions, and we also study this large family of mechanisms, which we call DA-SPR (Deferred Acceptance with Sequential Priority Reserves) mechanisms.

The idea of changing the sequencing of the reserved and unreserved positions when filling positions to achieve different results is not new. The effects of different sequencing (the "precedence order") have been studied previously in the school choice context by Dur et al. (2018[19]), (2020)[21]) in different models from ours, examining specific school choice reserve policies in Boston and Chicago respectively. In our model we have two exogenously given sets of agents, priority agents and non-priority agents, while in Dur et al. (2018)[19] each school has a reserve-eligible population, and in Dur et al. (2020)[21] each student belongs to a specific socio-economic tier and has a composite score which serves as a basis for prioritization. Dur et al. (2020)[21] presents the first optimality results for precedence policies in the framework of matching with slot-specific priorities. The model they study is based on a continuum model version of the matching with slot-specific priorities model introduced by Kominers and Sönmez (2016)[40]. The two applied school choice papers [19] and [21] study school-level selection only, and consequently do not use stability concepts or study the properties of entire matching mechanisms, which is the main subject of our paper. Apart from their empirical analyses, the theoretical results focus on studying small changes in the sequencing of a school's choice function by switching adjacent positions in the school's precedence order, which allows for unambiguous comparisons. Dur et al. (2018)[19] points out that switching the precedence order position of a reserve slot of a school with the position of an open slot following it, while fixing the other precedence order positions, weakly increases the number of reserve eligible students chosen by the school from a given pool of students. We examine more comprehensive criteria for selection sequences and establish the equivalence of three related concepts of comparability and consistency for the selection of priority agents (Theorem 12) for DA-SPR entity-level selection rules. These results serve as a foundation for comparing DA-SPR matching mechanisms (Theorem 14), since we study primarily entire matching mechanisms, similarly to Hafalir et al. (2013)[30], along with new stability concepts, and analyze their stability, efficiency, and incentive properties.

Another major difference from the previous two school choice studies that consider precedence orders is that both [19] and [21] view the specific sequencing as merely a means of affecting the number of accepted prioritized students. In this paper we argue, on the other hand, that different precedence orders correspond to different policy goals, namely representation versus protection, and that the latter objective is served by "targeting" priority agents who need priority treatment. In our approach the DA-MiR mechanism can be seen as the least targeted priority policy, since it selects agents for all reserved positions up front, before filling the unreserved positions, thereby giving reserved positions to agents who would qualify without taking up a reserved position. At the other extreme, the DA-TPR mechanism that we propose emerges as the most targeted priority policy, since it selects first from the general agent pool to fill the unreserved positions, allowing for qualified priority agents to take unreserved positions, and then it fills the reserved positions from the remaining lower-ranked priority agents.

Our motivation is also different from that of the previous papers which study specific school choice designs. Our main inspiration to identify and study the difference between the goals of representation and protection, and the consequences of this difference for designing a priority policy, comes from the refugee matching problem (see Jones and Teytelboym (2016[34], 2017[35], Andersson and Ehlers (2020)[7], and Van Basshuysen (2017)[59]). The shift from representation to protection is particularly important in the refugee protection context, since representation is not an issue for refugee settlement, but helping those refugee families that are in need of extra protection, for example refugees from war zones, is very relevant for designing a centralized refugee matching mechanism to solve the host country choice problem of refugee families (Sayedahmed (2017)[51], (2018)[52]). Sayedahmed (2018)[52] studies the design of a centralized refugee matching mechanism with priority classes, where refugee families in danger zones and/or with a long waiting period would be prioritized over others. In this context it is irrelevant to track which priority class the refugee families come from, and overrepresentation or underrepresentation of refugees from a particular priority class is not a matter of concern for countries, while the protection of refugee families from danger zones is. The protection goal is particularly appropriate during an emergency situation such as the Syrian refugee crisis, since it is not just a matter of matching Syrian refugee families to more desirable countries, but also whether we can match them to any country at all, given the shortage of resettlement opportunities, which may literally be a matter of life and death.

In this paper, which is meant to be a first theoretical study of mechanisms with different precedence orders serving different policy goals in a simple setting, we show that the DA-TPR mechanism retains the positive features of the DA-MiR mechanism. We prove that it satisfies our new stability concept called *protection-stability,* and that the DA-TPR mechanism is the optimal protection-stable mechanism (Theorem 5). We also prove that the DA-TPR mechanism, just like the DA-MiR mechanism, is weakly group-strategyproof, along with the entire class of DA-SPR mechanisms (Theorem 10). We also establish characterization results for the two main mechanisms: the DA-TPR is the only strategyproof mechanism that is *protection-stable* (Theorem 6), while the DA-MiR is the only strategyproof mechanism that is *representation-stable* (Theorem 7). Protection-stability is our new stability concept which aims for effective preferential treatment, and representation-stability corresponds to the stability concept with minority reserves introduced by Hafalir et al. (2013)[30]. Regarding their respective policy impact, we demonstrate that if agents have the same preferences over entities and entities have the same ranking over agents then the DA-TPR outcome weakly Pareto-dominates the DA-MiR outcome for priority agents (Theorem 14). Although there may be correlation among agent preferences and entity rankings in many real-life matching applications, we focus on homogeneous profiles primarily because they provide a benchmark that allows us to compare the two matching mechanisms which are not comparable in general for arbitrary ("noisy") profiles (Theorem 13). The comparison that we obtain based on the benchmark homogeneous profiles demonstrates that the DA-TPR mechanism increases the welfare of the priority group when compared to the DA-MiR mechanism, without increasing the reserve quotas. It is also clear that at the entity-selection level the DA-TPR mechanism unambiguously favors priority agents compared to the DA-MiR mechanism (Theorem 12), keeping in mind that preferential treatment is only provided for the set-aside reserved positions, guaranteeing the rights of the non-priority group.

The broader family of DA-SPR mechanisms that we study includes arbitrary precedence orders specified for each entity, and thus each member of this family is associated with a different set of selection rules for the entities. We explore this class of mechanisms and study their comparability in terms of their policy impact both at the entity level and at the level of the entire mechanism (Theorems 12, 13, and 14). We obtain the following characterization as our main result: a matching mechanism satisfies a weak reserve-stability axiom (weaker than both protection-stability and representation-stability) and strategyproofness if and only if it is a DA-SPR rule (Theorem 11). In addition, we show that DA-SPR rules are constrained efficient (Theorem 9) and weakly group-strategyproof (Theorem 10). These results demonstrate the importance of this family of priority treatment policies. We also identify and study a special subset of the family of DA-SPR mechanisms, the Split DA-SPR mechanisms, which stand out due to their transparency. These mechanisms offer a range of explicit policies between the two extremes of the DA-MiR and DA-TPR policies and represent a compromise between the two by dividing the reserved positions into non-targeted and targeted reserved positions, with the non-targeted reserved positions at the beginning of the sequence (as in the DA-MiR) and the targeted positions at the end of the sequence (as in the DA-TPR). It is easy to understand the policy goal behind these mechanisms, as opposed to arbitrary mixed sequences which would be difficult to justify for practical purposes due to their ambiguous priority treatment policies. Therefore, we recommend adopting a policy from the class of Split DA-SPR mechanisms which allow for clarity and flexibility when selecting a priority reserve policy, enabling the policy designer not only to set the number of reserved positions but also to determine the extent to which the priority policy is to be targeted at agents who need preferential treatment at each entity.

Finally, in addition to the school choice studies on preference orders, there are also several recent working papers that study various applications of affirmative action with reserved positions, which were written simultaneously with our paper and further demonstrate the relevance of our theoretical inquiry. Sönmez and Yenmez (2019) [56] study India's Supreme Court mandated procedure. They highlight that while a candidate can never lose a position to a less meritorious candidate from her own group, she can lose a position to a less meritorious candidate from a higher-privilege group. By studying choice rules that eliminate justified envy, they propose an alternative choice rule that resolves this deficiency. For additional literature on minority reserves in different contexts, see Aygün and Turhan (2017)[9], Aygün and Turhan (2020)[10], and Aygün and Bó (2020)[8]. In addition, Pathak et al. (2020)[44] study the 2019 H-1B visa allocation rule adopted by the Trump administration which maximizes the rate of high-skill awards and minimizes the rate of low-skill awards. They study visa allocation rules as functions of a given set of applicants, and by fixing the applicant set they focus their inquiry on the properties of matchings. Another study on choice rules is Sönmez and Yenmez (2019)[57] which examines overlapping reserves. Sönmez and Yenmez (2019)[55] study the constitutional implementation of reservation policies in India, and is somewhat more related to our study. They propose a new mechanism based on the Deferred Acceptance mechanism with an adapted choice function for heterogeneous positions as an improvement over the currently used Supreme-Court-mandated procedure that implements reservations when the positions are homogeneous. This paper is the first to propose a mechanism to implement both vertical and horizontal reservations in India which involves heterogeneous positions, that is, where the applicants apply for distinct institutions simultaneously.

Lastly, Pathak et al. (2020)[46] study the fair allocation of vaccines, ventilators, and antiviral treatments, and provides a precedence order comparative static analysis as part of a sequential reserve matching. Some of their results, along with some of the findings in Dur et al. (2018)[19] and Dur et al. (2020)[21], are related to our Theorem 12, which is the only theorem that we have on choice rules (or entity selection rules, as we call them), as opposed to matching mechanisms. Pathak et al. (2020)[46] also study matching results and find that a matching result complies with eligibility requirements, is non-wasteful, and respects priorities if, and only if, it is DA-induced. All these recent papers are quite specific to the particular application they study, while our aim is to provide some fundamental insights in a simple and more abstract setting. On the whole, these studies demonstrate that there are many interesting real-life allocation problems that can be addressed using priority reserve positions, and support our view that preferential treatment policies are important and worthwhile to study using a more theoretical approach as well.

3.2 Model with Priority Reserves

A matching market with priority reserves is defined by the following components.

1. A finite set of **agents** N.

2. A **priority group** $N^+ \subset N$. We call $N^0 = N \setminus N^+$ the **non-priority group.** Thus, $N = N^+ \cup N^0$. We will refer to agents in the priority group as priority agents for short, and to agents in the non-priority group as non-priority agents.

3. A finite set of **entities** E.

4. Each entity $e \in E$ has a **capacity** $q_e > 0$, which is the number of available positions at entity e. Let $q \equiv (q_e)_{e \in E}$.

5. **Reserved positions.** Let $q_e = q_e^r + q_e^u$, where $q_e^r \geq 0$ denotes the number of reserved positions for the priority group N^+, and $q_e^u \geq 0$ is the number of unreserved positions at e. Let $q^r \equiv (q_e^r)_{e \in E}$ and $q^u \equiv (q_e^u)_{e \in E}$. In order to rule out uninteresting cases and simplify the exposition, we assume that at least one entity has at least one reserved position, and each entity has at least one unreserved position.

6. Each $i \in N$ has **strict preferences** P_i are over $E \cup \{i\}$. Agents may not find all entities acceptable, and ranking an entity below i indicates that the entity is not acceptable.

7. Each $e \in E$ has a **strict** (priority) **ranking** $\succ_e$ over N. Entities are not considered active agents with their exogeneously given priorities over agents, similarly to schools in the school choice model of Abdulkadiroğlu and Sönmez (2003). Thus, when considering efficiency and incentives, we only consider the agents' side. We will refer to the entity priorities from now on as **entity rankings** in order to avoid any confusion in the terminology between entity priorities and the fixed set of priority agents N^+.

Let $P \equiv (P_i)_{i \in N}$ denote a **preference profile** for the agents and let $\mathcal{P}$ be the set of preference profiles. Let $\succ \equiv (\succ_e)_{e \in E}$ denote an **entity ranking profile** and let Π be the set of entity ranking profiles. Then a market is defined by $\langle N, N^+, E, q, q^r, P, \succ \rangle$. Given fixed N, N^+, E, q, q^r and $\succ$, a market is determined by a preference profile P. Although we fix the entity ranking profile $\succ \in \Pi$, we indicate it explicitly throughout the paper and work with a **profile** $(P, \succ)$ consisting of a preference profile $P \in \mathcal{P}$ and the fixed entity ranking profile $\succ$.

We will also need the following notation concerning the preferences. Let R_i denote the weak preferences of agent $i \in N$ associated with P_i. Since preferences are assumed to be strict, eR_ie' means that either eP_ie' or $e = e'$. We will write $P_i \in (e)$ if agent i ranks entity e first, $P_i \in (e, e')$ if agent i ranks entity e first and entity e' second, and so on. Then, for example, $P_i \in (e, i)$ indicates that agent i finds only entity e acceptable. The preferences of a coalition $S \subseteq N$ in preference profile P are denoted by P_S. We denote the preference profile of all the agents except for i by P_{-i}, and the preference profile of all the agents except the agents in coalition L by P_{-L}.

In order to simplify the exposition, we assume that for all $e \in E$, $|N^+| > q_e$ and $|N^0| > q_e$. These are mild assumptions which are easily satisfied in most applications and allow us to ignore some unimportant special cases in which there are too few agents.

Matching. A solution to a problem is a matching μ which is a function from N to $E \cup N$ such that, for every agent $i \in N$, $\mu(i) \in E \cup \{i\}$, where $\mu(i) = e$ means that agent i is matched to entity e and $\mu(i) = i$ means that agent i is unmatched. We will refer to $\mu(i)$ as agent i's *assignment* in μ. With a slight abuse of notation we denote that set of agents matched to i in matching μ by μ_e, and hence require that $\mu_e \subseteq N$ such that $|\mu_e| \leq q_e$. Let $\mathcal{M}$ denote the set of matchings.

A matching μ *Pareto-dominates* matching ν at preference profile $P \in \mathcal{P}$ if for all agents $i \in N$, $\mu_i R_i \nu_i$, and there exists agent $j \in N$ such that $\mu_j P_j \nu_j$. A matching μ *weakly Pareto-dominates* matching ν at preference profile P if either μ Pareto-dominates ν or $\mu = \nu$.

Matching mechanism. A (matching) mechanism is a function $f : \mathcal{P} \to \mathcal{M}$ that assigns a matching to each preference profile $P \in \mathcal{P}$. Given the fixed entity ranking profile $\succ \in \Pi$, we can also write that a mechanism assigns a matching to profile $(P, \succ)$ for all $P \in \mathcal{P}$.[2]

A *mechanism f Pareto-dominates mechanism g for $N' \subseteq N$* if for all preference profiles $P \in \mathcal{P}$, $f_{N'}(P)$ weakly Pareto-dominates $g_{N'}(P)$, and there exists $\tilde{P} \in \mathcal{P}$ such that $f_{N'}(\tilde{P}) \neq g_{N'}(\tilde{P})$.

[2]Alternatively, we could also define a mechanism for all $\succ \in \Pi$, so that $f : (\mathcal{P} \times \Pi) \to \mathcal{M}$. All of our results hold for this broader definition of a mechanism, except for Theorem 11.

3.3 Stability Axioms

Our analysis is focused on stability with priority reserves, and we collect the definitions of our key stability concepts in this section. Let us first recall some standard concepts of stability. Given a profile $(P, \succ)$, a matching μ is *blocked by an agent* $i \in N$ if $iP_i\mu(i)$. A matching μ is *blocked by an agent-entity pair* $(i, e) \in N \times E$ if the agent prefers entity e to her assignment in μ and the agent's ranking is violated in the entity ranking of e. Formally, (i, e) blocks a matching μ if $e \ P_i \ \mu(i)$ and, given that $i \notin \mu_e$, either $i \succ_e j$ for some $j \in \mu_e$ or $|\mu_e| < q_e$. In the first case agent i is higher in entity e's ranking than agent j, and in the second case e has at least one position left unfilled and thus i could be matched to e in addition to the other agents who have already been matched to e in μ. A *matching* is *stable* if it is not blocked by an agent or an agent-entity pair. A *matching mechanism* is *stable* if it assigns a stable matching to each profile $(P, \succ)$.

Before introducing our main stability axiom, let $\bar{N}_e^+(\mu_e, \succ_e) \equiv \{j \in (\mu_e \cap N^+) : \forall l \in (\mu_e \cap N^0), \ l \succ_e j\}$. We can interpret $\bar{N}_e^+$ as the set of *protected priority agents* for entity e with ranking $\succ_e$ at a given matching μ (as long as at least one non-priority agent is matched to e), since these are the priority agents who are matched to e and are ranked lower by entity e than *all* non-priority agents matched to the entity. Note that whether a priority agent is protected is not given exogenously, since it depends on the matching (which in turn depends on the profile in a mechanism) and on the entity's ranking.

Protection-stability (stability with protective reserves):

A matching μ is protection-stable at profile $(P, \succ)$ if the following hold:

1. For all $i \in N$, $\mu(i)R_ii$.

2. If there are $i \in N$ and $e \in E$ such that $eP_i\mu(i)$, then $|\mu_e| = q_e$ and

   (2a) if $i \in N^+$ then, for all $j \in \mu_e$, $j \succ_e i$, and $|\bar{N}_e^+(\mu_e, \succ_e)| \geq q_e^r$;

   (2b) if $i \in N^0$ then, for all $j \in (\mu_e \cap N^0)$, $j \succ_e i$, and $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| \leq q_e^r$.

Given $\succ \in \Pi$, a matching mechanism is protection-stable if it assigns a protection-stable matching to profile $(P, \succ)$ for each preference profile $P \in \mathcal{P}$.

Condition 1 in the definition of protection-stability requires that the matching is not blocked by any agent, and condition 2 rules out blocking by an agent-entity pair with respect to protection reserves. In particular, to prevent such a blocking, whenever there exists agent i who prefers an entity e to her assignment in μ, condition 2 requires that the entity's capacity be filled. Moreover, if agent i is a priority agent, (2a) stipulates that all agents who are matched to e are ranked higher by e than i, and the q_e^r lowest-ranked agents by entity e who are matched to e must be priority agents, ensuring that all reserved positions are filled with protected priority agents, agents *who would not have been selected by entity e without the reserve policy* (i.e., in our terminology, all reserved positions are *targeted*). Furthermore, if agent i is a non-priority agent then (2b) requires that all non-priority agents who are matched to entity e are ranked higher by e than i, and the number of priority agents matched to e who are ranked below i by entity e does not exceed the number of reserved positions q_e^r, ensuring the rights of non-priority agents.

Now we state the definition of stability with minority reserves, proposed by Hafalir et al. (2013, [30]), written in a form that makes it easily comparable to the definition of protection-stability. We call this notion representation-stability.

Representation-stability (stability with representative reserves):

A matching μ is representation-stable at profile $(P, \succ)$ if the following hold:

1. For all $i \in N$, $\mu(i) R_i i$.
2. If there are $i \in N$ and $e \in E$ such that $e P_i \mu(i)$, then $|\mu_e| = q_e$ and
   (2a) if $i \in N^+$ then, for all $j \in \mu_e$, $j \succ_e i$, and $|\mu_e \cap N^+| \geq q_e^r$;
   (2b) if $i \in N^0$ then, for all $j \in (\mu_e \cap N^0)$, $j \succ_e i$, and if $|\{\mu_e \cap N^+\}| > q_e^r$ then for all $j \in \mu_e$, $j \succ_e i$.

Given $\succ \in \Pi$, a matching mechanism is representation-stable if it assigns a representation-stable matching to profile $(P, \succ)$ for each preference profile $P \in \mathcal{P}$.

Representation-stability is similar to protection-stability, but it differs from it in some key aspects concerning the treatment of reserved positions. It requires that whenever there exists a priority agent i who prefers an entity e to her assignment in μ, *all* reserved positions at e are filled with priority agents, but note that it is not required that the reserved positions are targeted, that is, some or all of the priority agents matched to e could have been selected by entity e without the reserve policy. Furthermore, if there exists a non-priority agent i who prefers entity e to her assignment in μ, the number of priority agents matched to e can only exceed the number of reserved positions q_e^r if i has a lower rank at e than all the agents matched to e, implying that a priority agent cannot take up a reserved position unless there are enough highly-ranked priority agents that none of them require a reserved position, in which case the top q_e agents are selected regardless of their priority status.

