Essays on Choice and Matching

Kriti Manocha

Thesis Supervisor: Professor Arunava Sen

STATISTICAL INDIAN INSTITUTE

भिन्नेष्वैक्यस्य दर्शनम्

UNITY IN DIVERSITY

Thesis submitted to the Indian Statistical Institute in partial fulfillment of the requirements for the degree of Doctor of Philosophy

November, 2023

Kriti

To my family and teachers.

Acknowledgements

I wish to express my sincere gratitude to my supervisor, Prof. Arunava Sen, for his invaluable mentorship. His insightful perspectives have helped me understand how to approach research in general. I am thankful for the freedom he has given me to explore various topics in economic theory which has encouraged me to explore new avenues of research. I am truly fortunate to have had the privilege of working under his supervision.

I am also extremely grateful to Prof. Debasis Mishra, who has always been a constant support throughout this research journey. He has always been very generous with his time and expertise, providing enlightening feedback and counsel whenever I sought it. I am indebted to them for instilling in me the values of perseverance and determination.

I would like to acknowledge my co-authors – Bhavook Bhardwaj, Rajnish Kumar, Josue Ortega and Bertan Turhan– with whom I have had the privilege of collaborating. Their expertise and willingness to share knowledge have significantly enriched my research experience.

I would also like to extend my appreciation to Nikita, Ojasvita, Param, Sarvesh, Siddharth, Swati, and Vilok. Their presence and encouragement have made my research journey all the more enriching. I benefited immensely from interactions in reading groups and seminar sessions organized by the department. I have been blessed to be a part of an extremely supportive ecosystem provided by the Indian Statistical Institute.

On a personal note, I am deeply grateful to my husband and my parents for their unconditional love and constant encouragement. Their sacrifices and belief in my abilities have been a constant source of motivation throughout this endeavor.

Abstract

This thesis consists of three independent essays. The first chapter introduces a model of decision-making that is based on the procedure of *rejection*. Departing from the standard model of choice via preference maximization, the decision maker (DM) rejects *minimal* alternatives from a menu according to a preference relation. We axiomatically study the correspondence of non-rejected alternatives which we call the *acceptable correspondence* with different rationality conditions on the underlying preference relation. We also generalize our model to acceptable correspondences that are generated by the *successive* elimination of minimal alternatives. We find that the rejection approach developed in this chapter can offer explanations for various anomalies observed in decision theory, such as the *two-decoy* effect or the *two-compromise* effect (Tserenjigmid (2019)).

The second chapter proposes a sequential model of the college admissions problem. The selection criteria of institutions are formulated via choice rules that admit slot-specific priorities introduced by Kominers and Sönmez (2016). We show that the applicants can not be worse off in the subsequent stages when the candidates update their preferences that adhere to their assignment in the previous stage. Moreover, the mechanism that sequentially implements individual-proposing deferred acceptance is stable with respect to a generalized version of a sequential stability notion provided in this chapter. These results generalize the findings presented in Haeringer and Iehlé (2021). We use our results to analyze recently reformed admission procedures for engineering colleges in India (Baswana et al. (2019)), where applicants are provided various options to update their preferences in additional stages.

In the third chapter, we study the welfare consequences of merging Shapley–Scarf housing markets (Shapley and Scarf (1974)). We show that in the worst-case scenario, market integration can lead to large welfare losses and make the vast majority of agents worse off. However, on average, the integration is welfare enhancing and makes all agents better off ex-ante. The number of agents harmed by integration is a minority when all markets are small or the agent's preferences are highly correlated.

Contents

Chapter 0

Introduction

This thesis consists of three independent essays that span topics in Choice and Matching theory. Chapter 1 (co-authored with Bhavook Bhardwaj) studies the rejection behaviour of a decision-maker. Chapter 2 (co-authored with Bertan Turhan) proposes a sequential college admissions (two-sided matching) model that applies to engineering college admissions in India. Chapter 3 (co-authored with Josué Ortega and Rajnish Kumar) studies the consequences of merging Shapley–Scarf (one-sided matching) markets.

We provide a brief description of each chapter below.

Chapter 1. Rejection and Acceptable Correspondences

This chapter introduces a model of decision-making based on the procedure of *rejection*. The decision maker (DM) is endowed with an asymmetric binary relation, which we call a *rationale*. Departing from the standard model of choice via preference maximization, the DM *rejects* at least the "worst" alternatives according to the rationale. For any menu, the worst alternative, referred to as the *minimal* alternative (i) does not "dominate" any alternative, and (ii) is "dominated" by at least one alternative according to the rationale.

We axiomatically study the correspondence of non-rejected alternatives which we call the *acceptable correspondence* with different rationality conditions on the underlying rationale. The various classes of acceptable correspondences are characterized using modified versions of *Contraction Consistency* and *Expansion Consistency*, which are central to the characterization of maximal-element rationalizable (MR) correspondence (see Sen (1971), Bossert et al. (2005)). An important implication of our results is that selection by rejection does not follow from appropriate versions of selecting maximals.

In other words, the procedure of rejection is substantively different from selection by maximizing procedures.

The second part of our analysis considers a more general procedure where the DM rejects the $k-$minimal set that is obtained by successively rejecting the worst alternatives according to the underlying rationale. The DM is now represented by a pair (R, k), where R is the underlying transitive rationale and k is a threshold function. The threshold captures the idea of "satisficing" behaviour of Simon (1955), where the DM selects the alternatives above a threshold. It also provides a measure of the degree of rationality as discussed in Barberà et al. (2019), where a higher threshold indicates that the DM is closer to the benchmark of the maximal set. We present the characterization of two variants of the correspondence that rejects the $k-$minimal set, which we call the threshold-acceptable (TA) correspondence. The first is when the threshold is fixed across menus. The second is when the threshold can depend on the menu and accommodate phenomena like choice overload (Frick (2016)). Both variants of this correspondence have MR correspondence as a special case.

We find that the rejection approach developed in this chapter can offer explanations for various anomalies observed in decision theory. It includes behaviour like cyclic choices, choice reversals via the *decoy effect* and the *compromise effect* (see Huber et al. (1982), Simonson (1989)). Threshold acceptability on the other hand explains phenomenon like *two-decoy effect* and *two-compromise effect* (Tserenjigmid (2019)) (Teppan and Felfernig (2009) and Manzini and Mariotti (2010)). As k increases, the TA correspondence merges with the MR correspondence.

Chapter 2. sequential matching with affirmative action

In this chapter, we propose a sequential model of the college admissions problem. At each stage, applicants can decide whether to finalize their assignments or participate in the next stage by updating their submitted preferences. We assume that the selection criteria of institutions are formulated via choice rules that admit slot-specific priorities introduced by Kominers and Sönmez (2016). Each institution has a set of positions (slots) that can be assigned to different individuals. Positions have their own (potentially independent) rankings for individuals. Within each institution, a linear order – referred to as the precedence order – determines the sequence in which positions are filled.

We investigate the restrictions on the preferences of individuals across different rounds that result in *monotone* outcomes. We refer to a matching outcome as monotone when each individual is matched to an institution that is weakly higher than the match of previous rounds. The proof of this result utilizes ideas in Kojima and Manea (2010). It enables us to compare their static problem with our sequential model. We introduce a "backward-looking" notion of stability for sequential matching mechanisms that takes into consideration individual rationality, non-wastefulness and justified envy of individuals across different rounds. We establish a relationship between this notion of stability, which we refer to as sequential stability, and monotone outcomes.

We apply our theoretical findings to analyze the admissions problem in engineering colleges in India. Admissions to engineering colleges in India are based on a multi-period semi-centralized matching process and are subject to sophisticated affirmative action constraints. Baswana et al. (2019) claim that the welfare of candidates improves in every round of their mechanism, but do not provide a rigorous theoretical justification for their claim. We do that by adopting the model in Haeringer and Iehlé (2021).

3. On the integration of Shapley-scarf markets

In this chapter, we investigate the welfare effects of integrating disjoint Shapley–Scarf markets (Shapley and Scarf (1974)). We assume that the core allocation is implemented before and after the merge occurs. The gains/losses from integration are assessed by (i) the number of agents who obtain a better allocation, and (ii) the size of the welfare gains, measured by the rank of the assigned house. We present worst- and average-case results.

In the worst-case scenario, we show that when k communities with n_j agents each merge with n agents in total, it may harm a majority of agents (up to, but no more than, $n-k$ agents). Furthermore, the average rank of an agent's house can decrease asymptotically by, but not more than, 50% of the length of their preference list. These results are substantially worse than those for Gale–Shapley markets Ortega (2018, 2019). On the other side, the average-case results are more optimistic. We prove that the expected gains from integration in random housing markets are equal to $\frac{(n+1)[(n_j+1)H_{n_j}-n_j]}{n_j(n_j+1)n} - \frac{(n+1)H_n-n}{n^2}$, where H_n is the n-th harmonic number. Our computation shows that the expected welfare gains from integration are positive for all agents and larger for agents that are initially in smaller markets. We also provide an upper bound on the expected number of agents harmed by integration. This allows us to guarantee that a majority of agents benefit from integration when all the markets are of equal size. The average case results rely on the previous probabilistic analysis of Shapley–Scarf markets in the computer science literature (see Frieze and Pittel (1995); Knuth (1996)).

To guarantee that no more than half of the agents in any individual market are harmed by integration, we identify a preference domain, which we call the *sequential dual dictatorship*, that enforces a particular correlation among agent's preferences. This property is equivalent to assigning the title of the dictator to at most two agents at each step of the top trading cycle algorithm, thereby bounding the length of cycles that can occur.

Chapter 1

Rejection and Acceptable Correspondences[1]

1.1 Introduction

The standard theory of decision-making is founded on the principle of maximizing behaviour of the decision-maker, henceforth a DM. In the classical model, first developed in Samuelson (1938), the DM selects from subsets of a given set of alternatives, called menus. The DM chooses the set of maximal alternatives with respect to her underlying preference ordering over the set[2]. A recent strand in the literature refines this procedure to explain non-standard choice behaviour. The DM first forms a "consideration set" by means of some procedure[3] and in a subsequent stage, chooses a maximal alternative from the remaining set using the underlying preference ordering. Significantly, this initial stage involves some form of maximizing behaviour on the part of the DM, resulting in the elimination of non-maximal alternatives. In this chapter, we develop an alternative approach to consideration set formation where the DM, instead of maximizing, *rejects* a set of "worst" alternatives. We axiomatically study this set of non-rejected alternatives; referred to as the *acceptable* set.

Extensive research in consumer psychology and decision-making suggests that selection by choice and selection by rejection typically do not lead to the same outcome. In the experiments, Huber et al. (1987) found that rejection behaviour typically generates larger

[1]Co-authored with Bhavook Bhardwaj (Indian Statistical Institute, Delhi Centre).

[2]The early literature in choice theory has a thorough treatment of decision-making via preference maximization (see Houthakker (1950), Chernoff (1954) Arrow (1959), Sen (1971)).

[3]This includes "Shortlisting" (Manzini and Mariotti (2007)), "Categorization" (Manzini and Mariotti (2012)) and "Rationalization" (Cherepanov et al. (2013)). Lleras et al. (2017) provides a list of these procedures to illustrate that the formation of consideration sets varies significantly.

consideration sets compared to choice behaviour. Several explanations may account for the observed phenomenon. According to Shafir (1993), individuals tend to direct their focus towards positive attributes while choosing, and towards negative attributes while rejecting. Yaniv and Schul (2000) provide a theoretical framework suggesting that choice and rejection procedures imply different types of status quo for the alternatives, thereby invoking a different selection criterion for each procedure. Laran and Wilcox (2011) find (in their Study 4) that cognitive load affects rejection more than choice, suggesting that rejection might involve more deliberative processing[4].

In certain situations, we wish to consider, rejection seems more appealing and plausible than selection by choice. Consider a hiring process for a company looking to fill a managerial position. The DM (hiring manager) has received several applications from candidates. Before making the final selection through the interview, the process involves screening the candidates. The DM compares the candidates with either a Finance or Engineering degree based on years of managerial experience. The DM then narrows down the candidate pool by eliminating those with the lowest years of experience in each category of qualification. Similarly, consider a case where a Ph.D. candidate has a GRE score of X and research interests in Macroeconomics and Labour Economics. While applying, the DM rejects all universities that require a GRE score of more than X and do not have outstanding faculty in either of her research interests. There are situations —such as wishlisting, adding items in a cart, saving articles to read later, etc —where the selection is more appropriately described by rejection rather than by maximization. The process of selection may involve the approval of several alternatives such as described in Manzini et al. (2019).