Protection-stability and representation-stability are logically independent of each other. The differences in the definitions of representation-stability and protection-stability are restricted to the second condition in both (2a) and (2b). Representation-stability does not imply protection-stability since (2a) is stronger for protection-stability, as protection-stability gives more rights to priority agents. Formally, since $N_e^+(\mu_e, \succ_e) \subseteq (\mu_e \cap N^+)$, (2a) in protection-stability implies (2a) in representation-stability. On the other hand, protection-stability does not imply representation-stability since (2b) is stronger for representation-stability, as representation-stability gives more rights to non-priority agents. Suppose that $|\{h \in (\mu_e \cap N^+) : i \succ_e h\} > q_e^r$ and (2b) in representation-stability holds. Then $|\mu_e \cap N^+| > q_e^r$, and it follows that for all $j \in \mu_e$, $j \succ_e i$. This is a contradiction, since $q_e^r \geq 0$ implies that $|\{h \in (\mu_e \cap N^+) : i \succ_e h\} > 0$, and thus there exists at least one $h \in (\mu_e \cap N^+)$ such that $i \succ_e h$.

Our last stability axiom is a weak stability axiom with priority reserve positions which is implied by both protection-stability and representation-stability. It is defined by using the weaker condition in each case in the definitions of the two stability axioms, namely, (2a) specifying the rights for minority students is taken from representation-stability, and (2b) specifying the rights for majority students is taken from protection-stability.

Weak reserve-stability:

A matching μ is weakly reserve-stable at profile $(P, \succ)$ if the following hold:

1. For all $i \in N$, $\mu(i) R_i i$,

2. If there are $i \in N$ and $e \in E$ such that $e P_i \mu(i)$, then $|\mu_e| = q_e$ and

(2a) if $i \in N^+$ then, for all $j \in \mu_e$, $j \succ_e i$ and $|\mu_e \cap N^+| \geq q_e^r$

(2b) if $i \in N^0$ then, for all $j \in (\mu_e \cap N^0)$, $j \succ_e i$ and $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| \leq q_e^r$.

Given $\succ \in \Pi$, a matching mechanism is weakly reserve-stable if it assigns a weakly reserve-stable matching to profile $(P, \succ)$ for each preference profile $P \in \mathcal{P}$.

We remark that if $N^+ = \varnothing$ (or if $q_e^r = 0$ for each entity e) then all three stability axioms reduce to standard stability, since then (2a) and the second part of (2b) are satisfied vacuously in each definition and the first part of (2b) simply rules out blocking by an agent-entity pair.

We will say that if an agent-entity pair (i, e) violates either (2a) or (2b) in the definition of protection-stability for μ at $(P, \succ)$, then (i, e) is a *blocking pair* to (or *blocks*) *μ at $(P, \succ)$ with respect to protective reserves.* We will use similar terminologies based on representation-stability and weak reserve-stability, namely, *blocking with respect to representative reserves* and *blocking with respect to weak reserves,* respectively.

3.4 Deferred Acceptance with Targeted Priority Reserves

In the standard Gale-Shapley agent-proposing Deferred Acceptance mechanism [27] entities simply follow the entity rankings for selection, which corresponds to not having any priority agents or not having any reserved seats. It yields the *optimal* stable assignment at every profile, where optimal means that each agent receives her most preferred assignment that the agent may receive in a stable matching. In other words, the optimal stable matching Pareto-dominates any other stable matching.

All matching mechanisms that we study are agent-proposing Deferred Acceptance (DA) mechanisms, but with non-standard choice functions for entities. We refer to choice functions as *entity selection rules.* The selection rule for each entity specifies the set of selected agents from each applicant pool $S \subseteq N$ which may arise in some step of the iterated agent-proposing Deferred Acceptance algorithm. When there is a priority reserve policy in place, the entity selection rule for each entity depends on the overall capacity of the entity as well as its reserve quota, and different entity selection rules differ in their priority reserve policies.

Now we formally define our proposed matching mechanism.

Deferred Acceptance with Targeted Priority Reserves (DA-TPR)

Fix a preference profile $P \in \mathcal{P}$.

Step 1: Each agent applies to her most preferred entity according to P. Given the fixed entity ranking profile $\succ$, each entity $e \in E$ selects from all of its applicants as follows:

1a. Select the highest-ranked q_e^u agents from all the applicants according to entity ranking $\succ_e$. If there are fewer than q_e^u applicants then select all of them.

1*b*. Select the highest-ranked q_e^r priority agents from the remaining applicants according to $\succ_e$. If there are fewer than q_e^r priority agents remaining then select the highest-ranked non-priority agents from the remaining applicants to fill the remaining positions.

Tentatively assign all selected agents to e and reject the rest.

Step k : Each agent who was rejected in step $k-1$ applies to her next most preferred acceptable entity according to P among the entities that have not yet rejected her, if any such entity remains. Each entity $e \in E$ selects from the set of agents it has been holding tentatively, together with the set of its new applicants in this step as specified below. We refer to this set of agents as A_e^k, the applicant pool of entity e at step k.

ka. Select the highest-ranked q_e^u agents from A_e^k according to entity ranking $\succ_e$. If there are fewer than q_e^u agents in A_e^k then select all of them.

kb. Select the highest-ranked q_e^r priority agents from the remaining agents in A_e^k according to $\succ_e$. If there are fewer than q_e^r priority agents remaining then select the highest-ranked non-priority agents from the remaining agents in A_e^k to fill the remaining positions.

Tentatively assign all selected agents to e and reject the rest.

The algorithm terminates when there are no agents whose proposal is rejected, and the tentative assignments become final.

Hafalir et al.'s [30] proposed mechanism, the DA-MiR (Deferred Acceptance with Minority Reserves) is a DA matching mechanism with an entity selection rule which aims to implement affirmative action with a minority reserve policy in the school choice context. It differs from the DA-TPR in that it first selects agents for the q_e^r reserved positions from the set of N^+ agents (priority agents in our terminology correspond to minority students in Hafalir et al. [30]), and then it fills the rest of the positions from all the remaining applicants. We give the definition of the DA-MiR mechanism below using our terminology.

Deferred Acceptance with Minority Reserves (DA-MiR)

Fix a preference profile $P \in \mathcal{P}$.

Step 1: Each agent applies to her most preferred entity according to P. Given the fixed entity ranking profile $\succ$, each entity $e \in E$ selects from all of its applicants as follows:

1a. Select the highest-ranked q_e^r priority agents from all the applicants according to entity ranking $\succ_e$.

1b. Select the highest-ranked q_e^u agents according to $\succ_e$ from the remaining applicants. If there are fewer than q^e applicants in total then select all of them.

Tentatively assign all selected agents to e and reject the rest.

Step k : Each agent who was rejected in step $k-1$ applies to her next most preferred acceptable entity according to $\succ_e$ among the entities that have not yet rejected her, if any such entity remains. Each entity $e \in E$ selects from its applicant pool A_e^k (defined as the union of the set of agents it has been holding tentatively and the set of its new applicants in this step) as specified below.

ka. Select the highest-ranked q_e^r priority agents from A_e^k according to $\succ_e$.

kb. Select the highest-ranked q_e^u agents according to $\succ_e$ from the remaining agents in A_e^k. If there are fewer than q^e applicants in total then select all of them.

Tentatively assign all selected agents to e and reject the rest.

The algorithm terminates when there are no agents whose proposal is rejected, and the tentative assignments become final.

The difference between DA-MiR and DA-TPR is the time when the reserved positions are filled. The DA-MiR fills the reserved positions first, whereas the DA-TPR fills the reserved positions last. Thus, the DA-TPR allows for selecting priority agents for unreserved seats up front, without using up reserved seats, priority agents who would have been selected on their own right in the absence of a priority reserve policy.

Example 5. DA-TPR versus DA-MiR entity selection and matching.

Let $N^+ = \{i_1, i_4, i_5\}$, $N^0 = \{i_2, i_3\}$, $E = \{e_1, e_2, e_3\}$, $q = (2, 1, 2)$, $q^r = (1, 0, 0)$.

Agent Preference Profile

P_{i_1}	P_{i_2}	P_{i_3}	P_{i_4}	P_{i_5}
$\underline{e_1}$	$\underline{e_3}$	$\underline{e_1}$	e_1	e_1
e_3	e_1	e_2	$\underline{e_2}$	e_2
e_2	e_2	e_3	e_3	$\underline{e_3}$

Entity Ranking Profile

$\succ_{e_1}$	$\succ_{e_2}$	$\succ_{e_3}$
i_1	i_1	i_1
i_2	i_2	i_2
i_3	i_4	i_3
i_4	i_5	i_4
i_5	i_3	i_5

DA-TPR

Step	e_1	e_2	e_3
1	$\underline{i_1}$ i_3 $\underline{i_4}$ i_5		$\underline{i_2}$
2	$\underline{i_1}$ $\underline{i_4}$	i_3 $\underline{i_5}$	$\underline{i_2}$
3	$\underline{i_1}$ $\underline{i_4}$	$\underline{i_5}$	$\underline{i_2}$ $\underline{i_3}$

DA-MiR

Step	e_1	e_2	e_3
1	$\underline{i_1}$ $\underline{i_3}$ i_4 i_5		$\underline{i_2}$
2	$\underline{i_1}$ $\underline{i_3}$	$\underline{i_4}$ i_5	$\underline{i_2}$
3	$\underline{i_1}$ $\underline{i_3}$	$\underline{i_4}$	$\underline{i_2}$ $\underline{i_5}$

The matchings that result from the two different matching mechanisms are indicated in the preference profile of the agents: the DA-TPR matching is in bold and the DA-MiR matching is underlined.

Using the selection of entity e_1 in Step 1 of this example, we can demonstrate the fundamental difference in entity selection for the two matching rules. Entity e_1 has one reserved position and one unreserved position, and its applicant pool in Step 1 is $A^k_{e_1} = \{i_1, i_3, i_4, i_5\}$. Given $\succ_{e_1}$, following the DA-TPR entity selection rule entity e_1 chooses i_1 for the unreserved position, and then priority agent i_4 is selected for the reserved position. By contrast, following the DA-MiR entity selection rule entity e_1 chooses i_1 for the reserved position, and then agent i_3 is selected for the unreserved position. Note that the DA-MiR selection rule allocates the reserved position to priority agent i_1 in this example, who would have been selected without a priority policy anyway, while priority agents i_4 and i_5 would have benefited from the reserved position. The DA-TPR selection rule, on the other hand, selects the top-ranked piority agent i_1 for the unreserved position up front, and allocates the reserved position to i_4 (the higher-ranked between the two priority agents in need of a reserved position at e_1.) Indeed, here the DA-MiR selection rule ends up taking the two top-ranked agents, so the reserve policy makes no difference in the DA-MiR selection in this case, despite having a reserved position.

Overall, observe that priority agents i_4 and i_5 are better off under the DA-TPR mechanism compared to the DA-MiR mechanism, while priority agent i_1 is indifferent. $\diamond$

In the following we will call a matching *optimal with respect to a stability axiom* if every agent receives at least as good an assignment in this matching as in any other matching satisfying the stability axiom. Specifically, a **matching** μ is **optimal with respect to protection-stability** at profile $(P, \succ)$ if, for each agent $i \in N$, $\mu(i) = e$ is the most preferred entity among all entities that agent i could be matched to at any protection-stable matching at $(P, \succ)$, when there is any such entity. Given $\succ \in \Pi$, a **matching mechanism** is **optimal with respect to protection-stability** if, for each preference profile $P \in \mathcal{P}$, it assigns a matching to profile $(P, \succ)$ that is optimal with respect to protection-stability. We define the optimality of a matching and a matching mechanism similarly with respect to representation-stability.

Theorem 5 (**Stability and optimality of the DA-TPR mechanism**).
The DA-TPR mechanism satisfies protection-stability and it is optimal with respect to protection-stability.

Proof. First we prove the protection-stability of the DA-TPR mechanism f^{TPR}. Fix a preference profile $P \in \mathcal{P}$. It is readily checked that condition 1 and the first statement on quota saturation in condition 2 of the definition of protection-stability hold at profile $(P, \succ)$, so suppose that there exists a blocking pair (i, e) at $(P, \succ)$ with respect to protective reserves, and thus either (2a) or (2b) does not hold. Let $\mu \equiv f^{\text{TPR}}(P, \succ)$. Then $eP_i\mu(i)$

by assumption, which implies that i applied to entity e in some step of the DA-TPR procedure f^{TPR} at $(P, \succ)$ and was rejected by e at this step or in a later step.

(2a) Let $i \in N^+$. It is clear that for all $j \in \mu_e$, $j \succ_e i$. Suppose that $|\bar{N}_e^+(\mu_e, \succ_e)| < q_e^r$. Then, since all reserved positions are filled last by entity e in each step, there exists $j \in N^0$ who is selected for a reserved position in some step of the DA-TPR procedure at $(P, \succ)$ and who is eventually matched to e. However, this implies that all priority agents who applied to entity e in any step of the DA-TPR procedure at $(P, \succ)$ have been selected by e, which contradicts $eP_i\mu(i)$.

(2b) Let $i \in N^0$. It is clear that for all $j \in (\mu_e \cap N^0)$, $j \succ_e i$. Suppose that $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| > q_e^r$. This implies that there exists at least one priority agent h who was selected for an unreserved position in the last step in the DA-TPR procedure at $(P, \succ)$, prior to the selection of priority agents for the reserved positions, such that $i \succ_e h$. Since $eP_i\mu(i)$, this is a contradiction.

We show next that the DA-TPR mechanism is optimal with respect to protection-stability. We will use the following terminology: entity e is *achievable* for agent j at profile $(P, \succ)$ if there is a protection-stable matching μ at $(P, \succ)$ such that $\mu(j) = e$.

Fix a preference profile $P \in \mathcal{P}$. Suppose, by contradiction, that there exist an entity e and agent j such that e rejects agent j in some step k of the DA-TPR procedure at $(P, \succ)$ and e is achievable for j. Assume without loss of generality that no agent was rejected by an entity that is achievable for the agent prior to step k of the DA-TPR procedure at $(P, \succ)$. Since j is rejected, q_e agents are matched to e in step k. Let this set of agents be $L \subset N$. Given the DA-TPR procedure, if $j \in N^0$ then, for all $l \in L$, $l \succ_e j$. On the other hand, if $j \in N^+$ then, for all $h \in (L \cap N^+)$, $h \succ_e j$. Let $H \equiv \{h \in (L \cap N^+) : \forall l \in (L \cap N^0), l \succ_e h\}$. Then $j \in N^+$ also implies that $|H| = q_e^r$.

Since e is achievable for j, there exists a protection-stable matching μ at $(P, \succ)$ such that $\mu(j) = e$. Then, given that $|L| = q_e$ and $j \notin L$, there exists $i \in L$ such that $\mu(i) \neq e$. Then, as shown above, if $j \in N^0$ then $i \succ_e j$, and if both $j \in N^+$ and $i \in N^+$ then we also have $i \succ_e j$.

Note that $\mu(i)$ is achievable for i, since μ is protection-stable. Furthermore, since no agent was rejected by an entity that is achievable for the agent prior to step k, by our assumption, we have $eP_i\mu(i)$.

Case 1: $i \in N^+$

Since μ is protection-stable, for all $h \in \mu_e$, $h \succ_e i$. Thus, $j \succ_e i$. However, as shown above, given that $i \in N^+$, we have $i \succ_e j$, regardless of whether $j \in N^0$ or $j \in N^+$. This is a contradiction.

Case 2: $i \in N^0$

Since μ is protection-stable, for all $h \in (\mu_e \cap N^0)$, $h \succ_e i$. Thus, if $j \in N^0$, $j \succ_e i$. However, as shown above, if $j \in N^0$ then $i \succ_e j$, which is a contradiction, and hence $j \in N^+$. This implies that for all $h \in H$, $h \succ_e j$, as shown above. Moreover, for each $h \in H$, either e is her best achievable entity at $(P, \succ)$ or h cannot achieve e, given our assumption on step k. But since $\mu(j) = e$ and μ is protection-stable, given that for all $h \in H$, $h \succ_e j$, it follows that $H \subset \mu_e$. This implies that $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| > q_e^r$, which contradicts (2b) in the definition of protection-stability. Given that μ is protection-stable, we have a contradiction.

Thus, we can conclude that μ is optimal with respect to protection-stability at $(P, \succ)$ for all preference profiles $P \in \mathcal{P}$ for an arbitrary fixed entity ranking profile $\succ \in \Pi$, and therefore the DA-TPR mechanism is optimal with respect to protection-stability. □

A matching mechanism f is **strategyproof** if for all $i \in N$, for all $P \in \mathcal{P}$ and all alternative preferences P_i' for agent i, $f_i(P) R_i f_i(P_i', P_{-i})$. A coalition $L \subseteq N$ can **manipulate** matching $f(P)$ at P if there exists P_L' such that for all $i \in S$, $f_i(P_L', P_{-L}) P_i f_i(P)$. A matching mechanism f is **weakly group-strategyproof** if there is no coalition that can manipulate the matching $f(P)$ at any preference profile $P \in \mathcal{P}$. Thus, weak group-strategyproofness implies strategyproofness. Note that this is the weaker version of group-strategyproofness, which is satisfied by the standard Deferred Acceptance mechanism.

Theorem 6 (**Characterization of the DA-TPR mechanism**).

The DA-TPR mechanism is the only protection-stable and strategyproof mechanism.

We can establish a similar result for the DA-MiR mechanism, using representation-stability instead of protection-stability. Hafalir et al. (2013)[30] shows that the DA-MiR mechanism satisfies representation-stability and is optimal with respect to representation-stability. We prove the following characterization result which is parallel to the characterization of the DA-TPR mechanism.

Theorem 7 (**Characterization of the DA-MiR mechanism**).

The DA-MiR mechanism is the only representation-stable and strategyproof mechanism.

The proofs of Theorems 6 and 7 are in Appendix B. Although these two characterizations only require strategyproofness, we will prove a result in Section 5 that implies that both the DA-TPR and the DA-MiR mechanisms are also weakly group-strategyproof. This has already been shown for the DA-MiR rule by Hafalir et al. (2013)[30].

Based on the observation that both protection-stability and representation-stability simplify to standard stability when there are no priority agents (or, alternatively, no reserved positions at any entity) and given that both the DA-TPR and the DA-MiR mechanisms become the standard DA mechanism under these conditions, our characterizations of the DA-TPR and DA-MiR mechanisms (Theorems 6 and 7) both simplify to the characterization of the standard DA mechanism in Alcalde and Barberà (1994)[4], which states that the only stable mechanism which is strategyproof is the DA mechanism.

3.5 The Class of Sequential Priority Reserves Mechanisms

In this section we broaden our inquiry and study a class of matching mechanisms which includes both the DA-TPR and DA-MiR mechanisms and allows for an arbitrary ordering of reserved and unreserved positions for the entity selection rules. If each entity selection rule starts with the unreserved positions then we get the DA-TPR mechanism, and if each entity selection rule starts with the reserved positions then we get the DA-MiR mechanism. However, all orderings of reserved and unreserved positions and all combinations of different entity selection rules lead to a matching mechanism, and we refer to this family of mechanisms as **Deferred Acceptance with Sequential Priority Reserves,** or **DA-SPR** mechanisms for short. Choice functions (in our terminology entity selection rules) with an arbitrary sequence of minority reserve and unreserved school seats have been examined in different models in the Boston and Chicago school choice models by Dur et al. (2018)[19] and (2020)[21].