We consider a model where the DM is endowed with an asymmetric[5] binary relation, which we call a rationale[6]. For any menu, the set of worst alternatives with respect to the rationale is referred to as the *minimal set*. An alternative is minimal if (i) it does not "strictly dominate" any alternative and (ii) is strictly dominated by at least one alternative. In the earlier example, suppose candidates $\{a, b, c, d, e\}$ have applied for the position. Candidates a, b and d have Finance degrees while c and e have Engineering

[4]See Sokolova and Krishna (2016) for detailed empirical evidence suggesting that choice and rejection procedures result in different outcomes.

[5]When the underlying preference relation is not complete, Eliaz and Ok (2006) show that one can distinguish between indifference and indecisiveness of an agent by observing her choice behaviour. To avoid technical complications, we assume the underlying relation to be asymmetric. For further discussion on this issue, see Eliaz and Ok (2006).

[6]We follow the terminology of Manzini and Mariotti (2007)

degrees. Based on years of experience, the DM ranks a better than b and b better than d. Also, c is ranked better than e. The DM cannot compare candidates with different degrees. The minimal alternatives in this case are d and e.

Using the notion of minimal alternatives, we define the acceptable set for a given menu. An acceptable set is a subset of a menu that satisfies two properties. The first is that the minimal alternatives are rejected from the menu. Secondly, maximal alternatives are not rejected from the menu. In the earlier example, the acceptable sets are $\{a, c\}$ and $\{a, b, c\}$. An *acceptable correspondence* is a mapping that associates an acceptable set for every menu. Within the class of acceptable correspondences, we mainly focus on two subclasses. The first one, referred to as *maximal acceptable* (MA) correspondence, selects the largest acceptable set for every menu. In other words, the DM selects all alternatives except the minimal set in the menu. In the example above, the MA correspondence selects $\{a, b, c\}$ from the menu $\{a, b, c, d, e\}$. We explore refinements of this subclass of correspondences with additional restrictions imposed on the underlying rationale. These restrictions include acyclicity, completeness, transitivity, and linearity.

We provide characterization results of these classes of acceptable correspondences. The axioms we use are a modified version of *Contraction Consistency* and a modified version of *Expansion Consistency*, which are central to the characterization of maximal-element rationalizable (MR) correspondences (see Sen (1971), Bossert et al. (2005)). Acceptable correspondences are characterized by a weakening of both axioms, restricting to binary menus (see Theorem 1.1). Theorem 1.2 provides characterization results for MA correspondences, where the first one is a weakening of Contraction Consistency and the second one is a mild strengthening of Expansion Consistency. Theorem 1.3 and 1.4 characterize MA correspondences with added restrictions on the underlying rationale. An important implication of our results is that selection by rejection does not follow from appropriate versions of selecting maximals. In other words, the procedure of rejection is substantively different from selection by maximizing procedures (see discussion in Section 1.4.2).

The second part of our analysis considers a more general procedure where the DM rejects a larger *set* of worst alternatives using the underlying rationale. For convenience, we assume that the rationale satisfies transitivity. The DM is now represented by a pair (R, k), where R is the underlying transitive rationale and k is the threshold function. Such a DM rejects the "worst" k layers of minimal alternatives with respect to R. We call this rejected set as $k-$*minimal* set. The set of non-rejected alternatives is referred to as a *threshold-acceptable* set and the associated correspondence to be threshold-acceptable (TA) correspondence. The TA correspondence is a natural generalization of MA correspondences. The threshold captures the idea of "satisficing" behaviour of Simon (1955), where the DM selects the alternatives above a threshold. It also provides a measure of the degree of rationality (as discussed in Barberà et al. (2019)) where a higher threshold indicates that the DM is closer to the benchmark of the maximal set.

We consider two variants of TA correspondences. The first is when the threshold is fixed across menus. The second is when the threshold can depend on the menu and accommodate phenomena like choice overload (Frick (2016), Kovach and Ülkü (2020)). Both variants of this correspondence have MR correspondences as a special case. We provide a characterization result for both these models as Theorem 1.5 and 1.6 respectively.

Maximal acceptability can explain a wide variety of empirically observed choice behaviour that cannot be rationalized by preference maximization. This includes behaviour like cyclic choices, choice reversals via the *decoy effect* and the *compromise effect* (see Huber et al. (1982), Simonson (1989)). Threshold acceptability on the other hand explains phenomena like *two-decoy effect* and *two-compromise effect* (Tserenjigmid (2019), Teppan and Felfernig (2009) and Manzini and Mariotti (2010)). This can be done via low values of k. As k increases, the TA correspondence merges with the MR correspondence. These issues are discussed in detail in Section 1.6.

The layout of this chapter is as follows: the next subsection provides a brief literature review. Section 1.2 introduces the model formally and reviews MR correspondences. Section 1.3 introduces the concept of minimal sets and provides a characterization of acceptable correspondences. Section 1.4 and 1.5 characterize MA and TA correspondences respectively. Section 1.6 provides an application of the model.

1.1.1 Related Literature

There is a vast literature that studies various sequential models of elimination. Tversky (1972) provides a probabilistic theory of choice based on the sequential elimination of alternatives. Each alternative consists of a set of attributes. At each stage, an attribute is selected with some probability, and all the alternatives that do not include the attribute are eliminated. The model assumes no fixed ordering over the attributes. Mandler et al. (2012) presents a simplified (deterministic) version of Tversky (1972). They show that the two approaches of utility maximization and choosing by sequential elimination are nearly equivalent. Manzini and Mariotti (2007) looks at sequentially rationalizable choice functions where the DM eliminates non-maximal alternatives according to rationales until only one alternative remains as the final choice. The order in which the DM applies the rationales is fixed. In our model, the DM rejects a set of alternatives non-sequentially. The set of remaining alternatives is not necessarily a singleton set.

More recently, Masatlioglu and Nakajima (2007) introduced a deterministic theory of choice that is based on the elimination of alternatives. In their model, an alternative is eliminated only if it is dominated by another alternative in its "comparable" set. The decision procedure is characterized by a single property called the "axiom of choice by elimination". According to it, there is at least one alternative from each menu that is always chosen (not necessarily unique) if it is available in a sub-menu. This condition is necessary but not sufficient in our model.

Another paper closely related to ours is Barberà et al. (2019). In their model, the DM is endowed with a linear order and the assumption of perfect optimization is relaxed in favour of "Order$-k$ rationality". The DM chooses from one of the top-k alternatives according to the linear order. This concept is not equivalent to rejecting the bottom $m-k$ alternatives (m is the number of alternatives) in our framework for a number of reasons. First is that we do not restrict ourselves to linear orders. So, the concept of the worst k alternatives is not well defined in our framework. Secondly, we do not restrict to singleton-valued selections. As a result, our characterization results differ considerably from theirs.

Salant and Rubinstein (2008) develop a framework for modeling choice in the presence of framing effects. A frame includes observable information that is irrelevant in the assessment of the alternatives but affects choice. Our model can be regarded as an attempt to formalize a frame that is based on the rejection of alternatives.

There is substantial empirical evidence in the literature that the DM does not consider the entire menu when making a choice, restricting attention instead, to a subset of alternatives. This has been referred to as a *consideration set* of a menu and has been formalized in a variety of ways. Masatlioglu et al. (2012)) formalize this on the grounds of "limited attention" while Manzini and Mariotti (2007) do so in terms of "shortlisting" according to some criterion.

The concept of *partial dominance* proposed in Gerasimou (2016) is related to the notion of minimal alternatives introduced in this chapter. The partially dominant correspondence selects alternatives that are not dominated by any other alternative in the menu *and* dominates some alternatives according to a rationale. All alternatives are selected if none of the pair of alternatives are comparable in the menu. The correspondence that selects the minimal set from each menu differs from the partially dominant correspondence—this is discussed further in Section 1.4.2.

There are several two-stage models of choice (Manzini and Mariotti (2007), Manzini and Mariotti (2012), Cherepanov et al. (2013), Masatlioglu et al. (2016), Bajraj and Ülkü (2015)) where the first stage shortlisted set has different properties. Our analysis can also be interpreted as a contribution to this strand of literature. There are several differences between our model and these two-stage models. Firstly, we do not look at the second stage of final selection, and secondly, our shortlisted set allows for a much richer structure than their set of surviving alternatives. A shortlisting model close to ours is Bhardwaj and Manocha (2021). This paper studies a choice procedure where the set of minimal alternatives according to a transitive rationale are rejected in the first stage. In the second stage, a unique maximal alternative is chosen from the remaining set according to a possibly different rationale. The recent version Bhardwaj and Manocha (2023) introduces a simple model of stochastic choice along the lines of Echenique and Saito (2019). In this model, the DM first rejects the worst alternatives using a binary relation and then follows a Luce procedure to assign probabilities to the non-rejected alternatives.

There are other models that consider more general selections rather than final choices. For instance, in situations like wishlisting, the DM expresses a positive interest in an item by "approving" it. Manzini et al. (2019) provides a probabilistic model of this *approval* behaviour using a stochastic choice function from lists framework (Rubinstein and Salant (2006)). Instead, we work with a deterministic model of selection from *menus*.

1.2 Model

1.2.1 Preliminaries

Let X be a finite set of alternatives and $\mathcal{P}(X)$ be the set of all menus, that is, non-empty subsets of X. A DM makes non-empty selections from every menu by rejecting certain alternatives. A *decision correspondence* $d : \mathcal{P}(X) \to \mathcal{P}(X)$ is a mapping that satisfies the following properties (i) $d(A) \neq \emptyset$ and (ii) $d(A) \subseteq A$ for all $A \in \mathcal{P}(X)$.

A binary relation R over X is a subset of $X \times X$. We will denote by xRy if $(x, y) \in R$ and $\neg xRy$ if $(x, y) \notin R$. The DM in our model is endowed with a strict binary relation R (henceforth, a rationale), that is, an asymmetric relation where R is *asymmetric* (xRy implies $\neg yRx$). If xRy, we shall interpret it as "y is dominated by x".

A rationale is (i) *complete* **(C)** if for any $x, y \in X$ and $x \neq y$, if $\neg xRy$, then yRx (ii) *acyclic* **(Y)** if for any $x_1, \ldots x_k$, if $x_1Rx_2R\ldots x_k$, then $\neg x_kRx_1$ (iii) *transitive* **(T)** if for any x_1, x_2, x_3, if x_1Rx_2 and x_2Rx_3, then x_1Rx_3. Note that acyclicity is equivalent to transitivity under completeness. The implication does not hold without completeness. We will refer to a complete rationale as a *tournament* and a transitive rationale as a *partial order*. A rationale satisfying acyclicity (transitivity) along with completeness is a *linear order*, henceforth abbreviated as **(TC)**. We work with rationales with the interpretation that if DM is not able to rank x and y, she must be indecisive between the two alternatives[7].

1.2.2 Maximal Element Rationalizability Revisited

A decision correspondence is *maximal-element rationalizable* (Bossert et al. (2005)) if it selects just the set of maximal alternatives with respect to an underlying preference relation. The existence of a maximal set is guaranteed with a mild restriction of acyclicity on the preference relation. For a menu A and rationale R, the set of maximal alternatives is defined as

$$M(A, R) := \{x \in A | \ \neg yRx \ \forall y \in A\}$$

Definition 1.1. A decision correspondence $d : \mathcal{P}(X) \to \mathcal{P}(X)$ is **maximal element rationalizable** (MR) if there exists an acyclic rationale R, such that for all menus A,

$$d(A) = M(A, R)$$

Two simple consistency conditions characterize this behaviour, namely Contraction Consistency (CC) and Expansion (Exp) (see Sen (1971)). Also referred to as the Chernoff

[7]The assumption of asymmetry is a mild one. One can alternatively allow for indifferences and develop our model of rejection. However, without a richness assumption on the binary relation (*regularity*), one encounters the problem of distinguishing between indifference and incompleteness in the identification exercise (see Eliaz and Ok (2006))

axiom or Condition α, the CC condition states that if an alternative is selected in a menu, then it will still be selected if the alternative belongs to the sub-menu. Formally,

Axiom 1(a) (Contraction Consistency (CC)). For all menus A and $x \in A$, we have

$$[x \in d(A)] \implies [x \in d(A') \; \forall \; \{x\} \subset A' \subset A]$$

The second axiom requires consistency in selection as the set expands.

Axiom 2(a) (Expansion (Exp)). For all menus A, B and $x \in A \cap B$, we have

$$[x \in d(A) \cap d(B)] \implies [x \in d(A \cup B)]$$

According to the axiom, an alternative that is selected in two menus must be selected in their union as well. It is also referred to as Condition γ in the literature.

LEMMA **1.1** (Sen (1971))**.** *A decision correspondence is an* MR *correspondence if and only if satisfies CC and Exp.*

The analysis of necessary and sufficient conditions for maximal-element rationalizability by a binary relation has been explored thoroughly. We refer the reader to Moulin (1985) and Bossert et al. (2005) for a comprehensive treatment of this issue for choice correspondences.

1.3 Acceptable Correspondences

In this section, we first introduce the notion of a minimal alternative. Further, we formally define and characterize *acceptable correspondences* using the set of minimal alternatives of a menu.

DEFINITION 1.2. For a menu A and rationale R, the set of **minimal** alternatives $m(A, R)$ is defined as

$$m(A, R) \coloneqq \{x \in A | \; \neg xRy \; \forall y \in A \text{ and } yRx \text{ for some } y \in A\}$$

An alternative is minimal if it is dominated by some alternative of the menu *and* does not dominate any alternative. We illustrate the notion of minimal set with an example.

Example 1.1. Let $A = \{a, b, c, d, e\}$. Rationales R_1, R_2, R_3 and R_4 are depicted in Panel (i), (ii), (iii), and (iv) respectively of Figure 1.1. Arrow from x to y depicts xRy. Else, x and y are non-comparable. In Panel (i), the set of minimal alternatives is $\{c, d\}$. This is

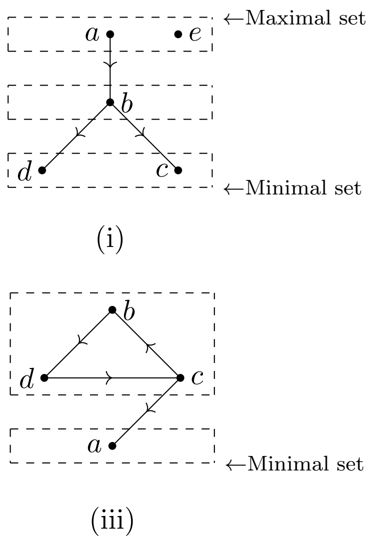

Figure 1.1: Rationales R_1, R_2, R_3, and R_4.

because both c and d are dominated by b, and do not dominate any other alternative in the set. Note that b is neither maximal nor minimal as it is dominated by a and dominate c and d. Alternative e is not minimal as it is not dominated by any other alternative. In Panel (ii), there are no minimal alternatives since each of the alternatives is either "isolated" or dominates at least one alternative in the set. The maximal set exists and is equal to $\{a, e\}$. In Panel (iii), there are no maximal alternatives, but the set of minimal alternatives is $\{a\}$. Lastly, in Panel (iv), there are no maximal or minimal alternatives as every alternative is both dominated and dominates some alternative in the menu.

This example illustrates that the set of maximal and minimal alternatives may be null in a menu. Observe that by definition, the set of minimal alternatives does not include any maximal alternative for a given rationale, that is, $M(A, R) \cap m(A, R) = \emptyset$[8].

For a menu A and rationale R, an alternative that does not belong to $m(A, R)$ is referred to as an *acceptable* alternative. We say that a decision correspondence is

[8]Our definition of minimal alternatives differs from the commonly known definition of minimality which states that an alternative is minimal if it does not strictly dominate any alternative. In that case, an alternative that is "isolated" in the menu is both a maximal as well as a minimal alternative. Therefore, we do not include such alternatives in our definition of minimal set. We thank Yves Sprumont for pointing this out to us.

acceptable if from each menu, (i) it selects a set of acceptable alternatives and (ii) it includes the maximal set if it exists. Formally,

Definition 1.3. A decision correspondence d is an **acceptable** correspondence if there exists a rationale R such that for all menus A,

$$M(A, R) \subseteq d(A) \subseteq A \setminus m(A, R)$$

If d is an acceptable correspondence and the underlying rationale is R, we say that "R generates d". The following observation is important.

Observation 1.1. Let d be an acceptable correspondence generated by a rationale R. Then, $R = R^d$ where (i) $d(\{x, y\}) = \{x\}$ if and only if $xR^d y$ and (ii) $d(\{x, y\}) = \{x, y\}$ iff $\neg xR^d y$ and $\neg yR^d x$.

The observation above says that if a rationale R generates an acceptable correspondence, then it must be unique. Also, it must be equal to the rationale revealed from the selection in binary menus of d. An example of a correspondence that is not acceptable is given below.

Example 1.2. Consider a decision correspondence d over the set $A = \{a, b, c\}$ as follows.

$$d(\{a, b\}) = \{a\} \qquad d(\{a, c\}) = \{a\} \qquad d(\{b, c\}) = \{b\} \qquad d(\{a, b, c\}) = \{a, c\}$$

We claim that d is not an acceptable correspondence. Suppose it is generated by rationale R. By Observation 1.1, we have aRb, bRc and aRc. Therefore, c is a minimal alternative according to R in the menu $\{a, b, c\}$, contradicting $c \in d(\{a, b, c\})$.

We now provide a characterization of acceptable correspondences. Our first axiom is a weakening of the CC condition. It places restrictions only on selection in binary menus.

Axiom 1(b) (Binary Contraction Consistency (BCC)). For all menus A and $x \in A$, we have

$$[x \in d(A)] \implies \big[x \in d(\{x, y\}) \ \forall\, y \in A \setminus \{x\}\big] \ \vee \ \big[\exists\, y \in A \setminus \{x\},\ d(\{x, y\}) = \{x\}\big]$$

This property has the following interpretation: if an alternative is selected in a menu, it must be either selected in *every* binary sub-menu, or it must be *uniquely selected* in some binary sub-menu. In the case of single-valued choice correspondences, this condition translates to a property called *Never Chosen*[9], which requires that if an alternative is not selected in every binary menu from a collection, then it must be rejected in their union. Such restriction on binary menus enables us to derive a rationale R that generates d such that for every menu A, $d(A) \subseteq A \setminus m(A, R)$ for some rationale R.

Our second axiom is a weakening of the Expansion condition. It requires that if an alternative is not rejected (is acceptable) in every binary menu from a collection, then it must be selected in their union as well. This property appears in Ehlers and Sprumont (2008) and is also known as *Always Chosen*[10]. The condition is formally defined below.

Axiom 2(b) (Binary Dominance Consistency (BDC)). For all menus A and $x \in A$, we have

$$[x \in d(\{x, y\}) \; \forall \; y \in A \setminus \{x\}] \implies [x \in d(A)]$$

This condition guarantees the existence of a rationale R that ensures $M(A, R) \subseteq d(A)$ for every menu A. Such a correspondence that contains an MR correspondence is referred to as *subrationalizable* (see Moulin (1985) and Echenique et al. (2011)). The independence of BCC and BDC is illustrated in the Example 1.3 below.

Example 1.3. Consider decision correspondences d over the set $X = \{a, b, c\}$ as follows

$$d(\{a, b\}) = \{a\} \qquad d(\{a, c\}) = \{a, c\} \qquad d(\{b, c\}) = \{b\} \qquad d(\{a, b, c\}) = \{b\}$$

This correspondence satisfies BCC but violates BDC as alternative a is selected in both $\{a, b\}$ and $\{a, c\}$. However, it is not selected in their union $\{a, b, c\}$. Now consider d in Example 1.2. It satisfies BDC, but not BCC as $c \in d(A)$ is not selected in $\{a, c\}$ and is also not selected uniquely in any other sub-menu.

The following straightforward result summarizes our discussion of this subsection.

Theorem 1.1. *A decision correspondence is an acceptable correspondence if and only if it satisfies BCC and BDC.*

Unlike MR correspondences, we do not need acyclicity of the rationale to guarantee the existence of an acceptable correspondence. The proof of this result is provided in Appendix 1.A.1.

[9] This property is used in characterizing two-stage model of Bhardwaj and Manocha (2021)

[10] It is the choice-rule formulation of what is sometimes called the Condorcet winner principle.

Note that MR correspondences, by definition, are a subclass of acceptable correspondences. However, not every sub-rationalizable[11] correspondence is acceptable. A trivial example of such a correspondence is the identity mapping $d^I : \mathcal{P}(X) \to \mathcal{P}(X)$ such that for all menus A, $d^I(A) = A$. Thus, the concept of acceptability strengthens the idea of sub-rationalizability to accommodate the rejection behaviour of the DM.

An important distinction between MR and acceptable correspondences is that the former satisfies a property of stability while the latter does not.

Definition 1.4. A correspondence d is stable if for all menus A,

$$d(A) = d(d(A))$$

Proposition 1.1. *MR correspondences are stable while acceptable correspondences are not.*

Stability of MR correspondences follows from the property of maximality. To show that acceptable correspondences are not stable, consider the example depicted in Figure 1.1 (Panel (i)). In the menu $A = \{a, b, c, d, e\}$, the set of acceptable alternatives is $\{a, b, e\}$. However, b is not an acceptable alternative in the smaller set $\{a, b, e\}$. Thus, $d(d(A)) \subsetneq d(A)$.

Note that, in this example $d(d(A)) = \{a, e\} = d(d(d(A)))$. In fact, an acceptable set is stable after a finite number of applications of the d operator. We will further investigate this behaviour of acceptable correspondences in Section 1.5.

1.4 Maximal Acceptable Correspondences

In this section, we characterize a natural class of acceptable correspondences that is, *maximal acceptable* correspondences. It selects the largest set of acceptable alternatives in the menu. We now formally define maximal acceptable correspondence.

Definition 1.5. A decision correspondence d is a **maximal acceptable** (MA) correspondence if there exists a rationale R such that for all menus A

$$d(A) = A \setminus m(R, A)$$

[11] A correspondence that contains a maximal-element rationalizable correspondence is referred to as a sub-rationalizable correspondence as defined in Moulin (1985).

We proceed to characterize the class of MA correspondences. Our first axiom is a mild weakening of the CC condition and a strengthening of the BCC condition. Unlike BCC, this puts restrictions on all the sub-menus considered. Formally,

Axiom 1(c) (Partial Contraction Consistency (PCC)). For all menus A and $x \in A$, we have

$$[x \in d(A)] \implies \big[x \in d(A') \; \forall \; \{x\} \subset A' \subset A\big] \;\vee\; \big[\exists \; y \in A \setminus \{x\}, \; d(\{x, y\}) = \{x\}\big]$$

According to the axiom, an alternative that is selected in a menu must be selected in *all* the sub-menus to which it belongs. Otherwise, it must be a *unique* choice in some binary menu.

The second axiom we propose is a strengthening of Expansion Consistency. Formally,

Axiom 2(c) (Strong Expansion (S-Exp)). For all menus A, B and $x \in A \cup B$, we have

$$\big[x \in d(A) \cap d(B)\big] \;\vee\; \big[d(A) = \{x\}\big] \implies [x \in d(A \cup B)]$$

The axiom has the following interpretation: consider two menus and an alternative x that belongs to the union of the two menus. If this alternative is selected in *both* menus, or *uniquely* selected in *one* of them, it must be selected in their union.

Our main result of this subsection is the following

Theorem 1.2. *A decision correspondence is an* MA *correspondence if and only if it satisfies PCC and S-Exp.*

The proof of the theorem can be found in Appendix 1.A.2.

1.4.1 Refinements of Maximal Acceptable Correspondences

In this subsection, we provide characterization results for MA correspondences obtained by imposing more structure on the rationale generating the correspondence. Note that no assumptions are made on the rationale in Theorem 1.2. Completeness and transitivity are standard rationality assumptions in the context of individual decision-making. In group decision-making, the Pareto relation is transitive (or quasi-transitive) but not complete. On the other hand, group decision-making via majority relations leading to *tournaments* is complete but not transitive.

We refer to an MA correspondence generated by an acyclic rationale as a $Y-$MA correspondence. Similarly, a correspondence generated by a transitive rationale is referred to as a $T-$MA correspondence, one generated by a complete rationale, a $C-$MA correspondence, and one generated by a linear order, a $TC-$MA correspondence. We introduce several axioms which in conjunction with the axioms deployed in Theorem 1.2 characterize $Y-$MA, $T-$MA, $C-$MA and $TC-$MA correspondences.

Axiom 3(a) (No Rejection Consistency (NRC)). For all menus A, we have

$$[d(A) = A] \implies [d(A') = A' \quad \forall A' \subset A]$$

According to NRC[12] if everything is acceptable in a menu, then everything must be acceptable in all its' sub-menus. This condition is necessary for the generating rationale to be acyclic and together with PCC and S-Exp is also sufficient.

Axiom 3(b) (No Binary Cycles (NBC)). For all $x, y, z \in X$, we have

$$[d(\{x, y\}) = \{x\}] \land [d(\{y, z\}) = \{y\}] \implies [d(\{x, z\}) = \{x\}]$$

The NBC axiom is well known in the literature (see Manzini and Mariotti (2007)). It is a straightforward restriction that guarantees that the generating rationale is transitive.

Axiom 3(c) (Resoluteness). For all menus A with $|A| = 2$, we have $d(A) \subsetneq A$.

The *resoluteness* axiom is introduced in Ehlers and Sprumont (2008). It requires the DM to be decisive in binary menus.

The following straightforward result characterizes $Y-$MA, $T-$MA, $C-$MA and $TC-$MA correspondences.