All the matching mechanisms studied in this paper are Deferred Acceptance mechanisms with specific selection rules (i.e., choice functions) for entities, where the entity selection rule is applied in each step by each entity to their applicant pool, given the entity's ranking of agents. We can define a general Deferred Acceptance mechanism with an *entity selection rule* specified for each entity, instead of simply selecting the highest-ranked agents from the applicant pool A_e^k in each step k, as in the standard Deferred Acceptance mechanism, where the applicant pool $A_e^k \subseteq N$ in step k is the union of the set of new applicants to e and the set of agents tentatively matched to e in step $k-1$. Before introducing the DA-SPR mechanisms, we first define entity selection rules in general.

Entity selection rules. The selection rule for entity e, denoted by c_e, maps from any set of agents (the applicant pool) $A \subseteq N$ and any entity ranking $\succ_e$ to a subset of A as follows:

i. if $|A| \leq q_e$ then $c_e(A, \succ_e) = A$.

ii. if $|A| > q_e$ then $c_e(A, \succ_e) \subset A$ such that $|c_e(A, \succ_e)| = q_e$.

For DA-SPR mechanisms specifically, the entity selection rules are given by the set of all feasible sequences of reserved and unreserved positions: for any entity $e \in E$ the sequence consists of an arbitrary order of q_e^r reserved positions and q_e^u unreserved positions. Thus, for an entity e with $q_e^r = 2$ and $q_e^u = 3$, the sequence (u, r, r, u, u) is an example of a feasible sequence, where r denotes a reserved position and u denotes an unreserved position. We will refer to such a feasible sequence henceforth as an **entity selection sequence** for entity e and denote it by s_e.

DA with Sequential Priority Reserves (DA-SPR)

Fix a profile $s \equiv (s_e)_{e \in E}$ of entity selection sequences. Then the DA-SPR mechanism f^s specifies a matching for each preference profile $P \in (\mathcal{P}$ based on the iterative steps of the Deferred Acceptance mechanism using the specific entity selection rule s_e for each entity $e \in E$. Given the definitions of the DA-TPR and DA-MiR rules, instead of the specific entity selection sequences that these two rules use in steps ka. and kb., the DA-SPR mechanism f^s uses entity selection sequence s_e for each entity e. Specifically, each entity e in each step k ($k \geq 1$) of the procedure at preference profile P fills positions according to the specific sequence of reserved and unreserved positions in s_e as follows: to fill a reserved position it chooses the next highest-ranked priority agent from the remaining agents in A_e^k, and to fill an unreserved position it chooses the next highest-ranked remaining agent in A_e^k according to $\succ_e$. In case there are not enough priority group applicants in A_e^k to fill all q_e^r positions with priority agents, the entity fills the position by the next highest-ranked non-priority agent remaining in A_e^k.

It is readily seen from the above definition that the DA-SPR mechanisms constitute a family of Deferred Acceptance mechanisms with specific selection rules for entities, which are determined by the entity selection sequence s_e for each entity e, and thus DA-SPR mechanisms are parameterized by the profile of entity selection sequences s. The DA-SPR mechanism with $s_e = (u, \ldots, u, r, \ldots r)$ for each entity e, consisting of the q_e^u unreserved positions first and then the q_e^r reserved positions, is the DA-TPR mechanism, while the DA-SPR mechanism with $s_e = (r, \ldots, r, u, \ldots u)$ for each entity, consisting of the q_e^r reserved positions first and then the q_e^u unreserved positions, is the DA-MiR mechanism.

Since the DA-TPR mechanism is protection-stable by Theorem 4.1, it is also weakly reserve-stable. Similarly, since the DA-MiR mechanism is representation-stable (Hafalir et al. (2013)[30]), it is also weakly reserve-stable. We extend these results to the entire family of DA-SPR mechanisms.

Theorem 8 (**Stability property of DA-SPR mechanisms**).
DA-SPR mechanisms are weakly reserve-stable.

Proof. Fix a profile of entity selection sequences $s \equiv (s_e)_{e \in E}$ and let f^s denote the corresponding DA-SPR mechanism. Fix a preference profile $P \in \mathcal{P}$ and let $\succ \in \Pi$ be the fixed entity ranking profile. It is readily checked that condition 1 and the first statement on quota saturation in condition 2 of the definition of weak reserve-stability hold at $(P, \succ)$, so suppose that there exists a blocking pair (i, e) at $(P, \succ)$ with respect to weak reserves, and thus either (2a) or (2b) does not hold in the definition. Let $\mu \equiv f^s(P, \succ)$. Then $eP_i\mu(i)$ by assumption, which implies that i applied to entity e in some step of the DA-SPR procedure f^s at $(P, \succ)$ and was rejected by e at this step or in a later step.

(2a) Let $i \in N^+$. If there exists $j \in \mu_e$ such that $i \succ_e j$, then if j is selected by e for any (reserved or unreserved) position in the last step of the DA-SPR procedure f^s at $(P, \succ)$ then i would not have been have been rejected by e in either the last step or in any previous step of the procedure. Suppose that $|\mu_e \cap N^+| < q_e^r$. This means that all priority agents who applied to entity e in any step of the DA-SPR procedure f^s at $(P, \succ)$ have been selected by e, contradicting the fact that i had applied.

(2b) Let $i \in N^0$. If there exists $j \in (\mu_e \cap N^0)$ such that $i \succ_e j$, then if j is selected by e

for any (reserved or unreserved) position in the last step of the DA-SPR procedure f^s at $(P, \succ)$ then i would not have been rejected by e in either the last step or in any previous step of the procedure. Suppose that $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| > q_e^r$. This implies that there exists at least one priority agent h who was selected for an unreserved position in the last step in the DA-TPR procedure at $(P, \succ)$, such that $i \succ_e h$. But then this unreserved position would have been filled in either the last or in some previous step by i, when i applied to e, given that $eP_i\mu(i)$, and this unreserved position would have been filled in the last step either by agent i or by some agent j such that $j \succ_e i$. Since $i \succ_e h$, $j \neq h$, and we have a contradiction.

In either case there is no pair (i, e) that blocks μ with respect to weak reserves at $(P, \succ)$. Therefore, f^s is weakly reserve-stable. Since f^s is an arbitrary DA-SPR mechanism, the result follows. □

Recall that the DA-TPR mechanism is optimal with respect to protection-stability (Theorem 5) and the DA-MiR mechanism is optimal with respect to representation-stability (Hafalir et al. (2013)[30]). Optimality implies uniqueness, so we cannot establish an optimality result for the class of DA-SPR mechanisms with respect to weak reserve-stability. We will show, however, that these mechanisms are constrained efficient subject to weak reserve-stability, which means that a DA-SPR mechanism assigns a matching to each profile such that this matching is not only weakly reserve-stable, but it is also not Pareto-dominated by another weakly reserve-stable matching at the given profile.

Constrained efficiency. A matching mechanism f is constrained efficient with respect to weak reserve-stability if, for all preference profiles $P \in \mathcal{P}$, whenever a matching ν Pareto-dominates $f(P)$ at P, ν is not weakly reserve-stable at P.

Theorem 9 (Constrained efficiency property of DA-SPR mechanisms). *DA-SPR mechanisms are constrained efficient with respect to weak reserve-stability.*

Proof. Fix a profile of entity selection sequences $s \equiv (s_e)_{e\in E}$ and let f^s denote the corresponding DA-SPR mechanism. Suppose, by contradiction, that $P \in \mathcal{P}$ is a preference profile such that $f^s(P)$ is Pareto-dominated by $\nu \in \mathcal{M}$, where ν is a weakly reserve-stable matching at $(P, \succ)$, given the fixed entity ranking profile $\succ \in \Pi$. Let $\mu \equiv f^s(P)$. Then there exists $j^* \in N$ such that $\nu_{j^*} P_{j*} \mu_{j^*}$. Let $e \equiv \nu(j^*)$. Assume without loss of generality that j^* gets rejected by e in step k of the DA-SPR procedure f^s applied to P such that there is no agent $h \in N$ who gets rejected by $\nu(h)$ in any step prior to k. Let $J \subset N$ be the set of those agents who are rejected by e in step k of the procedure and who are assigned to e in matching ν. Then, for all $j \in J$, $\nu(j) = e$ and $\mu(j) \neq e$. Note that $j^* \in J$ and thus $J \neq \varnothing$.

Let $I \subset N$ be a set of agents such that for all $i \in I$, i is tentatively matched to e in step k of the DA-SPR procedure f^s applied to (P) and $\nu(i) \neq e$. In addition, choose I to satisfy $|I| = |J|$. Note that such a set I exists, since q_e agents are matched temporarily to e in step k of the procedure, given that agents in J get rejected by e in step k, and $|\nu_e| \leq q_e$. If $\nu(i) P_i e$ for some $i \in I$ then our assumption on step k of the DA-SPR procedure f^s applied to P does not hold, and since $\nu(i) \neq e$ this implies that for all $i \in I$, $e P_i \nu(i)$.

Let $L \equiv (\nu_e \setminus J) \cup I$. Note that $|L| = q_e$ since $J \subset \nu_e$ and $\nu_e \cap I = \varnothing$. We will show that L is exactly the set of agents who are tentatively matched to e in step k of the DA-SPR procedure f^s applied to P. Note first that this set has q_e agents since j^* is rejected by e in step k. Let l be an arbitrary agent in this set. Then $l \notin J$. Thus, if $l \in \nu_e$ then $l \in L$. Moreover, if $l \notin \nu_e$ then $l \in I$, so $l \in L$ in this case as well. This proves that every agent who is tentatively matched to e in step k is in L. Since both sets have cardinality q_e, the equality of the two sets follows.

Case 1. There exist $i \in I$ and $j \in J$ such that $i, j \in N^+$.
Since $i, j \in N^+$ and since i is selected while j is rejected by e in step k of the DA-SPR procedure f^s applied to P, $i \succ_e j$ holds. Given $\nu(j) = e$ and $e P_i \nu(i)$, this contradicts the first requirement in part (2a) in the definition of weak reserve-stability for ν at $(P, \succ)$.

Case 2. There exist $i \in I$ and $j \in J$ such that $i, j \in N^0$.

Since $i, j \in N^0$ and since i is selected while j is rejected by e in step k of the DA-SPR procedure f^s applied to P, $i \succ_e j$ holds. Given $\nu(j) = e$ and $eP_i\nu(i)$, this contradicts the first requirement in part (2b) of the definition of weak reserve-stability for ν at $(P, \succ)$.

Case 3. For all $i \in I$, $i \in N^+$, and for all $j \in J$, $j \in N^0$.

Note first that if there exist $i \in I$ and $j \in J$ such that $i \succ_e j$ then we get the same contradiction as in Case 1. Thus, for all $i \in I$ and $j \in J$, $j \succ_e i$. Suppose $|L \cap N^+| > q_e^r$. Then, since L is the set of agents who are tentatively matched to e in step k of the DA-SPR procedure f^s applied to P, there exists $j \in J$ such that j should not have been rejected by e in step k. Thus, $|L \cap N^+| \leq q_e^r$. Then, given $|J| = |I| \geq 1$, $|\nu_e \cap N^+| < q_e^r$, which contradicts the second requirement in part (2a) of the definition of weak reserve-stability for ν at $(P, \succ)$.

Case 4. For all $i \in I$, $i \in N^0$, and for all $j \in J$, $j \in N^+$.

Fix $i \in I$. Since L is the set of agents who are tentatively matched to e in step k of the DA-SPR procedure f^s applied to P, $i \in (L \cap N^0)$, and $j^* \in J$ is a priority agent who is rejected by e in step k, $\{h \in (L \cap N^+) : i \succ_e h\} \geq q_e^r$. Note that for all $j \in J$, since i is selected while j is rejected by e in step k of the DA-SPR procedure f^s applied to P, $i \succ_e j$ holds. Then, given $|J| = |I| \geq 1$, $\{h \in (\nu_e \cap N^+) : i \succ_e h\} > q_e^r$, which contradicts the second requirement in part (2b) of the definition of weak reserve-stability for ν at $(P, \succ)$.

We have shown that there is a contradiction in each possible case. Therefore, $f^s(P)$ is not Pareto-dominated for any DA-SPR mechanism f^s at any preference profile P by any weakly reserve-stable matching, which means that all DA-SPR mechanisms are constrained efficient with respect to weak reserve-stability. □

We state next that not only the DA-TPR and DA-MiR mechanisms are strategyproof, as established before (see Theorems 4.2 and 4.3), but also the entire family of DA-SPR mechanisms is weakly group-strategyproof, implying that they are also strategyproof.

Theorem 10 (**Incentive property of DA-SPR mechanisms**).
DA-SPR mechanisms are weakly group-strategyproof.

Proof. We use the augmented market $M^{A(s)}$, a one-to-one matching market with one position for each entity, as defined in Appendix A, to show that an arbitrary DA-SPR mechanisms f^s is weakly group-strategyproof. By Lemma A.1 in Appendix A, the agent-proposing standard DA mechanism applied to the augmented market $M^{A(s)}$ leads to a matching that corresponds to the matching assigned by the DA-SPR mechanism f^s to the original market M. Now suppose, by contradiction, that there exists a market M with profile $(P, \succ)$ in which a coalition $L \subseteq N$ can manipulate the DA-SPR matching $f^s(P)$ at profile P, given $\succ$. Then coalition L can also manipulate the agent-proposing DA matching $f^{\text{DA}}(P^{A(s)})$ in the corresponding augmented market $M^{A(s)}$ at profile $P^{A(s)}$, given $\succ^A$. This is a contradiction, since the agent-proposing DA mechanism is weakly group-strategyproof (Dubins and Freedman (1981)[18], Roth (1982)[47]). □

With these results in hand, we are now ready to establish our main theorem stated below. Note that this characterization requires strategyproofness only, although all DA-SPR mechanisms satisfy weak group-strategyproofness as well, as we have just shown. The proof of Theorem 11 is relegated to Appendix B.

Theorem 11 (**Characterization of DA-SPR mechanisms**).
A matching mechanism is weakly reserve-stable and strategyproof if and only if it is a DA-SPR mechanism.

Intuitively speaking, the set of weakly reserve-stable mechanisms is large, since weak reserve-stability allows for all priority violations that are consistent with different ways of administering the priority reserve policy. In particular, it only requires that at least q_e^r positions are filled by priority agents at any entity e that is preferred by a priority agent to her assignment in a given matching, and that no more than q_e^r priority agents are assigned to an entity who are ranked lower by the entity than a non-priority agent who prefers this entity to his assignment in the given matching. Our characterization theorem above tells us that requiring strategyproofness narrows down the large set of weakly reserve-stable mechanisms to DA-SPR mechanisms.

3.6 Selection Sequences and Consistent Selection

Sequential priority reserve policies may be ambiguous in their policy impact. The following example demonstrates that, depending on the entity ranking, two entity selection sequences may produce inconsistent results when considering the number of selected priority agents.

Example 6. Ambiguous policy impact of DA-SPR mechanisms.

Fix an entity $e \in E$. Consider applicant pool $A = A^0 \cup A^+$, where $A^0 = \{i_1, i_2, i_3, i_4, i_5\} \subseteq N^0$ and $A^+ = \{i_6, i_7, i_8\} \subseteq N^+$. Let $q_e^r = 2$ and $q_e^u = 2$. Consider entity rankings $\succ_e$: $(i_6, i_1, i_2, i_7, i_3, i_8)$ and $\hat{\succ}_e$: $(i_1, i_6, i_7, i_3, i_2, i_8)$, and two entity selection sequences: $s_e = (u, r, r, u)$ and $s'_e = (r, u, u, r)$.

When the entity ranking is $\succ_e$, the two entity selection sequences result in the following selections from A (priority agents are indicated in bold below):

$$s_e \text{ yields } \begin{pmatrix} \boldsymbol{i_6} & \boldsymbol{i_7} & \boldsymbol{i_8} & i_1 \\ u & r & r & u \end{pmatrix}$$

with a total of *three* priority agents selected by e, and

$$s'_e \text{ yields } \begin{pmatrix} \boldsymbol{i_6} & i_1 & i_2 & \boldsymbol{i_7} \\ r & u & u & r \end{pmatrix}$$

with a total of *two* priority agents selected by e.

When the entity ranking $\hat{\succ}_e$ is used, the two entity selection sequences result in the following selections from A:

$$s_e \text{ yields } \begin{pmatrix} i_1 & \boldsymbol{i_6} & \boldsymbol{i_7} & i_3 \\ u & r & r & u \end{pmatrix}$$

with a total of *two* priority agents selected by e, and

$$s'_e \text{ yields} \begin{pmatrix} \boldsymbol{i_6} & i_1 & \boldsymbol{i_7} & \boldsymbol{i_8} \\ r & u & u & r \end{pmatrix}$$

with a total of *three* priority agents selected by e. ◇

The example shows that we cannot necessarily say which entity selection rule is more favorable for the priority group, since this depends on the entity rankings. In light of this, we define a property of a pair of entity selection rules which eliminates the above inconsistency.

Selection consistency. We call a pair of entity selection rules c_e and c'_e **selection inconsistent** if there exist an applicant pool $A \subseteq N$ and a pair of entity rankings $\succ_e$ and $\hat{\succ}_e$ such that

$$|\, c_e(A, \succ_e) \cap N^+| > |\, c'_e(A, \succ_e) \cap N^+|$$

and

$$|\, c_e(A, \hat{\succ}_e) \cap N^+| < |\, c'_e(A, \hat{\succ}_e) \cap N^+|.$$

Two entity selection rules are **selection consistent** if they are not selection inconsistent.

That is, if the number of priority agents chosen by selection rule c_e from A is greater than the number of priority agents chosen by c'_e from A under given entity preferences, then under any other entity preferences the number of priority agents selected by c_e from A cannot be less than the number selected by c'_e, ensuring that the two entity selection rules are consistent with each other in their policy impact.

We introduce next a more demanding criterion for comparing two entity selection rules.

N^+-superiority. A selection rule c_e is **N^+-superior** to another selection rule c'_e for entity e if, for all $A \subseteq N$ and all entity rankings $\succ_e$, we have $(c'_e(A, \succ_e) \cap N^+) \subseteq (c_e(A, \succ_e) \cap N^+)$. That is, if a priority agent is chosen by selection rule c'_e from A then this priority agent is also chosen by selection rule c_e from A, given a fixed entity ranking

$\succ_e$.

It can easily be verified that if one entity selection rule is N^+-superior to another one then the two entity selection rules are selection consistent. It is also clear that the converse does not hold, since selection consistency does not require set inclusion of the selected priority agents. However, all entity selection rules of interest follow the entity ranking among priority agents (i.e., do not allow for violations of entity rankings between any two priority agents), so the converse also holds for such entity selection rules. We obtain a more general equivalence result for DA-SPR selection sequences, which involves the comparability of entity selection sequences as defined below.

Let $\Sigma(s_e)$ denote the list of the cumulative reserved position count in the entity selection sequence s_e for entity $e \in E$. Thus, for example, if $s_e = (u, r, r, u, u)$ then $\Sigma(s_e) = (0, 1, 2, 2, 2)$. Two selection sequences s_e and s'_e for entity e are **comparable** if either $\Sigma(s_e) \leq \Sigma(s'_e)$ or $\Sigma(s_e) \geq \Sigma(s'_e)$, meaning that two corresponding entries in the two lists of cumulative reserved position counts are either the same or the strict inequality holds in the required direction consistently. If two distinct entity selection sequences s_e and s'_e are comparable, we call s_e **more targeted** than s'_e if $\Sigma(s_e) \leq \Sigma(s'_e)$, and **less targeted** if $\Sigma(s_e) \geq \Sigma(s'_e)$.