Theorem 1.3. *A decision correspondence is*

1. **Y**$-$MA *if and only if it satisfies PCC, S-Exp, and NRC.*

2. **T**$-$MA *if and only if it satisfies PCC, S-Exp, and NBC.*

[12] Sen (1971) introduces a property called δ which requires that if two alternatives are best in a menu S, then neither of them can be uniquely best in a larger menu T where $S \subset T$. Similarly, this says that if all the alternatives are best in a menu, then it cannot be that some alternatives are not the best in a smaller menu. A condition called *Aizerman* discussed in Moulin (1985) is in a similar vein and characterizes quasi-transitive rationalizability along with Contraction and Expansion Consistency.

3. **C**−MA *if and only if it satisfies PCC, S-Exp, and Resoluteness*[13].

Finally, we turn to MA correspondences that are generated by complete and acyclic rationales. Note that transitivity is implied by acyclicity in the presence of completeness and asymmetry. According to Theorem 1.3, the linearity of the generating rationale is guaranteed if the MA correspondence is both resolute and satisfies NRC (or NBC). Resoluteness along with NRC is equivalent to a stricter form of resoluteness, we call *strong resoluteness*[14].

Axiom 3(d) (Strong Resoluteness). For all menus A, we have $d(A) \subsetneq A$.

An alternate characterization can be provided using strong resoluteness and a CC condition for rejected alternatives.

Axiom 2(d) (Rejection Contraction Consistency (RCC)). For all menus A and $x \in A$,

$$[x \notin d(A)] \implies [x \notin d(A') \;\; \forall \; A' \subset A]$$

This condition is very restrictive as an alternative that is selected in at least one of a collection of menus must be selected in their union as well[15]. Observe that if the generating rationale is a linear order, the set of rejected alternatives is unique. The RCC condition can therefore be regarded as a counterpart of the CC condition for singleton-valued correspondence of rejected alternatives.

The relationship between Axioms **3(a)**, **3(b)**, **3(c)** and **3(d)** is established in Figure 1.2.

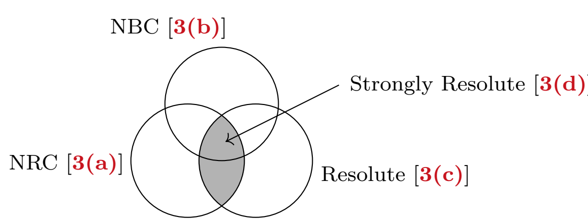

Figure 1.2: Relationship between Axioms **3(a)**, **3(b)**, **3(c)** and **3(d)**.

[13]A similar refinement result can be provided for acceptable correspondences, that follows from Observation 1.1. The implications of that result are subsumed in this result.

[14]This condition is introduced in Gerasimou (2016) under the name of *Choice implies Rejection*.

[15]The contrapositive of this statement suggests that it is a strengthening of Axiom **2(c)**.

Our final result in this subsection is stated below.

Theorem 1.4. *The following statements are equivalent. The correspondence d is*

1. **TC**−MA
2. *resolute and satisfies PCC, S-Exp, and NRC*
3. *resolute and satisfies PCC, S-Exp, and NBC*
4. *strongly resolute and satisfies PCC and S-Exp*
5. *strongly resolute and satisfies RCC*

The proof of Theorem 1.4 is provided in Appendix 1.A.2.

1.4.2 The Non-Equivalence of MR and MA Correspondences

It may be tempting to conclude that selection by maximization and selection by minimal rejection are "essentially" the same because one is the dual of the other. We argue in this section that it is not the case.

One possible interpretation of the equivalence claim is that an MA correspondence is an MR correspondence - what we obtain by rejecting minimal alternatives is precisely what we can obtain by maximizing, possibly different rationale. This claim is false which can be confirmed by referring to Example 1.4.

Example 1.4. Consider decision correspondences d over the set $X = \{a, b, c\}$ as follows

$$d(\{a,b\}) = \{a\} \qquad d(\{a,c\}) = \{a\} \qquad d(\{b,c\}) = \{b\} \qquad d(\{a,b,c\}) = \{a,b\}$$

The correspondence d is an MA correspondence generated by a linear order R : $aRbRc$. However, d violates CC since $b \in d(\{a,b,c\})$ and $b \notin \{a,b\}$. Therefore, it is not an MR correspondence. The falsity of the equivalence claim also follows the fact that the conditions that characterize MA and MR correspondences are distinct. The former is characterized by PCC and S-Exp and the latter by CC and Exp (Lemma 1.3 and Theorem 1.2). There is a special case when the equivalence holds. This is when MA correspondence is generated by a *dichotomous* rationale.

Definition 1.6. A rationale R is **dichotomous** if there exists a partition $\{G, B\}$ of X such that for all $x, y \in X$, xRy implies $x \in G$ and $y \in B$.

The equivalence result is stated below.

Proposition 1.2. *A decision correspondence d is an* MA *correspondence generated by a rationale R. Then the following statements are equivalent.*

1. *d is an* MR *correspondence.*

2. *R is dichotomous.*

Observation 1.1 provides an intuition of the result. A formal proof appears in Appendix 1.A.1.

An alternative interpretation of equivalence is that rejected alternatives are obtained by a maximizing procedure. In other words, the set of minimal alternatives according to a rationale is an MR correspondence for a possibly different rationale. Let d be an arbitrary correspondence and let $r^d(A) = A \setminus d(A)$ be the set of alternatives that are rejected while selecting $d(A)$. According to this notion of equivalence, there exists a rationale R' such that $r^d(A) = M(A, R')$ for every menu A. This claim is false as shown in Example 1.5 below.

Example 1.5. Consider decision correspondences d over the set $X = \{a, b, c\}$ as follows

$$d(\{a,b\}) = \{a\} \qquad d(\{a,c\}) = \{a,c\} \qquad d(\{b,c\}) = \{b,c\} \qquad d(\{a,b,c\}) = \{a,c\}$$

We note that d above is an MR correspondence generated by rationale R: aRb. Since $r^d(A)$ is empty for the menus $\{a, c\}$ and $\{b, c\}$, the equivalence does not hold. For r^d to be a decision correspondence, the condition of strong resoluteness is necessary. The unique case when equivalence in this sense holds is when d is a $TC-$MA correspondence.

Proposition 1.3. *A decision correspondence d is an* MA *correspondence generated by a rationale R. Then the following statements are equivalent.*

1. *r^d is an* MR *correspondence.*

2. *R is a linear order.*

The intuition of this result is the following: an alternative that is minimal in a menu A with respect to linear order R is the unique maximal alternative with respect to the linear order R^{-1}.

The notion of minimal alternatives defined in this chapter mirrors the concept of a *partially dominant correspondence* d^{PD} of Gerasimou (2016). Their correspondence is a refinement of MR correspondences. For every menu A and acyclic rationale R, $d^{PD}(A, R)$ consists of alternatives not dominated by any other alternative in A and dominate some alternatives according to R. If all alternatives are non-comparable in A, $d^{PD}(A, R) = A$. Example 1.5 also indicates that r^d associated with an MA correspondence is not a PD correspondence. However, the equivalence holds when the rationale is a linear order. This is reflected in Proposition 2 of Gerasimou (2016) that characterizes an MR correspondence using strong resoluteness and CC.

1.5 Threshold Acceptable Correspondences

In this section, we refine the class of acceptable correspondences by allowing the DM to reject more than the worst alternatives according to a rationale.

Definition 1.7. For all $k \in N$, menus A and rationales R, the **k−minimal** set $m^k(A, R)$ is defined recursively as follows:

1. $m^0(A, R) \coloneqq \emptyset$.
2. Given $m^{k-1}(A, R)$,

$$m^k(A, R) \coloneqq m^{k-1}(A, R) \cup m(A \setminus m^{k-1}(A, R), R)$$

The $k-$minimal set is obtained by successively removing the worst alternatives according to R. In the first step, the DM eliminates the worst alternatives, that is, the minimal set according to R. Then she removes the minimal set from the "survivor" set and so on for k steps. The example below illustrates the k-minimal set.

Example 1.6. Let $A = \{a, b, c, d, e, f, g\}$ and let R be a rationale shown in Figure 1.3.

An arrow from x to y signifies xRy. The rationale R is assumed to be transitive. For $k = 1$, the $m^1(A, R)$ is the set of minimal alternatives $m(A, R) = \{b\}$. For $k = 2$,

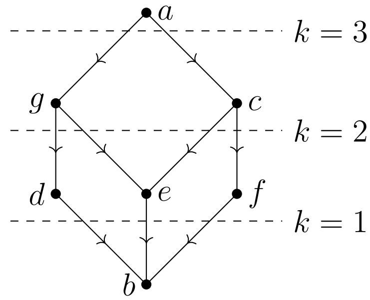

Figure 1.3: The rationale R

$m^2(A, R)$ is $\{b, d, e, f\}$ as this is the minimal set of the survivor set $\{a, c, d, e, f, g\}$. For $k \geq 3$, $m^k(A, R)$ is $A \setminus \{a\}$.

The integer k is referred to as a *threshold*. Note that in general, the threshold can depend on the menu and is specified by threshold function $k : \mathcal{P}(X) \to N$. Here $k(A)$ is the threshold for menu A.

It follows from the definition of $k-$minimal sets that for any menu A and rationale R, $m^k(A, R) = m^{k+1}(A, R)$ for some large enough k. In particular, this bound on k is less than $|A|$. We can therefore assume W.L.O.G that $k(A) \leq |A|$ for every menu A. We also assume that $k(A) \geq 1$ for all menus A to exclude consideration of trivial cases.

An important observation that follows immediately from the Definition 1.7 is that for any menu A and rationale R, $m^k(A, R) \cap M(A, R) = \emptyset$ for every $k \geq 1$.

Definition 1.8. A decision correspondence d is a **threshold acceptable** (TA) correspondence if there exists a transitive rationale R and a threshold function $k : \mathcal{P}(X) \to N$ such that for all menus A,

$$d(A) = A \setminus m^{k(A)}(A, R)$$

We assume transitivity on the generating rationale R for analytical simplicity. Our results can be extended to the acyclic case without generating any new significant insight. In view of our discussion, a DM can now be represented by a pair (R, k) where R is the rationale and k is the threshold function.

1.5.1 Fixed-TA Correspondences: Characterization

In this section we introduce and characterize TA correspondences where the threshold is menu-independent, that is, the threshold function is a constant function. This is consistent with a situation where the DM has a fixed processing capacity. Note that an MA correspondence discussed earlier is a special case of TA correspondence where $k(A) = 1$ for all the menus A. We will refer to such correspondences as *Fixed Threshold Correspondences* or fixed-TA correspondences.

The axioms that characterize fixed-TA correspondences are *Independence of Undominated Alternatives (IUA)*, *Expansion*, *NBC*, *Weakened CC (WnCC)* and *Neutrality*. We have already introduced Expansion and NBC in Sections 1.2.2 and 1.4.1 respectively. We proceed to formally introduce and discuss the other axioms.

Axiom 1(d) (Independence of Undominated Alternatives (IUA)). For all menus A and $x, y \in A$ such that $y \in d(\{x, y\})$, we have

$$[x \in d(A)] \implies [x \in d(A \setminus \{y\})]$$

Axiom 1(e) (Weakened Contraction Consistency (WnCC)). For all menus A and $x, y_1, y_2 \in A$ such that $d(\{y_1, y_2\}) = \{y_1, y_2\}$, we have

$$[x \in d(A)] \implies [x \in d(A \setminus \{y_i\}) \text{ for some } i \in \{1, 2\}]$$

The conditions IUA and WnCC are weakening of CC[16] condition that relates selections in larger menus and selections in appropriate sub-menus. Suppose alternative x is selected in menu A and y is "at least as strong as" x in the sense that y is selected in the binary menu $\{x, y\}$. Then the removal of y from the menu should make x "stronger" in the sub-menu. The IUA axiom requires x should continue to be selected in this sub-menu. However, this condition is stronger than PCC condition[17] (see Section 1.4).

The WnCC axiom considers a similar situation. Suppose there exist two alternatives

[16]The CC condition is also stated as Independence of Irrelevant Alternatives (IIA) in the literature. IIA says that for all menus A and $x, y \in A$ such that $x \neq y$, we have $x \in d(A) \implies x \in d(A \setminus \{y\})$. Axioms **1(d)** and **1(e)** are similar in spirit.

[17]Let d be correspondence and $X = \{a, b, c, d\}$ such that selections in binary menus are depicted in Figure 1.1, Panel (i). Let $d(X) = \{a, b, c\}$, $d(\{a, b, c\}) = \{a, b\} = d(\{a, b, d\})$, $d(\{a, c, d\}) = \{a, c\}$ and $d(\{b, c, d\}) = \{b\}$. This satisfies PCC but violates IUA. This is so because $d(X) = \{a, b, c\}$ and $c \notin d(\{b, c, d\}) = \{b\}$, when $d(\{a, c\}) = \{a\}$.

y_1 and y_2 in the menu which is equally strong relative to each other, that is, $d(\{y_1, y_2\}) = \{y_1, y_2\}$. Then the selection from the original menu must continue to be selected in one of the sub-menus where either y_1 or y_2 is removed. This condition does not require the x to be distinct from either y_1 or y_2. If $x = y_i$ for some $i \in \{1, 2\}$, then WnCC is implied by IUA. However, in the case when x, y_1, y_2 are distinct and $d(\{x, y_1\}) = d(\{x, y_2\}) = \{x\}$, then WnCC cannot be deduced from IUA. The relationship between the various CC axioms is discussed in Appendix 1.C.

The WnCC axiom is a modification of the Weak Contraction Consistency (WCC) condition that appears in Lahiri (2001), Ehlers and Sprumont (2008). According to WCC, a selection in a menu of size $T \geq 3$ must be the selection in at least one sub-menu of cardinality $T - 1$.

In order to define our last axiom, we introduce the concept of isomorphism between two menus for a given correspondence.

Definition 1.9 (**Isomorphic sets**). Let d be a correspondence. Menus A and B are isomorphic with respect to d if there exists a bijection $\sigma : A \to B$ such that for all $x, y \in A$, we have

$$[d\{x, y\} = \{x\}] \iff [d(\{\sigma(x), \sigma(y)\}) = \{\sigma(x)\}]$$

The notion of isomorphism only involves selections in binary menus. Consider a correspondence d and the menus $A = \{a, b, c, d\}$ and $B = \{b, c, e, f\}$ in Figure 1.4. Arrows depict selections in binary menus of the decision correspondence and dotted lines depict non-comparable pairs. For instance, $d(\{a, b\}) = \{a, b\}$ and $d(\{b, c\}) = \{b\}$. The menus $A = \{a, b, c, d\}$ and $B = \{b, c, e, f\}$ in Figure 1.4 are isomorphic with respect to the decision correspondence d for σ as,

$$\sigma(a) = b,\ \sigma(b) = e,\ \sigma(c) = f,\ \sigma(d) = c$$

Axiom 4 (Neutrality (N)). For all menus A, B and $x \in A$ such that A and B are isomorphic with respect to d with bijection σ, we have

$$[x \in d(A)] \implies [\sigma(x) \in d(B)]$$

The Neutrality axiom extends the selection behaviour of binary menus to the menus

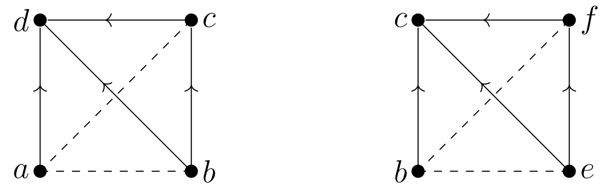

Figure 1.4: The Rationale R restricted to menus A and B.

themselves. This axiom requires that if a DM displays the same selection behaviour in the binary sub-menus of two menus, then she must display the same selection behaviour in the menus as well. Therefore, this precludes certain menu-dependent behaviour.

The independence of these axioms is illustrated in Appendix 1.B.1. Our main result in this section is the following:

Theorem 1.5. *A decision correspondence is a fixed-*TA *correspondence if and only if it satisfies IUA, WnCC, Exp, NBC, and N.*

The proof of Theorem 1.5 can be found in Appendix 1.A.2. Here, we provide a sketch of the proof which heavily relies on the notion of a *chain*. A chain is defined below.

DEFINITION 1.10 (**Chain**). Let d be a decision correspondence. Then, $\langle x_1, x_2, \ldots, x_n \rangle$ is a chain of length n if, $\forall\ i \in \{1, ..., n-1\}$, $d(\{x_i, x_{i+1}\}) = \{x_i\}$ and $d(\{x_i, x_j\}) = \{x_i\}$ implies $i < j$.

In Figure 1.4, $\langle b, c, d \rangle$ is a chain in menu $\{a, b, c, d\}$ of length 3 and $\langle e, f, b, c, d \rangle$ is a chain in menu $\{a, b, c, d, e, f\}$ of length 5. A menu is a chain if it comprises of all alternatives in the menu. Let R^d be a transitive rationale. We say that a "chain of length n below x" exists in menu A if for some $\{x_1, \ldots, x_n\} \subset A$, $\langle x, x_1, x_2, \ldots, x_n \rangle$ is a chain.

Sketch of the proof: In the necessity part of the proof of Theorem 1.5, the generating rationale R is recovered from the strict relation derived by selections in binary menus. In particular, $R = R^d$. It will be transitive due to NBC. In order to recover the threshold k, we identify the smallest cardinality menu, say A^*, such that there exists $x \in d(A^*)$ but x is not selected in some sub-menu to which it belongs. In other words, a violation of CC is observed. If such a set does not exist, then the threshold is set equal to $|X|$. In this case, d is actually an MR correspondence.

Suppose a menu A^* exists. By hypothesis, maximal alternatives with respect to R are selected for menus of cardinality smaller than $|A^*|$. This is proved using the characterization of MR correspondence (see Sen (1971)) since all the menus of cardinality lesser than $|A^*|$ satisfy Exp and CC conditions. In the next step, we prove that A^* is a chain. This is proved using IUA and WnCC, which together guarantee that only the largest chain below an alternative is relevant while selecting a non-maximal alternative. The threshold k is defined to be the cardinality of A^* minus the number of alternatives selected in this menu. In fact we show that this cardinality is equal to $|A^*| - 2$. This is so because exactly one non-maximal alternative is selected in this set. Since A^* is a chain, there is a chain of length $|A^*| - 2$ below it.

The result for general menus is established using induction on the cardinality of menus. The base case for menus with cardinality $|A^*|$ relies on an application of Neutrality. For larger menus A, the IUA condition ensures $k(A) \geq k$. On the other hand, the Exp condition ensures $k(A) \leq k$. Thus, k is fixed across all menus. $\square$

The characterization result in Theorem 1.5 is built on the characterization result of acceptable correspondences. Both the axioms, IUA and Exp are strengthening of BCC and BDC respectively (see Theorem 1.1). NBC ensures that the rationale generating d is transitive. (see Observation 1.1). Axioms WnCC and Neutrality impose restrictions on the correspondence so that the threshold in each menu is fixed.

This result is also related to the characterization result in Theorem 1.2 as $T-$MA correspondence is a special case of fixed-TA correspondence, with threshold $k = 1$. A Corollary of Theorem 1.5 gives an alternate characterization of $T-$MA correspondence.

Corollary 1.1. *A decision correspondence is a $T-$MA correspondence if and only if it satisfies IUA, S-Exp, and NBC.*

The S-Exp condition ensures $|A^*| = 3$. This is so for the following reason: for all $x, y \in X$, if $d(\{x, y\}) = \{x\}$, then for all $z \in X \setminus \{x, y\}$ it is true that $x \in d(\{x, y, z\})$. The other two axioms are redundant in this case.

1.5.2 TA Correspondences: Characterization

In this section, we characterize TA correspondences when the threshold may vary across menus. This includes situations where the threshold depends on factors like the size

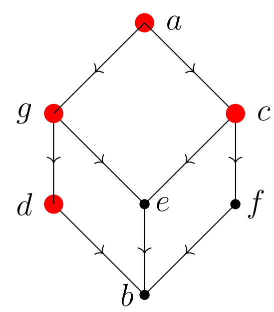

Figure 1.5: Violation of Monotonicity with $d(A) = \{a, c, d, g\}$

of the menu, complexity (in terms of selections in binary menus), number of desired selections from a menu, etc.

The characterization of TA correspondences is built on the characterization of acceptable correspondences (see Theorem 1.1.). In addition to BDC and BCC, we require a monotonicity axiom that gives a threshold structure to the correspondence. The intuition behind the monotonicity condition is the following: if an alternative is "stronger" than the other alternative in a menu, then rejection of the former implies rejection of the latter. The strength of an alternative is defined in terms of the length of the largest chain below the alternative. All maximal alternatives are the best alternatives in a menu. Thus, our monotonicity condition only compares the non-maximal alternatives.

For a decision correspondence d, let $\mathcal{L}(x, A)$ denote the set of chains below x in menu A with respect to d. For each menu A, we define a menu-specific revealed relation $\trianglerighteq_A$ using $\mathcal{L}(x, A)$ for every $x \in A$.

Definition 1.11. Let d be a decision correspondence. For all menus A and $x, y \in A$, we have

$$[\ x \trianglerighteq_A y\ ] \iff [\ |\mathcal{L}(x, A)| \geq |\mathcal{L}(y, A)|\ ]$$

where $|\mathcal{L}(x, A)|$ is the length of the longest chain in this collection.

The asymmetric and symmetric components of $\trianglerighteq_A$ are denoted by $\triangleright_A$ and $\sim_A$ respectively. It can be observed that $\trianglerighteq_A$ is complete and both $\triangleright_A$ and $\sim_A$ are transitive for any menu A. Thus, $\trianglerighteq_A$ is a menu-dependent weak order over the set of alternatives. A correspondence d with the strict relation derived by selections in binary sub-menus of the menu $A = \{a, b, c, d, e, f, g\}$ is depicted in Figure 1.3. The corresponding $\trianglerighteq_A$ is equal to the weak relation derived from selections in binary menus.

In view of the transitivity of R, it is legitimate to regard x being stronger than y if x and y belong to a chain and x is above y in the chain. If they do not belong to the same chain, then $x \sim_A y$ indicates that x and y are equivalent in terms of their strength. Likewise, $x \triangleright_A y$ indicates that there is an alternative z below x in the chain, and y is equivalent to z in terms of their strength. Note that the strength of an alternative in a menu is not proportional to the number of alternatives, or the number of chains that are below it. This is illustrated in Figure 1.6. Arrow from x to y depicts xRy. All other relations xRy are deduced by the transitivity of R. In Panel (i), there are 3 alternatives below b, and in Panel (ii), there are only 2 alternatives below it. Yet, b is stronger in Panel (ii) than in Panel (i).

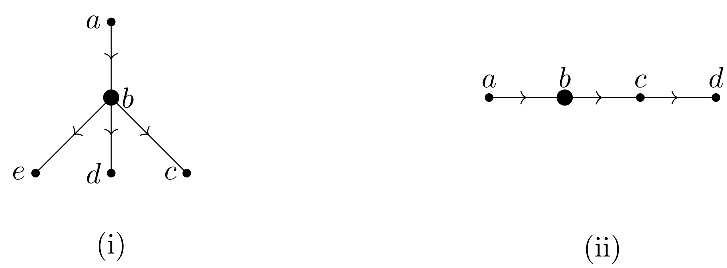

Figure 1.6: The Rationales R_1 and R_2.

The monotonicity axiom reflects a natural requirement: if a "weak" alternative y is selected in a menu, then a "stronger" alternative in the same menu must also be selected.

Axiom 5 (Monotonicity (MON)). For all menus A and $x, y \in A$ such that $d(\{y, z\}) = \{z\}$ for some $z \in A$ and $x \trianglerighteq_A y$, we have

$$[y \in d(A)] \implies [x \in d(A)]$$

This condition has an important implication: if $d(\{x, y\}) = \{x\}$, then $y \in d(A) \implies x \in d(A)$[18]. If R^d is transitive, the axiom ensures that every chain in the menu has an associated chain-specific threshold: all alternatives that are above the threshold are selected. This is illustrated in Figure 1.5.

For the case when x and y are not a part of any chain, MON ensures that an "internal neutrality" condition is satisfied. This is violated in the example in Figure 1.5.

The relationship between all the Contraction and Expansion Consistency axioms is

[18] This condition is referred to as *Condorcet Transitivity* in Moulin (1986) for a given tournament.

presented in Appendix 1.C.1. The following result is a characterization result of TA correspondences. The proof of this result is in Appendix 1.A.2.

Theorem 1.6. *A decision correspondence is a TA correspondence if and only if it satisfies BCC, BDC, NBC, and MON.*

1.6 Applications

The rejection approach developed in this chapter can offer explanations for various anomalies observed in decision theory. These anomalies include some common boundedly rational behaviours, such as the *attraction effect* and *compromise effect*[19]. Understanding these effects is essential for accurately predicting shifts in market shares resulting from the introduction or removal of products.

A succinct way to think about these phenomena is in terms of *choice reversals*. A decision correspondence d displays a *reversal* with respect to an alternative $x \in X$ if, for some $\{x\} \subset B \subset A$, we have $x \notin d(B)$ and $x \in d(A)$. The classical model of selection does not allow for any reversal and is characterized by the CC and Expansion conditions. However, in our model, we relax the assumption of the CC condition while still satisfying the Expansion condition, thereby accommodating reversals.