Note that in Example 6 s_e and s'_e are not comparable, since $\Sigma(s_e) = (0, 1, 2, 2)$ and $\Sigma(s'_e) = (1, 1, 1, 2)$. However, $s''_e = (u, r, u, r)$ is comparable to s_e, since $\Sigma(s''_e) = (0, 1, 1, 2) \leq \Sigma(s_e)$. In particular, s''_e is more targeted than s_e. Moreover, since $\Sigma(s''_e) \leq \Sigma(s'_e)$, s''_e is also more targeted than s'_e. Note, finally, that s''_e is also N^+-superior to both s_e and s'_e.

Theorem 12 (Consistency and comparability of selection sequences).
Given two DA-SPR entity selection sequences, the following statements are equivalent.

1. *The two entity selection sequences are selection consistent.*

2. *One entity selection sequence is N^+-superior to the other one.*

3. The two entity selection sequences are comparable.

Proof. 1. implies 2. since DA-SPR selection sequences follow the entity ranking when choosing priority agents. It is immediate that 3. implies 1. Thus, what remains to be shown is that 2. implies 3. Suppose, by contradiction, that there exist two DA-SPR entity selection sequences s_e and s'_e that are not comparable, but s_e is N^+-superior to s'_e. Since the two selection sequences are not comparable, s_e is not more targeted than s'_e, and thus there exists a position, say the kth position, in these sequences such that the cumulative count of reserved positions up to (and including) the kth position is strictly higher for s_e than for s'_e. Let the applicant pool be N and let $\succ_e$ rank agents according to s_e up to the kth rank as follows: for each reserved position $\succ_e$ ranks a priority agent and for each unreserved position it ranks a non-priority agent. Starting from the $k+1$st rank, let $\succ_e$ rank all remaining non-priority agents first and then all the remaining priority agents. Then s_e with ranking $\succ_e$ selects the top k agents regardless of their priority status for the first k positions, and then it selects additional priority agents for the remaining reserved positions only. This means that s_e selects exactly q_e^r priority agents. Note that s'_e with ranking $\succ_e$ cannot select fewer priority agents for the first k positions than s_e selected for the same positions, since selecting exactly the top k agents means selecting as few priority agents as possible for these positions. Starting from the $k+1$st position, s'_e selects additional priority agents for the remaining reserved positions only. Since $|N| > q_e$, $|N^0| > q_e^u$ and $|N^+| > q_e^r$ by assumption, and the cumulative count of reserved positions starting from the $k+1$st position is strictly higher for s'_e than for s_e, s'_e selects more priority agents from N than s_e, and the set of priority agents selected by s'_e is a strict superset of the set of priority agents selected by s_e: $(s_e(N, \succ_e) \cap N^+) \subset (s'_e(N, \succ_e) \cap N^+)$. This contradicts the assumption that s_e is N^+-superior to s'_e. □

We remark that the implication that 3. implies 2. generalizes the observations in Proposition 2 of Dur et al. (2018)[19] and can be verified in a straightforward manner. As mentioned in the Introduction, Theorem 12 is also related to some findings in Dur et al. (2020)[21] and Pathak et al. (2020)[46] in different contexts.

3.7 Comparing DA-SPR Mechanisms: Policy Implications

We specify two criteria for comparing matching rules in terms of the welfare of priority agents: the first one compares the outcomes at all profiles, which will be shown to be too strong a criterion, and thus we introduce a second criterion which restricts comparisons to the "benchmark" homogeneous profiles, where all entity rankings are the same and all the agents have the same preferences over entities. We focus on comparing the outcomes at homogeneous profiles because, as we show in Theorem 13, it is not possible to compare DA-SPR mechanisms on the basis of their priority policy impact when we consider all profiles. Unlike in Example 1, where the DA-TPR mechanism leads to a Pareto-improvement for priority agents compared to the DA-MiR mechanism, we can find profiles where some priority agents are worse off under the DA-TPR mechanism than under the DA-MiR mechanism. Restricting attention to homogeneous profiles is a reasonable relaxation in this context, partly because in many real-life applications a high correlation in both entity rankings and agent preferences is likely, but more importantly because homogeneous profiles serve as a litmus test of priority agent treatment, without the "noise" of the heterogeneous entity rankings and agent preferences.

N^+-domination. A matching mechanism f **N^+-dominates** another mechanism g if f Pareto-dominates g for the set of priority agents N^+ for an arbitrary entity ranking profile $\succ \in \Pi$. A matching mechanism f **homogeneously N^+-dominates** another mechanism g if, given that the fixed entity ranking profile $\tilde{\succ} \in \Pi$ is homogeneous, f Pareto-dominates g for the set of priority agents N^+ at every homogeneous preference profile $\tilde{P} \in \mathcal{P}$.

We show first that the lack of comparability holds in general for all mechanisms that satisfy weak reserve-stability and strategyproofness, and thus, by Theorem 11, for all DA-SPR mechanisms. Since no two DA-SPR mechanisms are comparable this way, it follows, perhaps somewhat surprisingly, that the DA-TPR mechanism does not Pareto-dominate the DA-MiR mechanism for priority agents, that is, it does not N^+-dominate it.

Theorem 13 (No N^+-domination).

Given two strategyproof and weakly reserve-stable mechanisms, neither N^+-dominates the other.

Proof. By Theorem 11, we only need to show that, for two given DA-SPR mechanisms, neither N^+-dominates the other. Fix two DA-SPR mechanisms, f and g. Since f and g are two distinct mechanisms, there exists entity $e \in E$ such that $s_e^f \neq s_e^g$, where s_e^f is the entity selection sequence for e in mechanism f, and s_e^g is the entity selection sequence for e in mechanism g. Assume without loss of generality that the kth position ($k > 0$) is the first one which is different between the two selection sequences.

Case 1: *The kth position in s_e^f is an unreserved position, while the kth position in s_e^g is a reserved position.*

We specify entity ranking $\succ_e$ for entity e as follows. Prior to the kth position, let $\succ_e$ rank a priority agent corresponding to each reserved position, and let it rank a non-priority agent corresponding to each unreserved position in the two entity selection sequences s_e^f and s_e^g. Moreover, let $\succ_e$ rank a priority agent kth. After the kth rank, let $\succ_e$ rank all the remaining non-priority agents, and finally all the remaining priority agents. Note that at least one priority agent is ranked at the end, lower than the priority agent in the kth rank, since $|N^+| > q_e^r$ by assumption. For all $i \in N$, let $P_i \in (e)$. Let $j \in N^+$ be the lowest-ranked priority agent by $\succ_e$ such that j is selected from N by e with ranking $\succ_e$, that is, $j \in s_e^f(N, \succ_e)$. Then, since there is one more reserved position after the kth position in s_e^f than in s_e^g, and given that $|N^0| > q_e^u$ and all remaining priority agents who were not selected for the first k positions by either entity selection rule are ranked last by $\succ_e$, $j \notin s_e^g(N, \succ_e)$. This means that $f_j(P) = e$ and $g_j(P) \neq e$, where the rest of the profile is arbitrary (i.e., with the exception of $P_i \in (e)$ for all $i \in N$ and the specification of $\succ_e$). Therefore, $f_j(P) P_j \, g_j(P)$, and since $j \in N^+$ it follows that g does not N^+-dominate f.

Case 2: *The kth position in s_e^f is a reserved position, while the kth position in s_e^g is an unreserved position.*

Note first that there exists k' with $k' > k$ such that the k'th position in s_e^f is an unreserved position, and the k'th position in s_e^g is a reserved position. Without loss of generality, let k' be the first such position. We specify entity ranking $\succ_e$ for entity e as follows. Up to and including the kth position, let $\succ_e$ rank a priority agent corresponding to each reserved position, and let it rank a non-priority agent corresponding to each unreserved position in s_e^f. Starting at the $k+1$st position, let $\succ_e$ rank all remaining priority agents first and then all remaining non-priority agents, with one exception: let $\succ_e$ rank a non-priority agent k'th. Such a non-priority agent exists by our assumption that $|N^0| \geq q_e^u$. Note that, given the applicant pool N, entity e with selection rule s_e^f would select a non-priority agent $h \in N^0$ who would not be selected by e with selection rule s_e^g, since s_e^f would lead to choosing a priority agent for the kth position but not for the k'th position, while s_e^g would lead to choosing a priority agent for both the kth and k'th positions (given the assumption that $|N^+| > q_e^r + 1$), and otherwise both entity selection rules would lead to selecting the same agents. Then there exists a priority agent $j \in N^+$ such that, given the applicant pool N, entity e with selection rule s_e^g would select agent j who would not be selected by e with selection rule s_e^f. Let all $q_e - 1$ agents who would be selected from N by e with ranking $\succ_e$ according to both s_e^f and s_e^g rank entity e first in preference profile P. Also, let $P_h \in (e, e')$ and $P_j \in (e, j)$. Furthermore, let $e' \in E \setminus \{e\}$ and let $q_{e'}$ agents who would not be selected by entity e with ranking $\succ_e$ from N according to either s_e^f or s_e^g rank e' first. Note that $q_{e'}$ such agents exist by our assumption that for all $e, e' \in E$, $|N| > q_e + q_{e'}$. Include in this set of $q_{e'}$ agents a priority agent $i \in N^+$. Note that, by our assumption that $|N^+| > q_e^r + 1$, such an agent exists. Let $\succ_{e'}$ rank i last and let $P_i \in (e')$. Let all other agents have no acceptable entities, and let the rest of profile $(P, \succ)$ be arbitrary.

At preference profile P under mechanism f entity e selects its top q_e-ranked agents in the first round of the procedure, including agent h but not agent j. The first round is the final round under mechanism f, since only j gets rejected in the first round and j has no acceptable entities other than e, the entity that rejected j in the first round. On the other hand, at profile $(P, \succ)$ under mechanism g entity e selects the same agents as under mechanism f, except that h is rejected and j is accepted by e. This means that h applies to e' in the second round under mechanism g, and since e' accepted all of its $q_{e'}$ applicants in the first round, including $i \in N^+$, and since $\succ_{e'}$ ranks i last, $h \in N^0$ is accepted and i is rejected by e' in the second round, given that entity e' has at least one unreserved position. This implies that $f_i(P) = e'$ and $g_i(P) \neq e'$. Thus, $f_i(P) P_i\, g_i(P)$ and, since $i \in N^+$, it follows that g does not N^+-dominate f. Finally, note that since f and g were two arbitrary DA-SPR mechanisms, the proof is completed. □

The following result follows directly from Theorems 11 and 13.

Corollary 3.

Given two DA-SPR mechanisms, neither N^+-dominates the other.

One DA-SPR **mechanism** is **more targeted** than another one if all the entity selection sequences are either the same or more targeted for the first one than for the second one. We define the concept of a **less targeted** DA-SPR mechanism similarly. Clearly, the unique most targeted DA-SPR mechanism is the DA-TPR mechanism, and the unique least targeted mechanisms is the DA-MiR mechanism, the two extreme members of the family of DA-SPR mechanisms. We will say that two DA-SPR **mechanisms** are **comparable** if one is more targeted than the other one.

In the proof of Theorem 13, Case 1 is the more intuitive case, since its proof implies that a less targeted mechanism does not N^+-dominate a more targeted one, but note that this proof applies much more generally. Case 2 is more surprising at first glance, however, since its proof implies, among others, that a more targeted mechanism does not N^+-dominate a less targeted one and, specifically, that the DA-TPR mechanism does not N^+-dominate the DA-MiR mechanism, contrary to expectation. Despite the fact that in each round of the DA-TPR mechanism each entity always selects at least as many priority agents as the DA-MiR mechanism (and often more) if they have the same applicant pool, as shown by Theorem 12, it is not necessarily the case that the DA-TPR outcome Pareto-dominates the DA-MiR outcome for priority agents at any given arbitrary profile. This is due to the fact that when the profile is not homogeneous and entities face different applicant pools depending on the selections made by other entities in previous steps, even one small change may trigger different rejection chains along the steps of the iterative procedure of the Deferred Acceptance mechanism, resulting in a series of differences in the assignments of agents, which makes a consistent comparison of matching outcomes at all profiles impossible. In particular, such rejection chains at non-homogeneous profiles can cause some priority agents to be worse off under the DA-TPR mechanism than under the DA-MiR mechanism, and this fact is captured in a more generally applicable form by the proof of Case 2. Thus, an understanding of how the Deferred Acceptance algorithm works makes this result more clear. A similar phenomenon has been identified by Kojima (2012)[39] when the reserve numbers are increased, as opposed to changing the entity selection sequences that we study here, which may also appear mystifying at first, just like the findings presented here. Although both types of changes in the priority policy clearly favor priority agents in the entity selection, due to the inner workings of the Deferred Acceptance mechanism neither of these changes favors the priority agents unambiguously when considering the entire mechanism, as reflected by a comparison of the final matching outcomes at an arbitrary profile.

Although N^+-domination is not possible among DA-SPR rules, we can establish N^+-domination results for any two DA-SPR mechanisms that are comparable. We know that priority agent selection is increased under a more targeted mechanism compared to a less targeted mechanism at the entity selection level without raising the number of reserved positions, just by changing the ordering of reserved and unreserved positions in the entity selection sequence (see Theorem 12). We prove below that this carries over to the entire matching mechanism when we restrict our attention to homogeneous profiles. Our result holds not only for the DA-TPR versus the DA-MiR mechanisms, but also extends to any two comparable DA-SPR mechanisms.

In the proof of the theorem we show that when the entities have homogeneous rankings and the agents have homogeneous preferences, each DA-SPR mechanism is equivalent to a serial dictatorship of the entities, and each entity chooses in turn from all the remaining agents according to the entity selection rule specified by the DA-SPR mechanism in question. This is shown to imply that when one DA-SPR mechanism is more targeted than another, the more targeted mechanism leads to a matching which weakly Pareto-dominates the matching outcome reached by the less targeted mechanism for priority agents. This is not by accident, and demonstrates that a more targeted mechanism is preferable to a less targeted mechanism when we wish to favor priority agents more. We note that it follows from the theorem that the DA-TPR outcome Pareto-dominates the DA-MiR outcome for the set of priority agents at any homogeneous profile.

Theorem 14 (**Homogeneous N^+-domination**).

Given two comparable DA-SPR matching mechanisms, the more targeted DA-SPR mechanism homogeneously N^+-dominates the less targeted DA-SPR mechanism.

Proof. Let f and f' be two DA-SPR mechanisms such that f is more targeted than f'. Let $s \equiv (s_e)_{e\in E}$ and $s' \equiv (s'_e)_{e\in E}$ be the profiles of entity selection sequences for f and f' respectively. Fix a homogeneous preference profile $\tilde{P} \in \mathcal{P}$ and a homogeneous entity ranking profile $\tilde{\succ} \in \Pi$. For all $i \in N$, let $P_i \in (e_1, e_2, \ldots, e_t, i)$, where $t \leq |E|$. Observe that, given $\tilde{\succ}$, both $f(\tilde{P})$ and $f'(\tilde{P})$ are obtained by the acceptable entities in P_i choosing sequentially among all remaining agents according to their respective entity selection sequences in f and f', where the sequence of the entities is determined by the ordering of the entities in the common preference ordering P_i. That is, first e_1 picks q_{e_1} agents, then e_2 picks q_{e_2} agents among the remaining agents, and so on, all the way to entity e_t. Since entity rankings are also homogeneous, both of these mechanisms correspond to one entity with T positions selecting according to the combined entity selection sequences of all entities in the order of $e_1, e_2, \ldots, e_t$, where $T \equiv \sum_{v=1}^{t} q_{e_v}$. Let s_T denote the combined sequence of length T consisting of selection sequences $s_{e_1}, s_{e_2}, \ldots, s_t$, in this order. Similarly, let s'_T denote the combined sequence of length T consisting of selection sequences $s'_{e_1}, s'_{e_2}, \ldots, s'_t$, in this order. Since for all $v = 1, \ldots, t$, s_{e_v} is more targeted than s'_{e_v}, it follows that s_T is more targeted than s'_T, that is, $\Sigma(s_T) \leq \Sigma(s'_T)$.

Observe that, given a fixed entity ranking, an entity with a more targeted entity selection sequence selects all the priority agents that it would select with a less targeted entity selection sequence, and sometimes more, as shown by Theorem 12. Therefore, mechanism f homogeneously N^+-dominates mechanism f'. □

3.8 Policies for Compromising Between Protection and Representation

Since not all DA-SPR mechanisms are comparable to each other, and arbitrary members of this family of mechanisms may be less than transparent in their policy implications, as demonstrated by Example 6, we now introduce a subclass of the DA-SPR mechanisms, the Split DA-SPR mechanisms, which are intuitive and provide consistent and comparable priority treatment policies.

Split DA-SPR mechanisms

A **Split DA-SPR entity selection sequence** s_e for entity $e \in E$ is of the following form:

$$s_e = (\underbrace{r, \ldots, r,}_{q_e^{rf}} \underbrace{u, \ldots, u,}_{q_e^u} \underbrace{r, \ldots, r}_{q_e^{rb}}),$$

where q_e^{rf} is the number of reserved positions at the front and q_e^{rb} is the number of reserved positions at the back of the sequence, with $q_e^r = q_e^{rf} + q_e^{rb}$. The reserved positions are split between the beginning and end of the sequence (the front and the back) with all unreserved positions between the reserved positions. Note that the DA-TPR and DA-MiR entity selection sequences are both Split DA-SPR sequences. In particular, a Split DA-SPR selection sequence with $q_e^{rb} = q^r$ is a DA-TPR entity selection sequence, and a Split DA-SPR selection sequence with $q_e^{rf} = q^r$ is a DA-MiR selection sequence. Observe that any two Split DA-SPR entity selection sequences are comparable: a lower

q_e^{rf} (or, equivalently, a higher q_e^{rb}) leads to a more targeted Split DA-SPR rule. Since the Split DA-SPR selection sequences are comparable, Theorem 12 tells us that they are unambiguous in their relative policy impact.

Split DA-SPR mechanisms are DA-SPR mechanisms such that each entity uses a Split DA-SPR entity selection sequence. Moreover, if two Split DA-SPR mechanisms are comparable, which requires that all of the selection sequences of the more targeted mechanism are either more targeted or the same as the selection sequences of the less targeted mechanism when compared for each entity, then Theorem 14 implies that their relative policy impact is unambiguous at homogeneous profiles. We summarize these results in the corollary below.

Corollary 4. *Let f and g be two Split DA-SPR mechanisms such that f is more targeted than g. Then for each entity $e \in E$ the entity selection rule used by f is either the same or N^+-superior to the entity selection rule used by g. Moreover, mechanism f homogeneously N^+-dominates mechanism g.*

As Corollary 2 shows, Split DA-SPR entity selection rules and mechanisms possess some nice properties regarding their policy impact. They are also quite intuitive, since in a Split DA-SPR mechanism each entity selects agents according to a split sequential priority reserve policy in which reserved positions are either at the front or at the end, with the unreserved positions in the middle. The reserved positions at the front can be seen as "non-targeted" positions, and the reserved positions at the back as "targeted" positions, positions to be given to priority agents who would be assigned to a less preferred entity without such a targeted position. Despite not being fully comparable, as shown by Corollary 3, due to their transparency Split DA-SPR mechanisms offer attractive policy choices for a designer who wishes to compromise between the fully targeted DA-TPR mechanism and the fully non-targeted DA-MiR mechanism.

3.9 Conclusion

This study brings a new perspective to the analysis of priority reserves and emphasizes the importance of clarifying the objective for providing priority treatment in diverse resource allocation problems, such as refugee settlement, centralized university admissions, or school choice. The previous literature on reserve policies focuses mainly on affirmative action and implicitly views representation as the primary goal of priority treatment, without making this assumption explicit. We seek an understanding of priority treatment policies in a more general framework by considering representation and effective preferential treatment (what we refer to simply as "protection") as two distinct objectives, and by evaluating priority reserve policies in terms of these objectives. In our formal treatment the objective of providing protection for the priority group is fulfilled by a preferential policy which is targeted at those members of the group who need it.

As a first step towards a more unified approach to priority reserve design, we study the normative and incentive properties of matching mechanisms with a reserve policy in a simple model with one group of agents who are to be prioritized over the rest of the agents, and analyze and characterize two contrasting reserve policies in order to understand the impact of their design. Our results identify Hafalir et al.'s (2013)[30] DA-MiR mechanism as the least targeted priority reserve mechanism, serving the primary goal of representation, as opposed to the DA-TPR mechanism which we propose as the most targeted priority reserve mechanism, serving the primary goal of providing effective preferential treatment.

These two matching mechanisms, the DA-MiR and the DA-TPR, are the two extreme members in the large family of DA-SPR mechanisms which includes any arbitrary sequence of reserved and unreserved positions for each entity when using the Deferred Acceptance mechanism with entity selection based on these sequences. We characterize this general class of mechanisms in terms of their desirable stability and incentive properties, which underlines the importance of these mechanisms. However, we argue that many members of this family of reserve policies are unclear in terms of their objectives and their impact on the welfare of the priority group. Therefore, we identify a subclass of the class of DA-SPR mechanisms, the Split DA-SPR mechanisms, which have a transparent design that explicitly specifies which reserved positions are targeted and which ones are not. The family of Split DA-SPR mechanisms enables the policy designer to combine the two goals of representation and protection in different proportions for each entity, since it consists of a whole range of selection policies which compromise between these two goals. These mechanisms are simple and are shown to be relatively consistent in their policy impact, and provide flexibility to choose a selection sequence for each entity which is transparent in its desired compromise.

Appendix A: Augmented Markets

We introduce an augmented one-to-one matching market for each DA-SPR mechanism, for which the standard agent-proposing DA mechanism produces the same matching as the DA-SPR mechanism.

Let the original market be $M = \langle N, N^+, E, q, q^r, P, \succ \rangle$. Fix a profile of entity selection sequences $s \equiv (s_e)_{e \in E}$ and let f^s denote the corresponding DA-SPR mechanism. We define an augmented market $M^{A(s)} = \langle N, E^A, P^{A(s)}, \succ^A \rangle$ for market M and DA-SPR rule f^s, as follows. Split each entity $e \in E$ into q_e entities with one position for each entity. Then the set of entities E^A in the augmented market consists of $\sum_{e \in E} q_e$ entities. The entity ranking $\succ_e^A$ of entity e in the augmented market is given as follows. For each entity $e \in E$, corresponding to each of the q_e^u unreserved positions of e, the entity e^u in the augmented market keeps the original ranking of entity e: $\succ_{e^u}^A \equiv \succ_e$. Corresponding to each of the q_e^r reserved positions of e, the entity e^r in the augmented market has priority-agent favoring rankings as defined below:

$$i \succ^A_{e^r} j \iff \begin{cases} i \in N^+ \text{ and } j \in N^0 \\ i, j \in N^+ \text{ and } i \succ_e j \\ i, j \in N^0 \text{ and } i \succ_e j \end{cases}$$

To specify the preferences $P_i^{A(s)}$ of agent $i \in N$ for the augmented market, for each agent i we replace entity e with its q_e copies of e in the preference ordering of agent i such that the ordering of the copies follows the sequence s_e of reserved and unreserved positions as specified by the DA-SPR mechanism f^s. Indexing the copies of reserved and unreserved positions available at each entity $e \in E$, the q_e^u unreserved positions become $e^{u_1}, e^{u_2}, \ldots$, and the q_e^r reserved positions become $e^{r_1}, e^{r_2}, \ldots$. Each agent keeps the relative ordering of the different entities $e \in E$ the same, and ranks the positions for each entity (which are ranked consecutively) following the sequence s_e specified by the DA-SPR mechanism f^s. For example, if under the DA-SPR mechanism f^s we have $s_{e_1} = (u, r, u)$ and $s_{e_2} = (r, r, u, u)$, and if $e_1 P_i e_2$ for agent $i \in N$, then $e_1^{u_1} \ P_i^{A(s)} \ e_1^{r_1} \ P_i^{A(s)} \ e_1^{u_2} \ P_i^{A(s)} \ e_2^{r_1} \ P_i^{A(s)} \ e_2^{r_2} \ P_i^{A(s)} \ e_2^{u_1} \ P_i^{A(s)} e_2^{u_2}$. This determines the one-to-one augmented market $M^{A(s)}$ for market M and DA-SPR rule f^s.

Any matching in $M^{A(s)}$ can be transformed into a matching in M in a straightforward manner: all agents who are matched to an entity in E^A in the augmented market $M^{A(s)}$ that corresponds to either an unreserved or a reserved position of e are matched to e in M. Given this definition of the augmented market $M^{A(s)}$, the following result is easy to verify .

Lemma A.1. *Fix a DA-SPR mechanism f^s. For each profile $(P, \succ) \in (\mathcal{P} \times \Pi)$, the DA-SPR matching $f^s(P, \succ)$ in the original market M corresponds to the agent-proposing DA matching $f^{DA}(P^{A(s)}, \succ^A)$ in the augmented market $M^{A(s)}$.*

Appendix B: Main Characterization Proofs

Proof of Theorem 11:

A DA-SPR mechanism is weakly reserve-stable and strategyproof by Theorems 8 and 10. We will show the converse: if a mechanism is weakly reserve-stable and strategyproof then it is a DA-SPR mechanism. Fix a mechanism g which is weakly reserve-stable and strategyproof. We first identify in Step 1 a mechanism within the family of DA-SPR mechanisms that is a candidate to be the same mechanism as g. Then in Steps 2 and 3 we use a proof by contradiction to prove that the two mechanisms are the same.

Step 1. In this step we construct the profile $s \equiv (s_e)_{e \in E}$ of entity selection sequences such that the corresponding DA-SPR mechanism f^s will be shown to be the same as the fixed weakly reserve-stable and strategyproof mechanism g, given the fixed entity ranking profile $\succ \in \Pi$. Fix $e \in E$, and let $P \in \mathcal{P}$ be the preference profile where each agent finds only entity e acceptable: for all $i \in N$, $P_i = (e, i)$. Let p_e denote the number of priority agents matched to e in addition to the q_e^r priority agents that are matched to e by g, given that $g(P)$ is weakly reserve-stable at $(P, \succ)$. That is, let $p_e \equiv |\mu_e \cap N^+| - q_e^r$. Given that $|N^+| > q_e$, $p_e \geq 0$ by condition (2a) in the definition of weak reserve-stability. Corresponding to each priority agent in $\succ_e$, let s_e have a reserved position, and corresponding to each non-priority agent in $\succ_e$, let s_e have an unreserved position, following the order of agents in $\succ_e$, except for the following two requirements.

1. Assign an unreserved position to the first p_e priority agents in the ordering instead of a reserved position.

2. Place all remaining reserved positions at the end of s_e to have exactly q_e^r reserved positions.

Note that the sequence s_e has a length of q_e. Repeat the same for each entity $e \in E$ in order to identify the sequence s_e for entity e. The constructed profile s of entity selection sequences identifies a specific DA-SPR mechanism. Let this mechanism be denoted by f^s.

In the rest of the proof we show that, given the fixed $\succ \in \Pi$, for all $P \in \mathcal{P}$, $g(P) = f^s(P)$. Since g is an arbitrary weakly reserve-stable and strategyproof mechanism, this will imply that any weakly reserve-stable and strategyproof mechanism is a DA-SPR mechanism.

Step 2. Suppose, by contradiction, that there exists $P \in \mathcal{P}$ such that $f^s(P) \neq g(P)$. We will show that then there exist $i, j \in N$ and $\bar{P} \in \mathcal{P}$ such that $i \in \bar{\mu}_e$, $e\bar{P}_j\,\bar{\mu}(j)$, $j \in \bar{\nu}_e$, and $e\bar{P}_i\bar{\nu}(i)$, where $\bar{\mu} \equiv f^s(\bar{P})$ and $\bar{\nu} \equiv g(\bar{P})$. We will refer to this as a *reversal* in the rest of this step. Specifically, this is a reversal with agents i, j and entity e at preference profile $\bar{P}$.

For ease of notation, let $\mu \equiv f^s(P)$ and $\nu \equiv g(P)$. Thus, we suppose that $\mu \neq \nu$. Since ν is weakly reserve-stable at P and f^s is a DA-SPR rule, Theorem 9 implies that ν does not Pareto-dominate μ. Hence, there exists $i_1 \in N$ such that $\mu(i_1)P_{i_1}\nu(i_1)$. Then, since g is weakly reserve-stable, condition 2 in the definition of weak reserve-stability implies that $|\nu'_{\mu(i_1}| = q_{\mu(i_1)}$. This implies that there exists $i_2 \in N\setminus\{i_1\}$ such that $\nu(i_2) = \mu(i_1) \neq \mu(i_2)$. If $\mu(i_1)P_{i_2}\mu(i_2)$ then we have a reversal with agents i_1, i_2 and entity $\mu(i_1)$ at profile P. Therefore, assume that $\mu(i_2)P_{i_2}\mu(i_1)$. Then we can repeat the same argument to find $i_3 \in N \setminus \{i_2\}$ such that $\nu(i_3) = \mu(i_2) \neq \mu(i_3)$. Assuming that there is no reversal with any agents i_t, i_{t+1} and entity $\mu(i_t)$ at preference profile P and given that there are a finite number of agents, we can continue the same way to find t, t' with $1 \leq t' < t \leq n$ such that $\nu(i_t) = \mu(i_{t-1}) \neq \mu(i_t)$, $i_t = i_{t'}$ and thus $\nu(i_{t'}) = \mu(i_{t-1}) \neq \mu(i_{t'})$. Let $I \equiv \cup_{v=t'}^{t-1} i_v$. Note that for all $v = t', \ldots, t-1$, $\mu(i_v)P_{i_v}\nu(i_v)$ and $\cup_{v=t'}^{t-1}\mu(i_v) = \cup_{v=t'}^{t-1}\nu(i_v)$.

For all $l \in N$, let $\tilde{P}_l$ denote the truncation of P_l at $\mu(l)$, that is, the preference ordering of the entities ranked above $\mu(l)$ is unchanged, with $\mu(l)$ as the last-ranked acceptable entity in $\tilde{P}_l$, while all entities ranked below $\mu(l)$ in P_l are unacceptable in $\tilde{P}_l$. If $g(\tilde{P}_I, P_{-I}) = \mu$ then g is not weakly group-strategyproof since coalition I can manipulate the matching at P, which is a contradiction. Therefore, $g(\tilde{P}_I, P_{-I}) \neq \mu$. However, since f^s is a DA-SPR rule and $\tilde{P}_l$ is a truncation of P_l at $\mu(l)$, where $\mu = f^s(P)$, we have $f^s(\tilde{P}_I, P_{-I}) = \mu$. Therefore, given that $f^s(\tilde{P}_I, P_{-I}) \neq g(\tilde{P}_I, P_{-I})$, we can apply a similar argument to $(\tilde{P}_I, P_{-I})$ as the one used for P above, and assuming that there is no reversal, we can show after iteratively applying this argument, that $f^s(\tilde{P}) = \mu \neq g(\tilde{P})$. Note that $g(\tilde{P})$ does not Pareto-dominate $f^s(\tilde{P})$, since f^s is constrained efficient with respect to weak reserve-stability by Theorem 9 and $g(\tilde{P})$ is weakly reserve-stable at $(\tilde{P}, \succ)$. Therefore, there exists $h \in N$ such that $\mu(h) \neq h$ and $\nu(h) = h$. Then, since g is weakly reserve-stable, condition 2 in the definition of weak reserve-stability implies that $|g_{\mu(h)}(\tilde{P})| = q_{\mu(h)}$. This implies that there exists $h' \in N \setminus \{h\}$ such that $g_{h'}(\tilde{P}) = \mu(h) \neq \mu(h')$. Since $\mu(h)\tilde{P}_{h'}\mu(h')$, this is a reversal with agents h, h' and entity $\mu(h)$ at profile $\tilde{P}$.

Step 3.

As shown in Step 2, there exists a reversal with agents $i, j \in N$ and entity e at profile $\bar{P} \in \mathcal{P}$ such that $i \in \bar{\mu}_e$, $e\bar{P}_j\,\bar{\mu}(j)$, $j \in \bar{\nu}_e$, and $e\bar{P}_i\bar{\nu}(i)$, where $\bar{\mu} \equiv f^s(\bar{P}, \succ)$ and $\bar{\nu} \equiv g(\bar{P}, \succ)$. We now consider different cases based on whether i and j are priority agents, and derive a contradiction in each case.

Case 1: $i, j \in N^+$.

Given that $\bar{\nu}$ is weakly reserve-stable at $(\bar{P}, \succ)$ and $i \in N^+$, condition (2a) in the definition of weak reserve-stability implies that for all $l \in \bar{\nu}_e$, $l \succ_e i$, and thus $j \succ_e i$. Similarly, given that $\bar{\mu}$ is weakly reserve-stable at $(\bar{P}, \succ)$, $j \in N^+$ implies that for all $l \in \bar{\mu}_e$, $l \succ_e j$, and thus $i \succ_e j$. This is a contradiction.

Case 2: $i, j \in N^0$.

Given that $\bar{\nu}$ is weakly reserve-stable at $(\bar{P}, \succ)$ and $i \in N^0$, condition (2b) in the definition of weak reserve-stability implies that for all $l \in (\bar{\nu}_e \cap N^0)$, $l \succ_e i$, and thus $j \succ_e i$. Similarly, given that $\bar{\mu}$ is weakly reserve-stable at $(\bar{P}, \succ)$ and $j \in N^0$, for all $l \in (\bar{\mu}_e \cap N^0)$, $l \succ_e j$, and thus $i \succ_e j$. This is a contradiction.

Case 3: $i \in N^0$, $j \in N^+$.

Since $\bar{\mu}$ is weakly reserve-stable, condition (2a) in the definition of weak reserve-stability implies $i \succ_e j$, and since $\bar{\nu}$ is also weakly reserve-stable, condition (2b) implies that $|\{h \in (\bar{\nu}_e \cap N^+) : i \succ_e h\}| \leq q^r_e$. Thus, at most q^r_e priority agents who are assigned to $\bar{\nu}_e$ are lower-ranked than i by $\succ_e$. However, since $\bar{\mu}$ is weakly reserve-stable, by condition (2a) we also have $|\bar{\mu}_e \cap N^+| \geq q^r_e$, that is, there are at least q^r_e higher-ranked priority agents than j according to $\succ_e$, since $j \notin \bar{\mu}_e$ and $eP_j\bar{\mu}_e$. Thus, since $i \in N^0$ gets an unreserved position at $\bar{\mu}_e$ in the sequence s_e identified for entity e in Step 1, i's ranking among N^0 agents by $\bar{\succ}_e$ is no higher than $q_e - q_e^r$. However, since $i \notin \bar{\nu}_e$, $j \in \bar{\nu}_e$, and $e\bar{P}_i\bar{\nu}(i)$, by condition (2b) in the definition of weak reserve-stability, for all $l \in (\bar{\nu}_e \cap N^0)$, $l \bar{\succ}_e i$, and no more than q_e^r priority agents are matched to e, which is a contradiction.

Case 4: $i \in N^+$, $j \in N^0$.

Since $\bar{\nu}$ is weakly reserve-stable, condition (2a) in the definition of weak reserve-stability implies that $j \succ_e i$ and $|(\bar{\nu}_e \cap N^+)| \geq q_e^r$. Therefore, at least q_e^r priority agents are ranked higher than i by $\succ_e$. Since $\bar{\mu}$ is weakly reserve-stable, condition (2b) in the definition of weak reserve-stability implies that for all $l \in (\bar{\mu}_e \cap N^0)$, $l \succ_e j$, and no more than q_e^r priority agents are matched to e in $\bar{\mu}_e$ who are ranked lower by $\succ_e$ than j, which implies that in the sequence s_e identified for entity e in Step 1 agent i gets a reserved position at $\bar{\mu}_e$, and i is ranked among the top q_e^r priority agents by $\succ_e$ among those priority agents who are ranked lower than j, which is a contradiction.

As we have derived a contradiction in all possible cases, $g(P) = f^s(P)$ for all $P \in \mathcal{P}$, and the proof is completed. □

We are now ready to present the proofs of the characterizations of the DA-TPR and DA-MiR mechanisms, which build on Theorem 11.

Proof of Theorem 6:

Protection-stability is shown by Theorem 5, and the strategyproofness of the DA-TPR mechanism follows from the more general result of Theorem 10. The other direction of the statement can be established as follows. Given that protection-stability implies weak reserve-stability, by Theorem 11 the only mechanisms that may satisfy protection-stability and strategyproofness are DA-SPR mechanisms. We will show that the only DA-SPR mechanism which satisfies protection-stability is the DA-TPR mechanism.

Suppose, by contradiction, that f^s is a DA-SPR mechanism for which there exists entity $e \in E$ with a selection sequence s_e such that it has at least one reserved position preceding an unreserved position. Let the first such reserved position in s_e be the kth position in the sequence s_e, where $k > 0$. Let $(P, \succ)$ satisfy the following. For all $i \in N$, $P_i \in (e)$, and $\succ_e$ ranks $k-1$ non-priority agents first, followed by one priority agent, say $i \in N^+$, then the rest of the non-priority agents and finally the rest of the priority agents. Then e selects the top q_e^r priority agents in step 1 of the DA-SPR procedure f^s at $(P, \succ)$, including priority agent i, and given $(P, \succ)$, this set of agents is $f_e^s(P)$. Note that a non-priority agent, say $j \in N^0$, is selected for the $k+1$st position by e, given our assumption that there are more than q_e^u non-priority agents, and thus $i \succ_e j$ and $i \notin \bar{N}_e^+(f_e^s(P), \succ_e)$. This implies that $|\bar{N}_e^+(f_e^s(P), \succ_e)| = q_e^r - 1$, which violates (2a) in the definition of protection-stability. Therefore, $f^s(P)$ is not protection-stable at $(P, \succ)$. This means that a DA-SPR mechanism can only satisfy protection-stability if for each entity e the selection sequence has no reserved position preceding any unreserved position. There is only one such DA-SPR mechanism, namely the DA-TPR mechanism, and the proof is completed. □

Proof of Theorem 7:

Representation-stability and strategyproofness of the DA-MiR mechanism are shown by Hafalir et al. (2013). The other direction of the statement can be established as follows. Given that representation-stability implies weak reserve-stability, by Theorem 11 the only mechanisms that may satisfy representation-stability and strategyproofness are DA-SPR mechanisms. We will show that the only DA-SPR mechanism which satisfies representation-stability is the DA-MiR mechanism.