It should be noted that there exist several other models that offer explanations for such behaviour, including Manzini and Mariotti (2007), Masatlioglu and Nakajima (2007), Gerasimou (2016), Lleras et al. (2017), and many others. However, it is important to highlight that these models allow for *at most* one reversal with respect to a particular alternative. A decision correspondence d displays a *single reversal* with respect to an alternative $x \in X$ if, for some $\{x\} \subset B \subset A \subset A'$, $x \in d(B)$ and $x \notin d(A)$, then $x \notin d(A')$.

In contrast, certain empirically observed phenomena, such as the *two-decoy* effect or the *two-compromise* effect, involve what can be described as a double reversal (see Tserenjigmid (2019); Manzini and Mariotti (2010); Teppan and Felfernig (2009)). This behaviour can also be accommodated in our model. To better understand this concept, consider the following example:

[19]These effects, also referred to as the asymmetric dominance effect or the decoy effect, were initially identified in experimental studies conducted by Simonson (1989) and Huber et al. (1982).

Example 1.7. Let $X = \{a, b, c\}$. Selections in all the menus is as follows:

$$\begin{array}{rcc} d(\{a,b\}) = \{a\} & d(\{a,b,c\}) = \{a,b\} & \\ \boxed{d(\{b,d\}) = \{b,d\}} & \boxed{d(\{a,b,d\}) = \{a\}} & \boxed{d(\{a,b,c,d\}) = \{a,b,d\}} \\ d(\{a,c\}) = \{a\} & d(\{a,c,d\}) = \{a,d\} & \\ d(\{b,c\}) = \{b\} & d(\{b,c,d\}) = \{b,d\} & \\ d(\{a,d\}) = \{a\} & & \\ d(\{c,d\}) = \{d\} & & \end{array}$$

The correspondence presented above is an MA correspondence generated by the rationale $R : aRbRc$, $aRdRc$. It exhibits more than two reversals with respect to the alternative 'd' as the menus expand from $\{b, d\}$ to $\{a, b, d\}$ and then to $\{a, b, c, d\}$.

Interestingly, our model allows for more than two reversals. In general, a decision correspondence exhibits a *t-reversal* with respect to an alternative $x \in X$ if there exist subsets $S_1, S_2, \ldots, S_{k+1}$ such that $\{x\} \subset S_1 \subset S_2 \subset \ldots \subset S_{k+1}$, and it satisfies the following conditions: $x \notin d(S_1)$, and for all $i \in 2, \ldots, k+1$, $x \in d(S_i)$ if and only if $x \notin d(S_{i-1})$. The models in the papers cited previously are not able to accommodate this general behaviour. Example 1.8 illustrates that acceptable correspondences can allow for up to $|X| - 2$ reversals with respect to an alternative.

Example 1.8. Consider a TA correspondence d and $X = \{x_1, x_2, x_3, \ldots, x_n\}$. Let the generating rationale R be a strict linear order such that $x_i R x_{i+1}$ holds for all $i = 1, \ldots, n-1$. The threshold function $k(.)$ is the following:

$k(\{x_1, x_2\}) = k(\{x_1, x_2, x_3\}) = 1$
$k(\{x_1, x_2, x_3, x_4\}) = k(\{x_1, x_2, x_3, x_4, x_5\}) = 3$
$k(\{x_1, x_2, x_3, x_4, x_5, x_6\}) = k(\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}) = 5$
and so on..

This results in a t-reversal with respect to x_2 where $t = |X| - 2$.

It may be tempting to conclude that every non-standard behaviour can be explained by TA correspondences. However, this is not the case, as indicated by the characterization of TA correspondences (see Theorem 1.6). A smaller number of reversals can be accommodated as the threshold k increases in the case of fixed-TA correspondences.

Appendix 1.A

1.A.1 Proof of Proposition 1.2

Proof. Let d be an MA correspondence and rationale R be such that for all menus A, $d(A) = A \setminus m(A, R)$.

First, we show that if d be an MR correspondence, that is, for all menus A, $d(A) = M(A, R')$ for some acyclic rationale R', then R is dichotomous. Consider some $x, y \in X$. If $d(\{x, y\}) = \{x\} = M(A, R')$, then $xR'y$ holds. Also, by Observation 1.1, $d(\{x, y\}) = \{x\}$ implies xRy. Therefore, $R = R'$ and for all menus A, $d(A) = M(A, R) = A \setminus m(A, R)$. Consider a partition of X, $\{G, B\}$ where $G = M(X, R)$ and $B = A \setminus M(X, R)$. By definition, if xRy then $y \in B$. It remains to show that $x \in G$. If possible, let zRx holds for some $z \in X \setminus \{x, y\}$, that is, $x \notin M(X, R)$. Then, it must be that $x \in m(X, R)$, which is a contradiction as xRy is true. We conclude that $x \in G$ and thus, R is dichotomous.

Now, we prove that if R is dichotomous, then d is an MR correspondence. By hypothesis, for all menus A, $d(A) = A \setminus m(A, R)$ for some dichotomous rationale R. That is, there exists a partition of X, $\{G, B\}$ where xRy implies $x \in G$ and $y \in B$. We now show that for all menus A, $d(A) = M(A, R)$. Since R is acyclic, $M(A, R) \subseteq A \setminus m(A, R)$ by definition, it remains to $A \setminus m(A, R) \subseteq M(A, R)$. Consider $x \in A \setminus m(A, R)$ and suppose $x \notin M(A, R)$. Then, there exists $y, z \in X \setminus \{x\}$ such that $yRxRz$ holds. By definition of the partition, $x \in G \cap B$ which is a contradiction. Thus, d is an MR correspondence. $\square$

1.A.2 Proofs of Theorems

Proof of Theorem 1.1

Proof. Let us first prove the necessity of axioms. Let d be an acceptable correspondence generated by a rationale R. By Observation 1.1, $R = R^d$. BDC follows from $M(A, R) \subseteq d(A)$ and BCC follows from $d(A) \subseteq A \setminus m(A, R)$.

To prove the "if" part, let $R = R^d$. Note that, R is asymmetric by definition. Consider any arbitrary $A \in \mathcal{P}(X)$. Suppose $x \in M(A, R)$. Then $x \in d(\{x, y\})$ for all $y \in A$ and by BDC, we have $x \in d(A)$. Therefore, $M(A, R) \subseteq d(A)$. Now, suppose that $x \in m(A, R)$, that is, yRx for some $y \in A \setminus \{x\}$ and $\neg xRy$ for all $y \in A$. Since $x \notin d(\{x, y\})$ for all $y \in A\setminus$ and $y = d(\{x, y\})$ for some $y \in A \setminus \{x\}$, by BCC, we have $x \notin d(A)$ and we are done. □

Proof of Theorem 1.2

Proof. Let us first prove the necessity of axioms. Let d be an MA correspondence generated by a rationale R. By Observation 1.1, $R = R^d$.

− *PCC:* Consider a menu A and $x \in A$ such that $x \in d(A)$. As d is generated by R, $x \notin m(A, R)$. There are two possible cases: (i) $x \in M(A, R)$, then $x \in M(A', R)$ for all $A' \subset A$. This implies $x \in d(A')$ (ii) $x \in A \setminus (M(A, R) \cup m(A, R))$. By definition, there exists a $y \in A \setminus \{x\}$ such that xRy is true. This implies, $d(A') = \{x\}$ for for $A' = \{x, y\}$.

− *S-Exp:* Consider menus A, B and $x \in A \cup B$ such that (i) $x \in d(A) \cap d(B)$. As d is generated by R, $x \notin m(A, R) \cup m(B, R)$. Therefore, $x \notin m(A \cup B, R)$, which further implies that $x \in d(A \cup B)$. Now consider the case (ii) $d(A) = \{x\}$. As d is generated by R, for every $y \in A \setminus \{x\}$, xRy holds. Therefore, $x \notin m(A \cup B, R)$, further implying $x \in d(A \cup B)$.

Now we prove the "if" part. Suppose d satisfies PCC and S-Exp. Define $R = R^d$. Note that, R is asymmetric by definition. Define d_R as $d_R(A) = A \setminus m(A, R)$ for all A. It is clear that d_R is well defined. Now, we show that $d = d_R$. It is immediate that $d(A) = d_R(A)$ for all A such that $|A| \leq 2$. Consider an arbitrary $A \in \mathcal{P}(X)$ and $x \in d_R(A)$. There are two possible cases: (i) $\neg yRx$ for all $y \in A \setminus \{x\}$, that is, $x \in d(\{x, y\})$ for all $y \in A \setminus \{x\}$. By repeated application of PCC, we have $x \in d(A)$. (ii)yRx for some $y \in A \setminus \{x\}$. Since $x \notin m(A, R)$, we have xRy for some $y \in A \setminus \{x\}$, that is, $d(\{x, y\}) = x$. By the second part of S-Exp, again we get $x \in d(A)$. Therefore, we conclude that $d_R(A) \subset d(A)$. Now, suppose that $x \in d(A)$. By PCC, there are possible two cases: (i) $x \in d(A')$ for all $\{x\} \subset A' \subset A$ which implies $x \in d(\{x, y\})$ for all $y \in A \setminus \{x\}$ and by definition of R, we have $\neg yRx$ for all $y \in A \setminus \{x\}$. Therefore $x \notin m(A, R)$, implying $x \in d_R(A)$. (ii) $\{x\} = d(A')$ for some $A' \subset A$. If $|A'| = 2$, then we have xRy for $y \in A' \subset A$, implying $x \notin m(A, R)$ and hence $x \in d_R(A)$. Suppose $|A'| \geq 3$. Let $A' = \{x, y_1, \ldots, y_k\}$ for some $k \geq 2$. Observe that $d(\{y_i, y_j\}) = \{y_i, y_j\}$ for all $i, j \in \{1, \ldots, k\}$ (suppose not, then S-Exp would imply $d(A) \neq \{x\}$). Consider any arbitrary $y_i \in A'$. Suppose $y_i \in d(\{x, y_i\})$. Then S-Exp implies that $y \in d(A')$, a contradiction. Therefore, $\{x\} = d(\{x, y_i\})$ implying xRy. Hence $x \notin m(A, R)$ which implies $x \in d_R(A)$. So, we have $d(A) = d_R(A)$. Since A was arbitrary, we have shown that $d = d_R$. □

Proof of Theorem 1.3

Proof. We only prove (1.). Proofs of (2.) and (3.) are straightforward.

Following the proof of Theorem 1.2, it remains to prove that R^d is acyclic. We first show necessity the of NRC.

− *NRC:* Consider a menu $A \in \mathcal{P}(X)$ such that $d(A) = A$. As d is generated by an acyclic R, $x \notin m(A, R)$ for every $x \in A$. Now there are two possible cases (i) for every $x, y \in A$, $\neg xRy$. In this case, $A = M(A, R)$. Thus, $A' = M(A', R)$ for every $A' \subset A$. In case (ii) there exists a $x, y \in A$ such that xRy holds. Suppose $A = \{x_1, x_2, \ldots, x_n\}$ such that x_1Rx_2 is true. If $n = 2$, then there is a contradiction as $x_2 \in m(A, R)$. Thus, the only case possible is (i). If $n > 2$, then given that $x_2 \notin m(A, R)$, there is a $y \in A \backslash \{x_1, x_2\}$ such that x_2Ry is true. This is because if $y = x_1$, then R is not a rationale. W.L.O.G, let this y be x_3. If $n = 3$, then by the argument above, we have a contradiction. Suppose then $n > 3$. By similar argument, there is a $y \in A \setminus \{x_2, x_3\}$ such that x_3Ry holds. Observe that if $y = x_1$, then there is a cycle in R, again leading to a contradiction as R is acyclic. By repeated application of these arguments and finiteness of A, we reach a contradiction as there would exist at least one alternative belonging to the minimal set of A.