Let f^s be a DA-SPR mechanism for which there exists entity $e \in E$ with a selection sequence s_e such that it has at least one unreserved position preceding a reserved position. Let the first such unreserved position in s_e be the kth position in the sequence s_e, where $k > 0$. Let $(P, \succ)$ satisfy the following. For all $i \in N$, $P_i \in (e)$, and $\succ_e$ ranks $k-1$ priority agents first, followed by all non-priority agents, and finally the rest of the priority agents. Then e selects the top k priority agents for the first k positions at $(P, \succ)$ in step 1 of the DA-SPR procedure f^s. Moreover, e selects an additional $q_e^r - (k-1)$ priority agents for the remaining reserved positions, all of whom are ranked lower by e than all of the non-priority agents. Given our assumptions, there is at least one such additional priority agent, say $j \in N^+$, and there is at least one non-priority agent, say $i \in N^0$ who is not selected by e. Given $(P, \succ)$, this set of agents is $f_e^s(P)$. Then $|\{f_e^s(P) \cap N^+\}| > q_e^r$ and there exist $i \in N^0$ and $j \in f_e^s(P)$ such that $i \succ_e j$, which violates (2b) in the definition of representation-stability. Therefore, $f^s(P)$ is not representation-stable at $(P, \succ)$. This means that a DA-SPR mechanism can only satisfy representation-stability if for each entity e the selection sequence has no unreserved position preceding any reserved position. There is only one such DA-SPR mechanism, namely the DA-MiR mechanism, and the proof is completed. □

Chapter 4

Partially Targeted Priority Reserve Policies

4.1 Introduction

The prevalent minority reserve policies in school choice limit the number of admitted majority students to give minority students a better chance to attend their desired schools. There are several studies on affirmative action policies in the context of matching mechanisms, such as Deferred Acceptance mechanisms with specific choice rules that accommodate priority reserves. Hafalir et al. (2013)[30] recommend an affirmative action policy with minority reserves, the DA-MiR (Deferred Acceptance with Minority Reserves) mechanism, which is a Deferred Acceptance algorithm with choice rules that have a quota set aside at each university to prioritize minority students. The DA-MiR mechanism gives higher priority to minority students up to the reserved number of positions at each university, however, it does not allow for admitting more minority students to a university than the number of reserved slots, unless all the minority students could gain entrance even without having an affirmative action policy. The DA-MiR mechanism serves the goal of minority representation well, given its attractive efficiency properties compared to majority quota rules (Kojima (2013)[39]). However, if the primary goal is not representation but rather to help prioritize a group of typically disadvantaged agents, such as refugee families in emergency situations, who may or may not be matched otherwise to entities that are highly ranked by them, then the DA-MiR mechanism is less attractive. Therefore, in the previous chapter we propose an alternative matching mechanism with *targeted priority reserves*, the DA-TPR (Deferred Acceptance with Targeted Priority Reserves) mechanism, which captures the goal of targeting exactly those agents in the priority group who are in need of preferential treatment, while recognizing that not all agents in the priority group need such treatment and thus allowing for a higher number of priority agents to be matched to an entity than the quota set aside for them.

A targeted reserve policy may be applicable to any situation where a disadvantaged group needs preferential treatment but some members of the group do not require this treatment, provided that the main objective is not representation, such as affirmative action, but to provide prioritization or protection for those who are deemed to have priority and who need this preferential treatment. Thus, the targeted priority reserve policy of the DA-TPR mechanism does not aim for representation, and the number of set-aside positions is determined by how many priority agents are to be helped, not by how many should be matched to a particular entity. Providing effective preferential treatment that makes a difference is especially important for refugee settlement ([51] [52]).

The DA-TPR mechanism that we propose in Chapter 3 fills the unreserved positions of the entities first, based on a general competition of all agents, which provides an opportunity for priority group agents to get accepted by entities on their own right, without taking up reserved positions. Once the unreserved positions are filled, when allocating the reserved positions entities target priority group agents who are in need of prioritization. The DA-TPR mechanism is a good alternative priority policy to the DA-MiR mechanism not only because it is more appropriate to meet the objective of providing effective preferential treatment, but also because it retains the positive features of the DA-MiR mechanism, as shown in Chapter 3. We also study the entire class of DA-SPR

(Deferred Acceptance with Sequential Priority Reserves) mechanisms in Chapter 3, which allow for any arbitrary sequence of the reserved and unreserved positions for each entity.

The impact of the sequence of filling reserved versus unreserved positions has been studied previously in the school choice context by Dur et al. (2018)[19]) and Dur et al. (2020)[21]. These papers study specific applications of providing priority treatment in the Boston and Chicago school choice programs, and their focus is school-level selection rather than the entire matching mechanism. Unlike our mechanism-level analysis, [19] and [21] study the properties of a school's choice rule in various models.[1] By contrast, we study entire matching mechanisms in the same simple model as Hafalir et al. (2013)[30], with one priority group only. Moreover, [19] and [21] view the specific selection sequencing as merely a means of affecting the number of accepted prioritized students, demonstrating that the representation of prioritized students can be changed by modifying only the sequencing but not the number of reserved positions. We argue, on the other hand, as seen in the previous chapter, that different precedence orders correspond to different policy goals.

In the class of DA-SPR mechanisms, the DA-MiR mechanism can be seen as the least targeted priority policy that serves the goal of *representation* and, at the other extreme, the DA-TPR mechanism that we propose in Chapter 3 emerges as the most targeted priority policy that is aimed at providing prioritization for *protection.* In the previous chapter we also propose the Split DA-SPR rules as good compromises between the two extreme rules for the designer who is interested in having both some representation and some effective protective treatment in the priority policy. Split DA-SPR rules split the reserved positions between the front (as in the DA-MiR mechanism) and the end of the entity selection sequence (as in the DA-TPR mechanism), with all the unreserved positions between them. These are intuitively appealing and transparent rules for policy makers, and we also demonstrated in Chapter 3 that these rules exhibit a consistent

[1] Other related studies of affirmative action mechanisms include Ehlers et al. (2014)[23] and Kominers and Sönmez (2016)[40], while Echenique and Yenmez (2015)[22] focus on the choice rules of a given school, similarly to [19] and [21].

policy impact when compared to each other.

The current chapter further explores priority treatment policies with reserves which compromise between representation and protection. The Split DA-SPR mechanisms were chosen from the class of DA-SPR rules as intuitive mechanisms for such a compromise. In this study we want to broaden this inquiry by not limiting our choices to DA-SPR mechanisms exclusively. We investigate potential alternatives for situations where the designer doesn't wish to use either a fully targeted mechanism (the DA-TPR) or a fully non-targeted mechanism (the DA-MiR), and we don't simply want to recommend a policy from the class of DA-SPR mechanisms. Instead, we want to specify directly how we compromise when we use Deferred Acceptance mechanisms with selection rules (i.e., choice rules) for each entity that follow the desired compromise, given the generally attractive properties of Deferred Acceptance mechanisms with choice rules. We start our investigation by making explicit inquiries regarding the design of the priority treatment policy in search of appropriate compromises, with the intention to pose questions that any policy maker could answer without being aware of the formal studies of priority reserves in matching theory. Specifically, we ask: *how many reserved positions have to be taken up by qualified priority agents?* Here *qualified priority agents* are agents who would be matched to the entity regardless of their priority status. If the answer is "all of them," then we have the DA-MiR mechanism, and if the answer is "none of them" then we have the DA-TPR mechanism. Slightly changing this question, we may also ask: *at most how many unreserved positions may be taken up by qualified priority group agents*? If the answer is "all of them," then we have the DA-TPR mechanism, and if the answer is "none of them" then we have the DA-MiR mechanism. Between these two extreme cases there is a range of policies with different degrees of compromise for both questions. Interestingly, even though these two policy inquiries are similar, they lead to two distinct sets of policies.

In the first class of mechanisms that we obtain by focusing on the allocation of reserved positions, each entity selection rule is based on the number of qualified priority agents who must get reserved positions, which applies to the applicant pool of each entity in each round of the Deferred Acceptance algorithm. Since these mechanisms are based on the allocation of the reserved positions, we call them the *r-PTPR mechanisms*. Here PTPR stands for Partially Targeted Priority Reserves mechanisms, and r-PTPR is short for Reserved-PTPR mechanisms. We show that the r-PTPR mechanisms are equivalent to a subclass of the class of DA-SPR mechanisms (Theorem 15), what we call the class of DA-ESPR mechanisms (Deferred Acceptance mechanisms with Explicit Sequential Priority Reserves). Interestingly, the class of DA-ESPR mechanisms is not the same as the Split DA-SPR mechanisms from Chapter 3. In fact, the only too common members of these two classes of mechanisms are the DA-TPR and DA-MiR mechanisms, the two extreme members of both classes. The family of DA-ESPR mechanisms is also intuitive, similarly to Split DA-SPR mechanisms, but they have a more subtle structure, and it is not obvious at first glance that these rules would correspond to our policy goals focusing on the allocation of reserved seats between qualified and unqualified priority group agents. The r-PTPR mechanisms also have some nice properties, in addition to their intuitive features, as seen from the equivalence with DA-ESPR mechanisms. We know from the previous chapter that the large family of DA-SPR mechanisms is characterized by a a weak stability axiom with priority reserves, called weak reserve-stability, and strategyproofness (see Theorem 11 in Chapter 3). Therefore, since the class of r-PTPR mechanisms is equivalent to the class of DA-ESPR mechanisms, the r-PTPR mechanisms are also weakly reserve-stable, constrained efficient with respect to weak reserve-stability, and weakly group-strategyproof, by Theorems 8, 9, and 10 in Chapter 3.

In the second class of mechanisms that we obtain by focusing on the allocation of unreserved positions, each entity selection rule is based on the maximum number of qualified priority group agents who are permitted to get unreserved positions, which applies to the applicant pool of each entity in each round of the Deferred Acceptance algorithm. Since these mechanisms are based on the allocation of the unreserved positions, we call them the *u-PTPR mechanisms* (short for Unreserved-PTPR mechanisms). We show that this class of mechanisms is not a subclass of the class of DA-SPR mechanisms (Theorem 16). However, they do satisfy weak reserve-stability (Theorem 17). Since these are not DA-SPR rules as shown by Theorem 16, Theorem 17 implies that these rules are not strategyproof, given our characterization result of DA-SPR mechanisms in the previous chapter (Theorem 11 in Chapter 3).

The policy goals behind the r-PTPR and u-PTPR mechanisms are easy to understand for the policy maker due to their consistent partially targeted priority policies specified by their policy numbers, as opposed to an arbitrary policy with inconsistent entity selection rules, which would be difficult to justify for practical use. The classes of r-PTPR and u-PTPR mechanisms allow for transparent choices when selecting a corresponding partially targeted priority reserve policy that a specific allocation problem calls for. In particular, these mechanisms enable the policy designer, in addition to setting the number of reserved positions at each entity, to also determine the extent to which the priority policy should be targeted. The flexible policies that these two classes of mechanisms offer have implications for refugee settlement, immigration, centralized university admissions, and public school choice systems, among other settings.

4.2 The Model

A matching market with priority reserves is defined by the following components.

1. A finite set of **agents** N.
2. A **priority group** $N^+ \subset N$. We call $N^0 = N \setminus N^+$ the **non-priority group.** Thus, $N = N^+ \cup N^0$. We will refer to agents in the priority group as *priority agents* for short, and to agents in the non-priority group as *non-priority agents.*
3. A finite set of **entities** E.
4. Each entity $e \in E$ has a **capacity** $q_e > 0$, which is the number of available positions at entity e. Let $q \equiv (q_e)_{e \in E}$.

5. **Reserved positions.** For each entity $e \in E$, $q_e = q_e^r + q_e^u$, where q_e^r denotes the number of reserved positions for the priority group N^+, and q_e^u is the number of unreserved positions at entity e. To simplify our exposition, we assume that $q_e^r > 0$ for all $e \in E$. Let $q^r \equiv (q_e^r)_{e \in E}$ and $q^u \equiv (q_e^u)_{e \in E}$.

6. Each $i \in N$ has **strict preferences** P_i over $E \cup \{i\}$. Agents may not find all entities acceptable, and ranking an entity below i indicates that the entity is not acceptable.

7. Each entity $e \in E$ has a **strict** (priority) **ranking** $\succ_e$ over N. Entities are not considered active agents with their exogeneously given priorities over agents, similarly to schools in the school choice model of Abdulkadiroğlu and Sönmez (2003). Thus, when considering efficiency and incentives, we only consider the agents' side. We will refer to the entity priorities from now on as **entity rankings** in order to avoid any confusion in the terminology between entity priorities and the fixed set of priority agents N^+.

Let $P \equiv (P_i)_{i \in N}$ denote a **preference profile** for the agents and let $\mathcal{P}$ be the set of preference profiles. Let $\succ \equiv (\succ_e)_{e \in E}$ denote an **entity ranking profile** and let Π be the set of entity ranking profiles. Then a market is defined by $\langle N, N^+, E, q, q^r, P, \succ \rangle$. Given fixed N, N^+, E, q and q^r, a market is determined by a **profile** $(P, \succ)$ consisting of a preference profile $P \in \mathcal{P}$ and an entity ranking profile $\succ \in \Pi$.

We will also need the following notation concerning the preferences. Let R_i denote the weak preferences of agent $i \in N$ associated with P_i. Since preferences are assumed to be strict, eR_ie' means that either eP_ie' or $e = e'$. We will write $P_i \in (e)$ if agent i ranks entity e first, $P_i \in (e, e')$ if agent i ranks entity e first and entity e' second, and so on. Then, for example, $P_i \in (e, i)$ indicates that agent i finds only entity e acceptable. The preferences of a coalition $L \subseteq N$ in profile P are denoted by P_L. We denote the preference profile of all the agents except for i by P_{-i}, and the preference profile of all the agents except the agents in coalition L by P_{-L}.

In order to simplify the exposition, we assume that for all $e \in E$, $|N^+| > q_e$ and $|N^0| > q_e$. These are mild assumptions which are easily satisfied in most applications and allow us to ignore some unimportant special cases in which there are too few agents.

Matching. An allocation in a market is a matching μ, which is a function from N to $E \cup N$ such that, for every agent $i \in N$, $\mu(i) \in E \cup \{i\}$, where $\mu(i) = e$ means that agent i is matched to entity e and $\mu(i) = i$ means that agent i is unmatched. We will refer to $\mu(i)$ as agent i's *assignment* in μ. With a slight abuse of notation we denote that set of agents matched to i in matching μ by μ_e, and hence require that $\mu_e \subseteq N$ such that $|\mu_e| \leq q_e$. Let $\mathcal{M}$ denote the set of matchings.

A matching μ *Pareto-dominates* matching ν at preference profile P if for all agents $i \in N$, $\mu_i R_i \nu_i$, and there exists agent $j \in N$ such that $\mu_j P_j \nu_j$.

Matching mechanism. A mechanism is a function $f : (\mathcal{P} \times \Pi) \to \mathcal{M}$ that assigns a matching to each profile $(P, \succ) \in (\mathcal{P} \times \Pi)$.

An **applicant pool** $A \subseteq N$ is a subset of N. Let A^+ denote the set of priority agents, and let A^0 denote the set of non-priority agents, in applicant pool A. That is, let $A^+ \equiv A \cap N^+$ and $A^0 \equiv A \cap N^0$.

Next we define the notion of qualified priority agents, since identifying these agents in an applicant pool plays an important role in the classes of mechanisms that we study.

Qualified priority agents. Let $e \in E$. Given a a fixed applicant pool $A \subseteq N$ with $|A| \geq q_e^u$ and a fixed entity ranking $\succ_e$, let $\text{top}^u(A, \succ_e) \subseteq A$ be the set of agents who are ranked by $\succ_e$ as the top q_e^u agents in A. Moreover, let $\text{top}^{u+}(A, \succ_e) = \text{top}^u(A, \succ_e) \cap A^+$ be the set of priority agents among the agents who are ranked by $\succ_e$ as the top q_e^u agents in A. We will refer to $\text{top}^{u+}(A, \succ_e)$ as the set of *qualified priority agents* in A according to $\succ_e$, since these are priority agents who would always be selected by an entity, even if they were not priority agents. Finally, let $v_e^+(A, \succ_e) \equiv |\ \text{top}^{u+}(A, \succ_e)|$ denote the number of qualified priority agents in A according to $\succ_e$. For ease of notation, we suppress the domain and denote the number of qualified priority agents by v_e^+ whenever this is unambiguous.

4.3 Deferred Acceptance Mechanisms

All matching mechanisms studied in this paper are agent-proposing Deferred Acceptance mechanisms with different entity selection rules (also known as choice rules). The selection rule for each entity specifies the set of selected agents from each applicant pool $A \subseteq N$ which arises in some step of the iterated agent-proposing Deferred Acceptance algorithm. We denote the applicant pool by $A_e^k \subseteq N$ in step k of the algorithm at a specific profile, where the applicant pool in step k for $k > 1$ is the union of the set of new applicants to entity e in step k and the set of agents tentatively matched to e in step $k-1$, and for $k = 1$ we let A_e^1 be simply the set of new applicants to e in step 1.

Entity selection rules. A selection rule for entity e, denoted by c_e, maps from any applicant pool $A \subseteq N$ and entity ranking $\succ_e$ to a subset of A as follows:

i. if $|A| \leq q_e$ then $c_e(A, \succ_e) = A$.

ii. if $|A| > q_e$ then $c_e(A, \succ_e) \subset A$ such that $|c_e(A, \succ_e)| = q_e$.

We now define the Gale-Shapley agent-proposing Deferred Acceptance mechanisms with entity selection rules, which only differ from the standard Deferred Acceptance algorithm [27], which simply follows the entity rankings for selection, in that they allow for arbitrary entity selection rules which determine the selection from the applicant pool of each entity in each step of of the algorithm.

Deferred Acceptance (DA) mechanisms with entity selection rules

Fix a profile $(P, \succ) \in (\mathcal{P} \times \Pi)$.

Step 1: Each agent applies to her most preferred entity according to P. Each entity $e \in E$ tentatively assigns its positions among its applicants (i.e., from applicant pool A_e^1)

according to its entity selection rule, up to its capacity. Any remaining applicants are rejected.

Step *k*: Each agent who was rejected in step $k-1$ applies to her next most preferred acceptable entity according to P among the entities that have not yet rejected her, if any such entity remains. Each entity selects from the set of agents it has been tentatively matched to, together with the set of its new applicants in this step (i.e., from applicant pool A_e^k) and tentatively assigns its positions according to its entity selection rule, up to its capacity. Any remaining agents in this set are rejected.

The algorithm terminates when no applicant is rejected in a particular step, and the tentative assignments in this last step become final.

When there is a priority reserve policy in place, the entity selection rule for each entity depends not only on the overall capacity of the entity q_e but also on its reserve quota q_e^r, and different entity selection rules may differ in their priority reserve policies. Specifically, recall that if the q_e^r reserved positions are filled first from the priority group applicants A^+ and then the q_e^u unreserved positions are filled from the remaining agents in the applicant pool A by each entity e then we get the DA-MiR mechanism of Hafalir et al. [30], and if the q_e^u unreserved positions are filled first from the general applicant pool A and then the q_e^r reserved positions are filled from the remaining priority agents in A^+ by each entity e then we get the DA-TPR mechanism studied in Chapter 3.[2]

4.4 Reserved-PTPR Mechanisms

In this section we study the class of r-PTPR mechanisms. These are Deferred Acceptance mechanisms with entity selection rules such that, for each member of the class of r-PTPR mechanisms, the entity selection rule for each entity specifies the number of qualified priority group agents who are required to take reserved positions. If there are not enough priority group agents in an applicant pool then this requirement may not be binding.

[2]See Chapter 3 for detailed definitions and explanations of the DA-MiR and DA-TPR mechanisms.

Specifically, if there are no more priority agents in the applicant pool than the number of reserved positions, then this requirement is satisfied automatically. Similarly, if there are no qualified priority agents then this requirement has no impact. We now formally define the r-PTPR mechanisms.

r-PTPR mechanisms

r-PTPR mechanisms are parameterized by the policy profile $(p_e^r)_{e\in E}$, which specifies a policy number p_e^r for each entity e, where $0 \leq p_e^r \leq q_e^r$ for all $e \in E$. This is the number of qualified priority agents in the applicant pool who must take up reserved positions.