To prove the "if" part, let $R = R^d$. Suppose d satisfies PCC, S-Exp, and NRC. By PCC and S-Exp, $d_R = d$ where $d_R(A) = A \setminus m(A, R)$ for all $A' \in \mathcal{P}(X)$. Assume for contradiction that R is not acyclic. So, there exists $x_1, \ldots, x_n$ such that $x_1Rx_2R \ldots x_nRx_1$ for some $n \geq 3$ (the case $n = 2$ is ruled out by asymmetry of R). Consider $A = \{x_1, \ldots, x_n\}$. We have $d_R(A) = d(A) = A$ and $d(A') = A'$ for all $A \subset A$ by S-Exp. However $d(\{x_1, x_2\}) = x_1$, a contradiction. Therefore, R is acyclic. □

Proof of Theorem 1.4

Proof. (1.) $\implies$ (2.) follows from Theorem 1.3.
(2.) $\implies$ (3.) follows from the equivalence of acyclicity and transitivity under the completeness of a rationale.
To prove (3.) $\implies$ (4.), it is sufficient to prove the following lemma:

LEMMA **1.2.** *A decision correspondence satisfies resoluteness and NRC if and only if it satisfies strong resoluteness*

Proof. The "if" part is straightforward. Let us now prove the "only if" part. Consider a correspondence d that satisfies resoluteness and NRC. For the sake of contradiction, there exists a $A \in \mathcal{P}(X)$ such that $d(A) = A$. By resoluteness, $|A| > 2$. Now by NRC, for every $A' \subset A$, $d(A') = A'$. Thus, there exists a pair $\{x, y\} \subset A$ such that $d(\{x, y\}) = \{x, y\}$ which contradicts resoluteness. □

To prove (4.) $\implies$ (5.), it is sufficient to prove that if a correspondence d is strongly resolute, then PCC and S-Exp imply RCC. Consider an arbitrary $A \in \mathcal{P}(X)$ and some $x \in d(A)$. To show the contra-positive of RCC, we need to prove that $x \in d(A')$ for all $A' \supset A$. As d satisfies PCC and S-Exp, by arguments in the proof of Theorem 1.2, d is an MA and underlying rationale R is defined as xRy iff $d(\{x, y\}) = \{x\}$. Strong resoluteness implies $d(\{x, y\}) \neq \{x, y\}$ for all $x, y \in X$. Thus, we have R that is complete. Now, as $x \in d(A)$, it must be that $x \notin m(A, R)$. By completeness of R, there exists a $y \in A \setminus \{x\}$ such that xRy holds. Thus, we get $x \notin m(A', R)$ for all $A' \supset A$. As d is generated by R, this completes the proof.

Now we prove (5.) $\implies$ (1.).
Suppose d is strongly resolute and satisfies RCC. Let $R = R^d$. By definition, R is asymmetric. Since d is strongly resolute, it is complete as well. Assume for contradiction that it is not transitive. Then there exists x, y, z such that $xRyRzRx$. Note, however that $d(A) \subsetneq A$ for $A = \{x, y, z\}$. W.L.O.G. assume that $z \notin d(A)$. Then by RCC, we must have $z \notin d(\{x, z\})$ a contradiction to zRx. Therefore, R is transitive. Now, let d_R be defined as in the proof of Theorem 1.2. Note that, since R is a strict linear order $|d_R(A)| = |A| - 1$ for all $A \in \mathcal{P}(X)$. We will show that $|d(A)| = |A| - 1$ for all $A \in \mathcal{P}(X)$. Suppose not, then there exists A such that $|d(A)| \leq |A| - 2$, that is, there exists $x, y \in A$ such that $x \notin d(A)$ and $y \notin d(A)$. However, RCC implies $x \notin d(\{x, y\})$ and $y \notin d(\{x, y\})$, a contradiction to the assumption that d is non-empty valued. Note

that $d_R(A) = d(A)$ for all A with $|A| = 2$. Now, consider any arbitrary A and assume for contradiction that $d_R(A) \neq d(A)$. Suppose $x \notin d(A)$ and $x \in d_R(A)$. RCC implies that $x \notin d(\{x, y\})$ for all $y \in A \setminus \{x\}$. However, since $x \in d_R(A)$, we have xRy for some $y \in A \setminus \{x\}$, that is, $d(\{x, y\}) = x$, which is contradiction. Therefore $d_R = d$. $\square$

Proof of Theorem 1.5

Proof. We begin with the "only if" part.

Necessity: Let d be a fixed-TA correspondence generated by a partial order R and a constant threshold $k \in N$. By Observation 1.1, $R = R^d$. We now prove the necessity of the axioms.

– *IUA:* Consider a menu $A \in \mathcal{P}(X)$ and $x \in A$ such that $x \in d(A)$. As d is represented by (R, k), $x \notin m^k(A, R)$. If for some $y \in A \setminus \{x\}$ (i) $d(\{x, y\}) = \{y\}$, then there exists a k length chain $\langle x, x_1, x_2, \ldots, x_k\rangle$ in menu A. As R is transitive, $y \neq x_i$ for all $i \leq k$. As this chain also belongs to $A \setminus \{y\}$, $x \notin m^k(A \setminus \{y\}, R)$. Consider (ii) $d(\{x, y\}) = \{x, y\}$. If $x \in M(A, R)$, then we are done. If not, then there is a chain as in case (i). By transitivity of R, $\langle x, x_1, x_2, \ldots, x_k\rangle$ is a linear order, that is, $d(\{x, x_i\}) = \{x\}$ for all $i \leq k$. This implies $y \neq x_i$ for all $i \leq k$, further implying $x \notin m^k(A \setminus \{y\}, R)$. As $d(A) = A \setminus m^k(A, R)$, $x \in d(A \setminus \{y\})$ holds.

– *Exp:* Let $A, B \in \mathcal{P}(X)$ and $x \in A$ be such that $x \in d(A) \cap d(B)$. This implies $x \notin m^k(A, R)$ and $x \notin m^k(B, R)$. If $x \in M(A, R)$ and $x \in M(B, R)$, then $x \in M(A \cup B, R)$ and we are done. If x is not the maximal alternative in either of the menu, say $x \notin M(A, R)$, then there is a chain $\langle x, x_1, x_2, \ldots, x_k\rangle$ in menu A. As this chain exists in $A \cup B$ as well, $x \notin m^k(A \cup B, R)$ holds which implies $x \in d(A \cup B, R)$.

– *NBC:* This condition follows directly from transitivity of R.

– *WnCC:* Consider a menu $A \in \mathcal{P}(X)$ and $x \in A$ such that $x \in d(A)$. Let $y_1, y_2 \in A$ be such that $d(\{y_1, y_2\}) = \{y_1, y_2\}$. If $x \in \{y_1, y_2\}$ or $y_i \in d(\{x, y_i\})$ for some $i \in \{1, 2\}$, then we are done by the arguments in proving the necessity of IUA. Let $x \notin \{y_1, y_2\}$ and $x = d(\{x, y_i\})$ for all i. If $x \in M(A, R)$, then $x \in M(A \setminus \{y_i, R\})$ and we are done. Consider the case when $x \notin M(A, R)$. This implies there is a chain $\langle x, x_1, x_2, \ldots, x_k\rangle$ in menu A. As $d(\{y_1, y_2\}) = \{y_1, y_2\}$, atmost one of them belongs to this chain. If y_1, y_2 both do not belong to this chain, then the removal of any y_i assures the existence of this chain and hence $x \in d(A \setminus \{y_i\}, R)$. Let y_1 W.L.O.G belong to this chain. Then removal of y_2 assures the existence of this chain and as argued above, $x \in d(A \setminus \{y_2\}, R)$.

– *N:* As $m^k(A, R)$ is a the function of R restricted to a menu, N is satisfied. □

We now prove the "if" part in several steps.

Sufficiency: Consider d satisfying our axioms.

Step 1. First we define R and k. Let $R = R^d$. Note that R is asymmetric by construction and transitive by NBC. We need to prove that $d(A) = A \setminus m^k(A, R)$ for some $k \geq 1$. Before we define k, let us prove an intermediate lemma. We say a menu $A \in \mathcal{P}(X)$ satisfies property $\mathbb{P}$ if d_A satisfies property $\mathbb{P}$, where $d_A : \mathcal{P}(A) \to \mathcal{P}(A)$ is restriction of d on $\mathcal{P}(A)$.

Lemma **1.3.** *Consider a menu $A \in \mathcal{P}(X)$ that satisfies CC and Exp, then $d(A) = M(A, R)$.*

Proof. By [T9] of Sen (1971), d_A is normal as A satisfies CC and Exp.[20]. Hence, $d_A(A) = M(A, R_A)$ where R_A is the relation R^d restricted to a menu A. As $d(B) = d_A(B)$ for all $B \in \mathcal{P}(A)$, this concludes, $d(A) = M(A, R)$. □

Let $A^* \in \mathcal{P}(X)$ be the smallest cardinality menu such that there exists $x \in d(A^*)$ and $x \notin d(A)$ for some $\{x\} \subset A \subset A^*$, that is, there is a violation of CC. If there is no such A^*, then d satisfies CC and Exp, and by Lemma 1.3, $d(A) = M(A, R)$ for all $A \in \mathcal{P}(X)$. Note that for all $A \in \mathcal{P}(X)$ and $k' \geq 1$, $M(A, R) \subset A \setminus m^{k'}(A, R)$. Let $k = |X|$ in this case. Then, $M(A, R) = A \setminus m^{|X|}(A, R)$ holds.

Suppose such A^* exists. There may be multiple such menus. Pick anyone arbitrarily. Note that CC is satisfied for all $A \in \mathcal{P}(X)$ such that $|A| = 2$ and therefore $|A^*| > 2$. Define $k = |A^*| - 2$.

[20]Here a correspondence d is normal if for all $A \in \mathcal{P}(X)$, $d(A) = \{x \in A \mid \forall y \in A, \exists B_y \in \mathcal{P}(X)$ such that $x \in d(B_y)$ and $y \in B_y\}$. Condition [T9] is as follows: A correspondence is normal if and only if it satisfies CC and Exp.

Now, we show that for all $A \in \mathcal{P}(X)$ such that $|A| < |A^*|$, we have $d(A) = M(A,R) = A \setminus m^k(A,R)$. We consider two cases: (i) $|A^*| = 3$. In this case, since $k = 1$, it is straightforward to see that $d(A) = A \setminus m^k(A,R) = M(A,R)$. (ii) $|A^*| > 3$. In this case, CC and Exp is satisfied for all $A \in \mathcal{P}(X)$ such that $|A| < |A^*|$. By Lemma 1.3, it is true that $d(A) = M(A,R)$. Since at most $|A| - 1$ steps are required in successive elimination to reach the maximal set for any A and $k = |A^*| - 2$, we have $M(A,R) = A \setminus m^k(A,R)$.

Step 2. Before showing $d(A) = A \setminus m^k(A,R)$ for all A, we will first show that IUA and Exp imply WCC which is formally stated as follows.

Definition 1.12 (**Weak Contraction Consistency**(WCC)). For all menus A and $x \in A$, we have

$$x \in d(A) \implies x \in d(A \setminus \{y\}) \text{ for some } y \in A \setminus \{x\}$$

Lemma **1.4.** *A decision correspondence d satisfies IUA and EXP, then d satisfies WCC.*

Proof. Consider any arbitrary $A \in \mathcal{P}(X)$ and $x \in d(A)$. We have two cases possible cases: (i) Suppose $y \in d(\{x,y\})$ for some $y \in A \setminus \{x\}$. Then by IUA, we have $x \in d(A \setminus \{y\})$. (ii) $y \notin d(\{x,y\})$ for all $y \in A \setminus \{x\}$, that is, $\{x\} = d(\{x,y\})$. By Exp, we have $x \in d(B)$ for all $\{x\} \subset B \subset A$ and therefore d satisfies WCC. □

Now, we prove the following important lemma.

Lemma **1.5.** *A^* is a chain.*

Proof. Let $A^* = \{x, y_1, \ldots, y_{k+1}\}$ (recall $|A^*| = k+2$) such that $x \in d(A^*)$ and $x \notin d(B)$ for some $B \subset A^*$. Now, since d satisfies WCC, there exists $y_i \in A^* \setminus \{x\}$ such that $x \in d(A^* \setminus \{y_i\})$. By the previous step, we know that $x \in M(A^* \setminus \{y_i\}, R)$ implying $\neg y_j R x$ for all $j \neq i$. As $x \notin d(B)$ for some $B \subset A^*$, by the previous step, we have $x \notin M(B,R)$, that is, $y_l R x$ for some $l \in \{1, \ldots, k+1\}$. Also, $\neg y_j R x$ for all $j \neq i$ implies $l = i$. Since $y_i R x$, $x \notin M(A^* \setminus \{y_j\}, R)$ for all $j \neq i$ and by our previous step, we have $x \notin d(A^* \setminus \{y_j\})$. By IUA, we must have $y_j \notin d(\{x, y_j\})$, that is, $\{x\} = d(\{x, y_j\})$ for all $j \neq i$. Therefore, by definition of R, we have $x R y_j$ for all $j \neq i$. Now, consider any y_j and y_l with $j, l \neq i$. Suppose $d(\{y_j, y_l\}) = \{y_j, y_l\}$. Then, by WnCC, we must have $x \in d(A^* \setminus \{y_j\})$ or $x \in d(A^* \setminus \{y_l\})$, a contradiction. Therefore, we must have $d(\{y_j, y_k\}) \neq \{y_j, y_l\}$ implying $y_j R y_l$ or $y_l R y_j$. We have shown that R is complete on

$A^* \setminus \{y_i\}$. Since $y_i R x$, NBC implies that R is complete and transitive on A^*, that is, A^* is a chain. $\square$

Remark. Observe that since x was arbitrary, we have shown that only the top two alternatives will be selected in A^*. To see this, suppose $y_j \in d(A^*)$ for some $j \neq i$, then IUA implies that $y_j \in d(A^* \setminus \{y_i\})$ which is a contradiction since $y_j \notin M(A^* \setminus \{y_i\}, R)$. Further, we have $\{y_i\} = d(\{y_i, y_j\})$ for all $j \neq i$ and $\{y_i\} = d(\{y_i, x\})$ implying $y \in d(A^*)$ by Exp. Therefore $d(A) = \{x, y_i\}$.