Fix a policy profile $p^r \equiv (p_e^r)_{e\in E}$. The r-PTPR mechanism specified by p^r is defined by the following entity selection rule. For each entity $e \in E$ with entity ranking $\succ_e$, and for each applicant pool $A \subseteq N$:

i. If $|A| \leq q_e$, select A.

ii. If $|A| > q_e$:

ii.a. Select $\max(v_e^+(A, \succ_e) - p_e^r, 0) + q_e^r$ priority agents according to entity ranking $\succ_e$ (i.e., select the highest-ranked priority agents according to $\succ_e$). If $|A^+| \leq \max(v_e^+(A, \succ_e) - p_e^r, 0) + q_e^r$, then select all priority agents.

ii.b. Fill the remaining positions from the remaining agents in the applicant pool according to $\succ_e$ such that a total of q_e agents are selected.

Examining the lower and upper limits of the policy number p_e^r, we can verify that the extreme members of this class of mechanisms are the DA-MiR and DA-TPR mechanisms. If the policy number is set at zero ($p_e^r = 0$), meaning that none of the qualified priority agents have to take up reserved positions, then we get the DA-TPR entity selection rule, since $p_e^r = 0$ implies that $v_e^+(A, \succ_e) + q_e^r$ priority agents are selected from the applicant pool, which corresponds to the the number of qualified priority agents plus the number of reserved positions, implying that none of the qualified priority agents take up reserved positions. Thus, the policy profile $p_e^r = 0$ for all $e \in E$ leads to the DA-TPR mechanism.

If the policy number is set to equal the number of reserved positions ($p_e^r = q_e^r$), meaning that all of the reserved positions have to be taken by qualified priority agents (when there are enough such agents), then we get the DA-MiR entity selection rule, since if $p_e^r = q_e^r$ then q_e^r priority agents are selected first when $v_e^+(A, \succ_e) \leq p_e^r$. When $v_e^+(A, \succ_e) > p_e^r$, the entity selection rule requires that $v_e^+(A, \succ_e)$ priority agents are selected first, but notice that in this case the DA-MiR selection rule would select all qualified priority agents overall, since the ones not selected for the reserved positions up front would be selected for the unreserved positions afterwards. In fact, in this case potentially even more priority agents than the number of qualified priority agents may be selected in total, since this case corresponds to the rare scenario where there are so many qualified priority agents that the DA-MiR selection rule chooses more priority agents than the number of reserved positions, and it selects exactly the q_e highest-ranked agents regardless of their priority status. Thus, the policy profile $p_e^r = q_e^r$ for all $e \in E$ leads to the DA-MiR mechanism.

4.4.1 An Equivalence

We will establish next that r-PTPR mechanisms are DA-SPR mechanisms, and we identify the subclass of DA-SPR mechanisms that is equivalent to the class of r-PTPR mechanisms. We call mechanisms in this subclass DA-ESPR mechanisms (Deferred Acceptance mechanisms with Explicit Sequential Priority Reserves).

Before stating the equivalence result for r-PTPR mechanisms, we need to define the class of DA-ESPR mechanisms. Let us first recall the definition of the general class of DA-SPR mechanisms. DA-SPR mechanisms are Deferred Acceptance mechanisms with entity selection rules, where the entity selection rules are specified as entity selection sequences of reserved and unreserved positions. For an entity $e \in E$, an entity selection sequence consists of an arbitrary ordering of q_e^r reserved positions and q_e^u unreserved positions. Given an applicant pool $A \subseteq N$ and an entity ranking $\succ_e$, each reserved position is filled with the highest-ranked remaining priority agent, and each unreserved position is filled with the highest-ranked remaining agent from the applicant pool A. These selections follow the order of the entity selection sequence of reserved and unreserved positions, and the highest-ranked agents are determined according to entity ranking $\succ_e$. We denote an entity selection sequence for entity e by s_e, and write entity selection sequences as, for example, $s_e = (u, u, r, r, u, r, u, u, u, u)$, where r indicates a reserved position and u indicates an unreserved position, and entity e has $q_e^r = 3$ reserved positions, $q_e^u = 7$ unreserved positions, and hence $q_e = 10$ overall capacity.

Deferred Acceptance with Explicit Sequential Priority Reserves (DA-ESPR) mechanisms

A DA-ESPR entity selection sequence s_e for entity $e \in E$ is of the following form:

$$s_e = (\underbrace{r, \ldots, r,}_{q_e^{rf}} \underbrace{u, \ldots, u,}_{q_e^u - q_e^{rf}} \mid \underbrace{r, \ldots, r,}_{q_e^{rb}} \underbrace{u, \ldots, u}_{q_e^r - q_e^{rb}})$$

where there are q_e^{rf} reserved positions at the front and q_e^{rb} reserved position starting with the $(q_e^u + 1)$st position, such that $q_e^{rf} + q_e^{rb} = q_e^r$, and each unreserved position is either between the two sequences of reserved positions or at the end of the sequence.

We think of these selection sequences as "explicit," since it is transparent where the reserved positions are located. Notice that the DA-ESPR selection policies allow the designer to explicitly identify the non-targeted reserved positions which are at the front, and the rest of them are targeted reserved positions. The targeted reserved positions for DA-ESPR entity selection rules are different from the targeted reserved positions for Split DA-SPR entity selection rules, but nonetheless we call these targeted as well, although they are not at the very end of the sequence. DA-ESPR entity selection sequences have an intriguing structure which is related to both the DA-TPR and DA-Mir sequences in a subtle way. The overall structure of a DA-ESPR entity selection sequence resembles the DA-TPR entity selection sequence, since the first portion in the sequence consisting of reserved and then unreserved positions contains q_e^u positions, and the second portion in the sequence consisting of reserved and then unreserved positions contains q_e^r positions.

However, within each portion the structure of a DA-ESPR entity selection sequence corresponds to DA-MiR entity selection sequences, leading with the reserved positions and followed by the unreserved positions. The DA-TPR and DA-MiR entity selection sequences themselves are also the two extreme cases of the DA-ESPR entity selection sequences: when $q_e^{rf} = q_e^r$ and thus $q_e^{rb} = 0$, this is the DA-MiR selection sequence, and when $q_e^{rf} = 0$ and thus $q_e^{rb} = q_e^r$, this is the DA-TPR selection sequence. Therefore, the class of DA-ESPR mechanisms also contains the DA-MiR and DA-TPR mechanisms as its extreme members.

Theorem 15 (Equivalence).

The class of r-PTPR mechanisms is equivalent to the class of DA-ESPR mechanisms.

Proof. We show that each r-PTPR mechanism is equivalent to a DA-ESPR mechanism. Fix an r-PTPR mechanism $p^r = (p_e^r)_{e \in E}$. Let $e \in E$ and fix an entity ranking $\succ_e$. Fix an applicant pool $A \subseteq N$. Given the policy p_e^r for entity e, we show that the number of targeted positions in the equivalent DA-ESPR mechanism for each entity $e \in E$ is $q_e^{rb} = q_e^r - p_e^r$. Given that the r-PTPR entity selection sequence begins selection by choosing $\max(v_e^+ - p_e^r, 0) + q_e^r$ priority agents first according to $\succ_e$ for each entity e, we consider the following two cases.

Case 1. $v_e^+(A, \succ_e) > p_e^r$

In this case the r-PTPR entity selection rule selects $v_e^+ - p_e^r + q_e^r$ priority agents first and fills the rest of the positions regardless of priority status. Since there are more qualified priority agents than p_e^r, the additional qualified priority agents get unreserved positions. Given p_e^r, the number of qualified priority applicants who must take up reserved positions, if we can express the r-PTPR entity selection as a DA-SPR sequential entity selection rule then the sequence must start with p_e^r reserved positions, to ensure that p_e^r qualified agents take up reserved positions. Then the sequence must have $q_e^u - p_e^r$ unreserved positions next, to ensure that the remaining $v_e^+ - p_e^r$ qualified priority agents don't take up reserved positions. Note that this specifies in total the first q_e^u positions, which means that so far in the sequence exactly the v_e^+ qualified priority agents have been selected among all priority agents, and the rest of the agents selected for the first q_e^u positions in the sequence are non-priority agents. After the first q_e^u positions the remaining $q_e^r - p_e^r$ reserved positions must follow in the sequence, which ensures that exactly $p_e^r + v_e^+ - p_e^r + q_e^r - p_e^r = v_e^+ - p_e^r + q_e^r$ priority agents are selected first if there are at least $q_e^r - p_e^r$ unqualified priority agents (if there are fewer unqualified priority agents then some of these reserved positions would not be given to priority agents), corresponding to the r-PTPR entity selection requirement. This leaves the remaining p_e^r unreserved positions last in the sequence, which may allow for selecting additional priority agents. In sum, the r-PTPR selection leads to the following selection sequence for entity e:

$$s_e = (\overbrace{\underbrace{r, \ldots, r,}_{p_e^r} \ u, \ldots, u,}^{q_e^u} | \ \overbrace{\underbrace{r, \ldots, r,}_{q_e^r - p_e^r} \ u, \ldots, u,}^{q_e^r}),$$

This is a DA-ESPR entity selection sequence with $p_e^r = q_e^{rf}$ and $q_e^r - p_e^r = q_e^{rb}$.

Case 2. $v_e^+(A, \succ_e) \leq p_e^r$

In this case the r-PTPR entity selection rule selects q_e^r priority agents first and fills the rest of the positions regardless of priority status. Since $v_e^+(A, \succ_e) \leq p_e^r$, all qualified priority agents have to take up a reserved position, and this corresponds to the DA-MiR selection sequence, which is a DA-ESPR entity selection sequence.

Therefore, in both cases the r-PTPR entity selection rule is equivalent to a DA-ESPR entity selection sequence. Since this argument applies to all entities and entity rankings, we can conclude that each r-PTPR mechanism is equivalent to a DA-ESPR mechanism. Finally, note that since both $0 \leq q_e^{rf} \leq q_e^r$ and $0 \leq p_e^r \leq q_e^r$, and we can always find an r-PTPR selection rule with $p_e^r = q_e^{rf}$, as in Case 1, each DA-ESPR mechanism is also equivalent to an r-PTPR mechanism. This completes the proof. □

Theorem 15 is somewhat surprising, since we started with a policy question concerning how many qualified priority agents have to take up reserved positions, without having any sequence of positions in mind, and we ended up with a sequential DA-SPR entity selection rule. Moreover, one would expect that if there is an equivalence with DA-SPR entity selection sequences then these would be Split DA-SPR sequences, but this is only true in the extreme cases. However, DA-ESPR selection sequences still explicitly divide the reserved positions into targeted versus non-targeted positions, although the targeted positions are different from the ones in Split DA-SPR sequences.

The following example illustrates Theorem 15 and its proof.

Example 7 (r-PTPR versus DA-ESPR entity selections).

Fix an applicant pool $A \subseteq N$ and entity $e \in E$. Let $A^0 = \{i_1, ...i_7\}$ and $A^+ = \{i_8, ...i_{14}\}$. Let $q_e^r = 4$ and $q_e^u = 7$. Let entity e's ranking[3] over the agents in applicant pool A be given by

$$\succ_e: i_1, i_2, i_{12}, i_{14}, i_8, i_4, i_{10}, \,|\, i_{11}, i_5, i_3, i_6, i_9, i_7, i_{13}.$$

Observe that $v_e^+(A, \succ_e) = 4$, since there are four priority agents among the $q_e^u = 7$ highest-ranked agents according to $\succ_e$. We will consider several different policy numbers $p_e^r \leq q_e^r$.

Let $p_e^r = 4$. Then the r-PTPR entity selection rule requires that $q_e^r = 4$ priority agents are selected first, namely $i_{12}, i_{14}, i_8, i_{10}$, plus 7 more agents from the remaining applicants, which means selecting $i_1, i_2, i_4, i_{11}, i_5, i_3, i_6$ in addition. This corresponds to the DA-MiR entity selection sequence $(r, r, r, r, u, u, u, \,|\, u, u, u, u)$, which leads to the following set of selected agents:

$$\{i_{12}, i_{14}, i_8, i_{10}, i_1, i_2, i_4, \,|\, i_{11}, i_5, i_3, i_6\}.$$

Let $p_e^r = 3$. Then the r-PTPR entity selection rule requires that $1 + q_e^r = 5$ priority agents are selected first, namely $i_{12}, i_{14}, i_8, i_{10}, i_{11}$, plus 6 more agents from the remaining applicants, which means selecting $i_1, i_2, i_4, i_5, i_3, i_6$ in addition. This corresponds to

[3] In the examples we use the notation $|$ to mark the cut-off for the first q_e^u positions in the entity ranking, the entity selection sequences and in the selected applicant sets which are ordered according to the relevant selection sequence.

the DA-ESPR entity selection sequence $(r, r, r, u, u, u, u, | \, r, u, u, u)$, which leads to the following set of selected agents:

$$\{i_{12}, i_{14}, i_8, i_1, i_2, i_4, i_{10}, | \, i_{11}, i_5, i_3, i_6\}.$$

Let $p_e^r = 1$. Then the r-PTPR entity selection rule requires that $3 + q_e^r = 7$ priority agents are selected first, namely $i_{12}, i_{14}, i_8, i_{10}, i_{11}, i_9, i_{13}$, plus 4 more agents from the remaining applicants, which means selecting i_1, i_2, i_4, i_5 in addition. This corresponds to the DA-ESPR entity selection sequence $(r, u, u, u, u, u, u, | \, r, r, r, u)$, which leads to the following set of selected agents:

$$\{i_{12}, i_1, i_2, i_{14}, i_8, i_4, i_{10}, | \, i_{11}, i_9, i_{13}, i_5\}.$$

Let $p_e^r = 0$. Then the r-PTPR entity selection rule requires that $4 + q_e^r = 8$ priority agents are selected first, namely $i_{12}, i_{14}, i_8, i_{10}$, i_{11}, i_9, i_{13}, i_1, plus 3 more agents from the remaining applicants, which means selecting i_2, i_4, i_5 in addition. Note that we couldn't select 8 priority agents since there are only 7 priority agents. This corresponds to the DA-TPR entity selection sequence $(u, u, u, u, u, u, u, | \, r, r, r, r)$, which leads to the following set of selected agents:

$$\{i_1, i_2, i_{12}, i_{14}, i_8, i_4, i_{10}, | \, i_{11}, i_9, i_{13}, i_5\}.$$

4.4.2 Properties of Reserved-PTPR Mechanisms

Since DA-ESPR mechanism are DA-SPR mechanisms, Theorem 15 implies that the class of r-PTPR mechanisms is also a subclass of the class of DA-SPR mechanisms. Therefore, r-PTPR mechanisms possess all the nice properties of DA-SPR mechanisms that are presented in Chapter 3.

Weak reserve-stability:

A matching μ is weakly reserve-stable at profile $(P, \succ)$ if the following hold:

1. For all $i \in N$, $\mu(i) R_i i$,

2. If there are $i \in N$ and $e \in E$ such that $e P_i \mu(i)$, then $|\mu_e| = q_e$ and

(2a) if $i \in N^+$ then, for all $j \in \mu_e$, $j \succ_e i$ and $|\mu_e \cap N^+| \geq q_e^r$

(2b) if $i \in N^0$ then, for all $j \in (\mu_e \cap N^0)$, $j \succ_e i$ and $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| \leq q_e^r$.

A matching mechanism is weakly reserve-stable if it assigns a weakly reserve-stable matching to each profile.

Constrained efficiency. A matching mechanism f is constrained efficient with respect to weak reserve-stability if, for all profiles $(P, \succ) \in (\mathcal{P} \times \Pi)$, whenever a matching ν Pareto-dominates $f(P)$ at P, ν is not weakly reserve-stable at P.

A matching mechanism f is **strategyproof** if for all $i \in N$, for all $(P, \succ) \in (\mathcal{P} \times \Pi)$ and all P'_i, $f_i(P, \succ) R_i \; f_i((P'_i, P_{-i}), \succ)$. A coalition $L \subseteq N$ can **manipulate** matching $f(P, \succ)$ at $(P, \succ)$ if there exists P'_L such that for all $i \in S$, $f_i((P'_L, P_{-L}), \succ) P_i f_i(P, \succ)$. A matching mechanism f is **weakly group-strategyproof** if there is no coalition that can manipulate the matching $f(P, \succ)$ at any profile $(P, \succ) \in (\mathcal{P} \times \Pi)$.

Now we are ready to state the corollary to Theorem 15 and Theorems 8, 9, and 10 in Chapter 3, showing the properties of r-PTPR mechanisms.

Corollary 5. *The r-PTPR mechanisms are weakly reserve-stable, constrained efficient with respect to weak reserve-stability, and weakly group-strategyproof.*

Note that since r-PTPR mechanisms are weakly group-strategyproof, it follows that they are also strategyproof.

Furthermore, similarly to Split DA-SPR entity selection sequences, any two DA-ESPR entity selection sequences are also comparable to each other, as defined in Chapter 3, and thus we can state the following corollary, which is similar to Corollary 2 on Split DA-SPR entity selection rules and mechanisms in Chapter 3. For the relevant definitions we refer to Chapter 3.

Corollary 6. *Let f and g be two r-PTPR (or, equivalently, DA-ESPR) mechanisms such that f is more targeted than g. Then for each entity $e \in E$ the entity selection rule used by f is either the same or N^+-superior to the entity selection rule used by g. Moreover, mechanism f N^+-dominates mechanism g.*

4.5 Unreserved-PTPR Mechanisms

We now begin our analysis of the class of u-PTPR mechanisms, which is the class of DA mechanisms with entity selection rules where each member specifies for each entity the maximum number of qualified priority agents who are allowed to get unreserved positions. First we formally define the u-PTPR mechanisms.

u-PTPR mechanisms

u-PTPR mechanisms are parameterized by the policy profile $(p_e^u)_{e \in E}$, which specifies a policy number p_e^u for each entity e, where $0 \leq p_e^u \leq q_e^u$ for all $e \in E$. p_e^u is the number of qualified priority agents in the applicant pool who are allowed to fill unreserved positions, and consequently don't use up a reserved position.

Fix a policy profile $p^u \equiv (p_e^u)_{e \in E}$. The u-PTPR mechanism specified by p^u is defined by the following entity selection rule. For each entity $e \in E$ with entity ranking $\succ_e$, and for each applicant pool $A \subseteq N$:

i. If $|A| \leq q_e$, select A.

ii. If $|A| > q_e$:

ii.a. Select $\min(p_e^u, v_e^+(A, \succ_e)) + q_e^r$ priority agents according to entity ranking $\succ_e$

(i.e., select the highest-ranked priority agents according to $\succ_e$). If $|A^+| \leq \min(p_e^u, v_e^+(A, \succ_e)) + q_e^r$, then select all priority agents.

ii.b. Fill the remaining positions from the remaining agents in the applicant pool according to $\succ_e$ such that a total of q_e agents are selected.

Examining the lower and upper limits of the policy number p_e^u, we can verify that the extreme members of this class of mechanisms are also the DA-MiR and DA-TPR mechanisms. If the policy number is set at zero ($p_e^u = 0$), meaning that none of the qualified priority agents are allowed to fill unreserved positions, then we get the DA-MiR entity selection rule. Thus, the policy profile $p_e^u = 0$ for all $e \in E$ leads to the DA-MiR mechanism. If the policy number is set to equal the number of unreserved positions ($p_e^u = q_e^u$), meaning that all of the unreserved positions can be filled by qualified priority agents (but this depends on the number of qualified priority agents, of course), then we get the DA-TPR entity selection rule, since if $p_e^u = q_e^u$ then $v_e^+(A, \succ_e) + q_e^r$ priority agents are selected first, and then the remaining positions are filled from the remaining set of applicants. This corresponds to selecting all qualified priority agents, and choosing in addition q_e^r priority agents who are not qualified (if there are enough such priority agents), which is exactly the DA-TPR entity selection rule. Thus, the policy profile $p_e^u = q_e^u$ for all $e \in E$ leads to the DA-TPR mechanism.

Example 8 (u-PTPR entity selection.).
Fix an applicant pool $A \subseteq N$ and entity $e \in E$. Let $q_e^r = 3$ and $q_e^u = 6$. Let $A^0 = \{i_1, ...i_8\}$ and $A^+ = \{i_9, ...i_{14}\}$. Let entity e's ranking over the agents in applicant pool A be given by

$$\succ_e: i_1, i_2, i_{10}, i_6, i_9, i_{11}, | \, i_4, i_3, i_{12}, i_5, i_7, i_8, i_{13}, i_{14}.$$

Observe that $v_e^+(A, \succ_e) = 3$, since there are three priority agents among the $q_e^u = 6$ highest-ranked agents according to $\succ_e$. We will consider several different policy numbers $p_e^u \leq q_e^u$.

Let $3 = v_e^+ \leq p_e^u \leq q_e^u = 6$. Here p_e^u is not binding, so $\min(p_e^u, v_e^+) = v_e^+ = 3$. Then

e selects $v_e^+ + q_e^r = 3 + 3 = 6$ priority agents first: $i_{10}, i_9, i_{11}, i_{12}, i_{13}, i_{14}$. Entity e fills next the remaining 3 positions with i_1, i_2, i_6. Observe that this result corresponds to the DA-TPR selection and the number of selected priority agents is 6, which is the highest possible number.

Let $p_e^u = 2$. Here p_e^u is binding, thus $\min\{p_e^u, v_e^+\} = p_e^u = 2$. Then e selects $p_e^u + q_e^r = 2 + 3 = 5$ priority agents first: $i_{10}, i_9, i_{11}, i_{12}, i_{13}$. Entity e fills the remaining 4 positions with i_1, i_2, i_6, i_4. The number of selected priority agents is 5.

Let $p_e^u = 1$. Here p_e^u is binding, thus $\min\{p_e^u, v_e^+\} = p_e^u = 1$. Then e selects $p_e^u + q_e^r = 1 + 3 = 4$ priority agents first: $i_{10}, i_9, i_{11}, i_{12}$. Entity e fills the remaining 5 positions with i_1, i_2, i_6, i_4, i_3. The number of selected priority agents is 4. Observe that this result coincides with the DA-MiR selection.

Let $p_e^u = 0$. Here p_e^u is binding, thus $\min\{p_e^u, v_e^+\} = p_e^u = 0$. Then e selects $p_e^u + q_e^r = 0 + 3 = 3$ priority agents first: i_{10}, i_9, i_{11}. Entity e fills the remaining 6 positions with $i_1, i_2, i_6, i_4, i_3, i_{12}$. This result also corresponds to the DA-MiR selection. Note that a priority agent, i_{12}, is selected in the second step, which implies that this is a DA-MiR selection where the reserve is exceeded. The number of selected priority agents is 4 which is greater than $q_e^r = 3$. ◇

When we check for DA-SPR entity selection sequences that correspond to a u-PTPR entity selection rule, we find that the equivalent DA-SPR selection sequence is a function of the applicant pool, which rules out the equivalence to any DA-SPR entity selection sequence. Thus, we cannot obtain a similar result to Theorem 15 for u-PTPR mechanisms. We state this result next.

Theorem 16. *u-PTPR mechanisms are not equivalent to DA-SPR mechanisms.*

Proof. Let $N^0 = \{i_1, ...i_8\}$ and $N^+ = \{i_9, ...i_{15}\}$. Fix an entity $e \in E$, and let $q_e^r = 3$ and

$q_e^u = 4$. Let entity e's ranking be given by

$$\succ_e: i_1, i_2, i_3, i_{12}, | \, i_8, i_4, i_{10}, i_7, i_5, i_{11}, i_9, i_6, i_{14}, i_{13}, i_{15}.$$

Fix the policy number $p_e^u = 1$.

Consider applicant pool $A = \{i_1, i_2, i_3, i_{12}, | \, i_4, i_7, i_5, i_{11}, i_9, i_6, i_{14}, i_{15}\}$. Then $v_e^+(A, \succ_e) = 1$. Observe that the u-PTPR selection rule with $p_e^u = 1$ requires selecting $1 + q_e^r = 4$ priority agents first, namely $i_{12}, i_{11}, i_9, i_{14}$, plus 3 more agents from the remaining applicant pool, which results in selecting i_1, i_2, i_3. The same set of agents is selected by the DA-TPR entity selection sequence $(u, u, u, u, | \, r, r, r)$. Notice that this is the only DA-SPR entity selection sequence that leads to this selection, because the policy number p_e^u is set at 1, and thus the qualified priority agent i_{12} is allowed to take one of the unreserved positions, which would be impossible using any entity selection sequence in which the first reserved position is in one of the first four positions in the sequence.

Now consider applicant pool $A' = \{i_1, i_{12}, i_{10}, i_7, | \, i_5, i_{11}, i_9, i_6, i_{14}, i_{15}\}$. Then $v_e^+(A', \succ_e) = 2$. Observe that the u-PTPR selection rule with $p_e^u = 1$ requires selecting $1+q_e^r = 4$ priority agents first, namely $i_{12}, i_{10}, i_{11}, i_9$, plus 3 more agents from the remaining applicant pool, which results in selecting i_1, i_7, i_5. However, the DA-SPR entity selection sequence $(u, u, u, u, | \, r, r, r)$ leads to selecting $\{i_1, i_{12}, i_{10}, i_7, | \, i_{11}, i_9, i_{14}\}$, which does not give the same result. Hence, we cannot find a DA-SPR selection sequence that is equivalent to this u-PTPR entity selection rule with $p_e^u = 1$.

It is easily verified that such a counterexample can be generalized, as long as there are enough priority and non-priority agents in the applicant pool to compete for selection. In general, depending on the applicant pool A, the equivalent DA-SPR selection sequence varies under the u-PTPR procedure. Therefore, while we can always find a DA-SPR entity selection sequence that picks the number of priority agents selected by a u-PTPR entity selection rule for a given applicant pool, we cannot find any DA-SPR entity selection sequence that works for all applicant pools and all entity rankings. $\square$

Although the class of u-PTPR mechanisms is not a subclass of the class of DA-SPR mechanisms, which is characterized by weak reserve-stability and strategyproofness, they are nonetheless weakly reserve-stable.

Theorem 17. *The u-PTPR mechanisms are weakly reserve-stable.*

Proof. Fix a policy profile $(p_e^u)_{e\in E}$ and let f denote the corresponding u-PTPR mechanism. Fix a preference profile $P \in \mathcal{P}$ and let $\succ \in \Pi$ be the fixed entity ranking profile. It is readily checked that condition 1 and the first statement on quota saturation in condition 2 of the definition of weak reserve-stability hold at P, so suppose that there exists a blocking pair (i, e) at P with respect to weak reserves, and thus either (2a) or (2b) does not hold in the definition. Let $\mu \equiv f(P)$. Then $eP_i\mu(i)$ by assumption, which implies that i applied to entity e in some step of the u-PTPR procedure f at P and was rejected by e at this step or in a later step.

Case 1. Let $i \in N^+$. If there exists $j \in \mu_e$ such that $i \succ_e j$, then if j is selected by e according to entity ranking $\succ_e$ in the last step of the u-PTPR procedure f at $(P, \succ)$ (either as part of the $\min(p_e^u, v_e^+(A, \succ_e)) + q_e^r$ priority agents in Step (ii.a) or as part of filling the remaining positions from the remaining agents in the applicant pool according to $\succ_e$ such that a total of q_e agents are selected in Step (ii.b)), then i would not have been rejected by e in either the last step or in any previous step of the procedure. Suppose that $|\mu_e \cap N^+| < q_e^r$. Since the u-PTPR entity selection rule selects $\min(p_e^u, v_e^+) + q_e^r \geq q_e^r$ priority agents in each step when there are enough applicants in the applicant pool, this means that all priority agents who applied to entity e in any step of the u-PTPR procedure f at $(P, \succ)$ have been selected by e, contradicting the fact that i had applied.

Case 2. Let $i \in N^0$. If there exists $j \in (\mu_e \cap N^0)$ such that $i \succ_e j$, then if j is selected by e in the last step of the u-PTPR procedure f at $(P, \succ)$ (either as part of the agents selected in Step (ii.a) of the u-PTPR procedure when there are not enough priority agents or as part of filling the remaining positions from the remaining agents in the applicant pool according to $\succ_e$ such that a total of q_e agents are selected in Step (ii.b)) then i

would not have been rejected by e in either the last step or in any previous step of the procedure. Suppose that $|\{h \in (\mu_e \cap N^+) : i \succ_e h\}| > q_e^r$. This implies that there exists at least one priority agent h with $i \succ_e h$ such that h was selected in the last step in the u-PTPR procedure at $(P, \succ)$ as part of filling the remaining positions from the remaining agents in the applicant pool in Step (ii.b). But then this position that h filled would have been filled in either the last or in some previous step by i, when i applied to e, given that $eP_i\mu(i)$, and this position would have been filled in the last step either by agent i or by some agent j such that $j \succ_e i$. Since $i \succ_e h$, $j \neq h$, and we have a contradiction.

In either case there is no pair (i, e) that blocks μ with respect to weak reserves at $(P, \succ)$. Therefore, f is weakly reserve-stable. □

Since u-PTPR mechanisms are not DA-SPR mechanisms by Theorem 16, but satisfy weak reserve-stability by Theorem 17, it follows from the characterization of DA-SPR mechanisms (Theorem 11 in Chapter 3) that u-PTPR mechanisms are not strategyproof. We state this result below.

Corollary 7. *The u-PTPR mechanisms are not strategyproof.*

It is striking that the classes of r-PTPR and u-PTPR mechanisms lead to such different results. Specifically, while both the u-PTPR and r-PTPR mechanisms are weakly reserve-stable, r-PTPR mechanisms are weakly group-strategyproof by Corollary 5, while u-PTPR mechanisms are not even strategyproof. Given that the difference between the policy options for the two different classes of mechanisms do not appear to be significant, we can conclude that for the policy designer r-PTPR mechanisms would very likely be more attractive.

4.6 Conclusion

Both the r-PTPR and u-PTPR mechanisms offer a range of policies between the DA-TPR and the DA-MiR mechanisms, providing the designer with flexibility and clarity when choosing a priority reserve policy. Given that the policy choices of the two matching mechanisms only differ in a subtle way, it is surprising to see that these two classes of matching mechanisms lead to substantially different different results. When comparing the two classes of matching mechanisms, both exhibit good stability properties, since both satisfy weak reserve-stability. However, based on their incentive properties, the r-PTPR mechanisms appear to be more attractive, as they are strategyproof (in fact, weakly group-strategyproof), while u-PTPR mechanisms are not strategyproof. The r-PTPR mechanisms are also more transparent, since we can easily understand how they work based on their equivalence to the DA-ESPR mechanisms with explicit entity selection sequences. This equivalence result also highlights that the r-PTPR mechanisms allow for a compromise between the two goals of protection and representation by dividing the reserved positions into targeted and non-targeted reserved positions. It would be more difficult to make a similar distinction for unreserved positions, which may explain why focusing on how the reserved positions are divided between the two different objectives leads to better priority policy alternatives when the designer's aim is to compromise between these objectives.

Bibliography

[1] Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match. *The American Economic Review 99*, 5 (2009), 1954–1978.

[2] Abdulkadiroglu, A., and Sönmez, T. School choice: A mechanism design approach. *The American Economic Review 93*, 3 (2003), 729–747.

[3] Afacan, M. O., Alioğullari, Z. H., and Barlo, M. Sticky matching in school choice. *Economic Theory 64*, 3 (2017), 509–538.

[4] Alcalde, J., and Barberà, S. Top dominance and the possibility of strategy-proof stable solutions to matching problems. *Economic Theory 4*, 3 (1994), 417–435.

[5] Alfred, C. What history can teach us about the worst refugee crisis since WWII. *The Huffington Post 12* (2015).

[6] Alva, S., and Manjunath, V. Strategy-proof Pareto-improvement. *Journal of Economic Theory 181* (2019), 121–142.

[7] Andersson, T., and Ehlers, L. Assigning refugees to landlords in Sweden: Stable maximum matchings.

[8] Aygün, O., and Bó, I. College admission with multidimensional privileges: The Brazilian affirmative action case. *Available at SSRN 3071751* (2020).

[9] Aygün, O., and Turhan, B. Large-scale affirmative action in school choice: Admissions to IITs in India. *American Economic Review 107*, 5 (2017), 210–13.

[10] Aygün, O., and Turhan, B. Designing direct matching mechanisms for India with comprehensive affirmative action. *Available at SSRN 3582669* (2020).

[11] Aygün, O., and Turhan, B. Dynamic reserves in matching markets. *Journal of Economic Theory 188* (2020), 105069.

[12] Bansak, K., Ferwerda, J., Hainmueller, J., Dillon, A., Hangartner, D., Lawrence, D., and Weinstein, J. Improving refugee integration through data-driven algorithmic assignment. *Science 359*, 6373 (2018), 325–329.

[13] Biró, P., and Gudmundsson, J. Complexity of finding Pareto-efficient allocations of highest welfare. *European Journal of Operational Research* (2020).

[14] Chambers, C. P., and Yenmez, M. B. Choice and matching. *American Economic Journal: Microeconomics 9*, 3 (2017), 126–47.

[15] Delacrétaz, D., Kominers, S. D., Teytelboym, A., et al. Matching mechanisms for refugee resettlement. Tech. rep., 2019.

[16] Doğan, B. Responsive affirmative action in school choice. *Journal of Economic Theory 165* (2016), 69–105.

[17] Doğan, B., Doğan, S., and Yildiz, K. Lexicographic choice under variable capacity constraints. *arXiv preprint arXiv:1910.13237* (2019).

[18] Dubins, L. E., and Freedman, D. A. Machiavelli and the Gale-Shapley algorithm. *The American Mathematical Monthly 88*, 7 (1981), 485–494.

[19] Dur, U., Kominers, S. D., Pathak, P. A., and Sönmez, T. Reserve design: Unintended consequences and the demise of Boston's walk zones. *Journal of Political Economy 126* (2018), 2457–2479.

[20] Dur, U., Mennle, T., and Seuken, S. First-choice maximal and first-choice stable school choice mechanisms. In *Proceedings of the 2018 ACM Conference on Economics and Computation* (2018), pp. 251–268.

[21] Dur, U., Pathak, P., and Sönmez, T. Explicit vs. statistical targeting in affirmative action: Theory and evidence from Chicago's exam schools. *Journal of Economic Theory 187* (2020), 104996.

[22] Echenique, F., and Yenmez, M. B. How to control controlled school choice. *American Economic Review 105*, 8 (2015), 2679–94.

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

[24] Ehlers, L., and Klaus, B. Object allocation via Deferred-Acceptance: Strategy-proofness and comparative statics. *Games and Economic Behavior 97* (2016), 128–146.

[25] Ergin, H. I. Efficient resource allocation on the basis of priorities. *Econometrica 70*, 6 (2002), 2489–2497.

[26] Fleiner, T. A fixed-point approach to stable matchings and some applications. *Mathematics of Operations Research 28*, 1 (2003), 103–126.

[27] Gale, D., and Shapley, L. S. College admissions and the stability of marriage. *The American Mathematical Monthly 69*, 1 (1962), 9–15.

[28] Giuffre, M., and Costello, C. Tragedy'and responsibility in the Mediterranean. *Open Democracy* (2015).

[29] Goto, M., Iwasaki, A., Kawasaki, Y., Kurata, R., Yasuda, Y., and Yokoo, M. Strategyproof matching with regional minimum and maximum quotas. *Artificial intelligence 235* (2016), 40–57.

[30] Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. Effective affirmative action in school choice. *Theoretical Economics 8*, 2 (2013), 325–363.

[31] Hatfield, J. W., and Kominers, S. D. Hidden substitutes. In *Proceedings of the 16th ACM Conference on Electronic Commerce* (2016).

[32] Hatfield, J. W., and Milgrom, P. R. Matching with contracts. *American Economic Review 95*, 4 (2005), 913–935.

[33] Hylland, A., and Zeckhauser, R. The efficient allocation of individuals to positions. *Journal of Political economy 87*, 2 (1979), 293–314.

[34] Jones, W., and Teytelboym, A. Choices, preferences and priorities in a matching system for refugees. *Forced Migration Review*, 51 (2016), 80.

[35] Jones, W., and Teytelboym, A. The international refugee match: A system that respects refugee preferences and the priorities of states. *Refugee Survey Quarterly 36*, 2 (2017), 84–109.

[36] Kesten, O. On two competing mechanisms for priority-based allocation problems. *Journal of Economic Theory 127*, 1 (2006), 155–171.

[37] Kesten, O., and Kurino, M. Strategy-proof improvements upon deferred acceptance: A maximal domain for possibility. *Games and Economic Behavior 117* (2019), 120–143.

[38] Knuth, D. E. *Mariages stables et leurs relations avec d'autres problemes combinatoires: introduction a l'analysis mathematique des algorithmes-.* Les Presses de l'Universite de Montreal, 1976.

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

[40] Kominers, S. D., and Sönmez, T. Matching with slot-specific priorities: Theory. *Theoretical Economics 11*, 2 (2016), 683–710.

[41] Koser, K. *Responding to migration from complex humanitarian emergencies: Lessons learned from Libya.* Chatham House London, 2011.

[42] Morrill, T. An alternative characterization of top trading cycles. *Economic Theory 54*, 1 (2013), 181–197.

[43] Pápai, S. Strategyproof assignment by hierarchical exchange. *Econometrica 68*, 6 (2000), 1403–1433.

[44] Pathak, P. A., Rees-Jones, A., and Sönmez, T. Immigration lottery design: Engineered and coincidental consequences of H-1B reforms. *National Bureau of Economic Research* (2020).

[45] Pathak, P. A., and Sönmez, T. School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation. *American Economic Review 103*, 1 (2013), 80–106.

[46] Pathak, P. A., Sönmez, T., Ünver, M. U., and Yenmez, M. B. Fair allocation of vaccines, ventilators and antiviral treatments: Leaving no ethical value behind in health care rationing. *arXiv preprint arXiv:2008.00374* (2020).

[47] Roth, A. E. The economics of matching: Stability and incentives. *Mathematics of operations research 7*, 4 (1982), 617–628.

[48] Roth, A. E. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of political Economy 92*, 6 (1984), 991–1016.

[49] Roth, A. E. A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom. *The American economic review* (1991), 415–440.

[50] Roth, A. E., and Peranson, E. The effects of the change in the nrmp matching algorithm. *JAMA 278*, 9 (1997), 729–732.

[51] Sayedahmed, D. Host country choice of refugees: A mechanism design approach. *Working Paper, Concordia University* (2017).

[52] Sayedahmed, D. Centralized refugee matching mechanisms with hierarchical priority classes. *Working Paper, Concordia University* (2019).

[53] Schuck, P. H. Refugee burden-sharing: A modest proposal. *Yale J. Int'l L. 22* (1997), 243.

[54] Shapley, L., and Scarf, H. On cores and indivisibility. *Journal of mathematical economics 1*, 1 (1974), 23–37.

[55] Sönmez, T., and Yenmez, M. B. Constitutional implementation of reservation policies in India. *Working paper* (2019).

[56] Sönmez, T., Yenmez, M. B., et al. Affirmative action in India via vertical and horizontal reservations. *Working paper* (2019).

[57] Sönmez, T., Yenmez, M. B., et al. Affirmative action with overlapping reserves. *Working Paper, Boston College Department of Economics* (2019).

[58] Szobolits, A., and Sunjic, M. H. Many refugees misunderstand Schengen expansion. *UN High Commissioner for Refugees (UNHCR)* (2007).

[59] van Basshuysen, P. Towards a fair distribution mechanism for asylum. *Games 8*, 4 (2017), 41.