Step 3. To show $d(A) = A \setminus m^k(A, R)$ for all A such that $|A| \geq |A^*|$, we use strong induction on $|A|$ with the base case, $k' = |A^*| = k + 2$. All $A \in \mathcal{P}(X)$ with cardinality $k + 2$ are of two types: (i) A satisfies CC, then by argument in Step 1, $d(A) = M(A, R) \subseteq A \setminus m^k(A, R)$. If A is a chain, given that $k = |A^*| - 2$, $A \setminus m^k(A, R)$ consists of top two alternatives of R. By Neutrality, $d(A) = A \setminus m^k(A, R)$ and thus CC is violated at A. Since A is not a chain, $d(A) = M(A, R) = A \setminus m^k(A, R)$ as each non-maximal alternative will have a chain less than size k below it (ii) A violates CC. Note that by Lemma 1.5, A is a chain of length $k + 2$. Neutrality implies that the top two alternatives of A will be selected in A, therefore we get $d(A) = A \setminus m^k(A, R)$

For the inductive step, consider any $k' \geq |A^*|$ and suppose that $d(A) = A \setminus m^k(A, R)$ for all A such that $|A| \leq k'$. Now, we will show that $d(A) = A \setminus m^k(A, R)$ for all A such that $|A| = k' + 1$.

Consider an arbitrary A such that $|A| = k' + 1$ and suppose $x \in d(A)$. Then WCC implies that $x \in d(A \setminus \{y\})$ for some $y \in A \setminus \{x\}$. By our inductive hypothesis, $x \notin m^k(A \setminus \{y\})$. There are two possible cases: (i) $x \notin M(A \setminus \{y\}, R)$, that is, there exists $z \in A \setminus \{y\}$ such that zRx and there exists a chain of length at least k below x implying $x \notin m^k(A, R)$. (ii) $x \in M(A \setminus \{y\}, R)$. Suppose $\neg yRx$, that is, $x \in d(\{x, y\})$, then, $x \in M(A, R)$ and implying $x \notin m^k(A, R)$. Now, suppose yRx. If there exist $z, z' \in A \setminus \{x, y\}$ such that $d(\{z, z'\}) = \{z, z'\}$, then by WnCC, we must have $x \in d(A \setminus \{z\}) \cup d(A \setminus \{z'\})$. W.L.O.G, let $x \in d(A \setminus \{z\})$. Since yRx, by our inductive hypothesis, we know that there exists a chain of length k below x and hence $x \notin m^k(A, R)$. Therefore, $d(\{z, z'\}) \neq \{z, z'\}$ for all $z, z' \in A \setminus \{x, y\}$ which implies $A \setminus \{x, y\}$ is a chain by NBC. If possible, there exists some $z \in A \setminus \{x, y\}$ such that $z \in d(\{x, z\})$. IUA then implies that $x \in d(A \setminus \{z\})$ and hence by our inductive hypothesis, since yRx, there exists a k length chain below x. Therefore, $\{x\} = d(\{x, z\})$ for all $z \in A \setminus \{x, y\}$. Since $A \setminus \{x, y\}$ is a chain, we have a

k long chain below x implying $x \notin m^k(A, R)$.

Now, to show the other direction, suppose $x \notin m^k(A, R)$ and assume for contradiction that $x \notin d(A)$. Note that $|A| \geq k+3$ (since $|A^*| = k+2$) and let $A = \{x, y_1, \ldots, y_p\}$ where $p \geq k+2$. Suppose, $x \in d(A \setminus \{y_i\}) \cap d(A \setminus \{y_j\})$ for some $y_i \neq y_j$, then by Exp, we have $x \in d(A)$. Therefore $x \in d(A \setminus \{y_i\})$ for at most one $y_i \in \{y_1, \ldots, y_p\}$. W.L.O.G, let $i = 1$. Then we must have $x \notin d(A \setminus \{y_j\})$ for all $j \in \{2, \ldots, p\}$. By our inductive hypothesis, $x \in m^k(A \setminus \{y_j\})$ for all $j \in \{2, \ldots, p\}$. Since $x \notin m^k(A, R)$, we have xRy_j for all $j \in \{2, \ldots, p\}$ and there is a k length chain below x. W.L.O.G, let the chain be $\langle x, y_2, \ldots, y_{k+1} \rangle$. Since $p \geq k+2$, there exists y_j with $j \in \{k+2, \ldots, p\}$ such that this chain is below x in $A \setminus \{y_j\}$. Thus, $x \notin m^k(A \setminus \{y_j\}, R)$. By our induction hypothesis, $x \in d(A \setminus \{y_j\})$ and since $j \neq 1$, we have a contradiction. Therefore, we have established that $x \in d(A)$ and the proof is complete. $\square$

$\square$

Proof of Theorem 1.6

Proof. The "only if" is straightforward. Condition BCC and BDC hold true as correspondence d is an acceptable correspondence (Theorem 1.1). NBC follows from the arguments in proof of Theorem 1.5. Let us now prove MON.

− *MON:* Consider a menu $A \in \mathcal{P}(X)$ and $x, y \in A$ such that $x \trianglerighteq_A y$ and $d(\{y, z\}) = \{z\}$ for some $z \in A$. As d is a TA correspondence, there exists threshold function $k(.)$ and partial order R such that $d(A) = A \setminus m^{k(A)}(A, R)$. $x \notin m^k(A, R)$. If $y \in d(A)$, then $y \notin m^{k(A)}(A, R)$. As $d(\{y, z\}) = \{z\}$ for some $z \in A$, by definition of R, it is true that zRy. This implies there exists at least $k(A)$ length chain below y in menu A. As the longest chain below x is larger than the longest chain below x in A, it must be that there also exists at least $k(A)$ length chain below x in menu A. Thus, $x \notin m^{k(A)}(A, R)$, further implying $x \in d(A)$.

We now show the "if" part. Define R as xRy if and only if $\{x\} = d(\{x, y\})$ and $k : \mathcal{P}(X) \to N$ as

$$k(A) = min\{max_{x \in A}|\mathcal{L}(x, A)|, 1\}$$

where $max_{x \in A}|\mathcal{L}(x, A)|$ denotes the length of the largest chain in A. If the largest chain is empty, then k is considered equal to 1. Relation R is asymmetric by construction and transitive by NBC. Now, we show that $d(A) = A \setminus m^{k(A)}(S, R)$. There are two possible cases: (i) $d(A) = A$. By BCC and NBC, it must be that $d(\{x, y\}) = \{x, y\}$ for all $x, y \in A$. It is straightforward to see that $d(A) = m^{k(A)}(A, R)$ for all $k(A)$. (ii) $d(A) \neq A$. Suppose $x \in d(A)$ and assume for contradiction that $x \in m^{k(A)}(A, R)$. Then there exists a $y \in A \setminus d(A)$ such that $y \trianglerighteq_A x$ and by MON, we must have $y \in d(A)$, a contradiction. Therefore $d(A) \subset A \setminus m^{k(A)}(A, R)$. Now, suppose $x \notin m^{k(A)}(A, R)$. If $\neg yRx$ for all $y \in A$, then $x \in d(A)$ by BDC. Suppose yRx for some $y \in A$. Assume for contradiction that $x \notin d(A)$. Then by the definition of $k(A)$, $x \in m^{k(A)}(R, A)$, a contradiction. $\square$

Appendix 1.B

1.B.1 Independence of axioms in Theorem 1.5

Example 1.9. (Exp, NBC, WnCC, N is satisfied, but IUA is violated)

Let $X = \{a, b, c\}$. Selections in all the menus is as follows:

$$d(\{a, b\}) = \{a\} \qquad d(\{a, c\}) = \{a\} \qquad d(\{b, c\}) = \{b\} \qquad d(\{a, b, c\}) = \{a, c\}$$

Condition IUA is violated as $c \in d(\{a, b, c\})$ and $b \in d(\{b, c\})$, but $c \notin d(\{a, c\})$.

Example 1.10. (IUA, NBC, WnCC, N is satisfied, but Exp is violated)

Let $X = \{a, b, c\}$. Selections in all the menus is as follows:

$$d(\{a, b\}) = \{a\} \qquad d(\{a, c\}) = \{a\} \qquad d(\{b, c\}) = \{b\} \qquad d(\{a, b, c\}) = \{b\}$$

Condition Exp is violated as $a \in d(\{a, b\}) \cap d(\{a, c\})$, but $a \notin d(\{a, b, c\})$.

Example 1.11. (IUA, Exp, WnCC, N is satisfied, but NBC is violated)

Let $X = \{a, b, c\}$. Selections in all the menus is as follows:

$$d(\{a, b\}) = \{a\} \qquad d(\{a, c\}) = \{c\} \qquad d(\{b, c\}) = \{b\} \qquad d(\{a, b, c\}) = \{a, b\}$$

Condition NBC is violated as $d(\{a, b\}) = \{a\}$ and $d(\{b, c\}) = \{b\}$, but $d(\{a, c\}) \neq \{a\}$.

Example 1.12. (IUA, Exp, NBC, N is satisfied, but WnCC is violated)

Let $X = \{a, b, c\}$. Selections in all the menus is as follows:

$$\begin{array}{llll} d(\{a,b\}) = \{a\} & d(\{a,c\}) = \{a\} & d(\{b,c\}) = \{b\} & d(\{a,b,c\}) = \{a\} \\ d(\{a,d\}) = \{a\} & d(\{b,d\}) = \{b\} & d(\{c,d\}) = \{c,d\} & d(\{b,c,d\}) = \{b\} \\ d(\{a,c,d\}) = \{a\} & d(\{a,b,d\}) = \{a\} & d(\{a,b,c,d\}) = \{a,b\} & \end{array}$$

Condition WnCC is violated as $b \in d(\{a,b,c,d\})$ and $d(\{c,d\}) = \{c,d\}$, but $b \notin d(\{a,b,c\}) \cup d(\{a,b,d\})$.

Example 1.13. (IUA, Exp, NBC, WnCC is satisfied, but N is violated)

Let $X = \{a, b, c\}$. Selections in all the menus is as follows:

$$\begin{array}{llll} d(\{a,b\}) = \{a\} & d(\{a,c\}) = \{a\} & d(\{b,c\}) = \{b\} & d(\{a,b,c\}) = \{a,b\} \\ d(\{a,d\}) = \{a\} & d(\{b,d\}) = \{b\} & d(\{c,d\}) = \{c,d\} & d(\{b,c,d\}) = \{b\} \\ d(\{a,c,d\}) = \{a\} & d(\{a,b,d\}) = \{a\} & d(\{a,b,c,d\}) = \{a,b\} & \end{array}$$

Condition WnCC is violated as $b \in d(\{a,b,c,d\})$ and $d(\{c,d\}) = \{c,d\}$, but $b \notin d(\{a,b,c\}) \cup d(\{a,b,d\})$.

Appendix 1.C

1.C.1 Relationship between classes of acceptable correspondences

We denote the class of acceptable correspondences by $\mathcal{A}$ (acceptable), $\mathcal{A}^k$ (TA), $\mathcal{A}^{\bar{k}}$ (fixed-TA), $\mathcal{A}^1$ (MA) and $\mathcal{A}^\star$ (MR). Figure 1.7 illustrates the set-inclusion property of all sub-classes of acceptable correspondences.

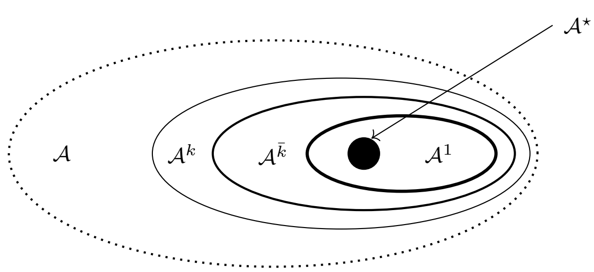

Figure 1.7

We summarize the characterization results presented in this chapter (for a transitive rationale) in the Table below. The condition NBC is satisfied for all the sub-classes.

	$\mathcal{A}^\star$	$\mathcal{A}^1$	$\mathcal{A}^{\bar{k}}$	$\mathcal{A}^k$	$\mathcal{A}$
Modified CC axioms	CC	PCC	IUA+WnCC	BCC	BCC
Modified Exp axioms	Exp	S-Exp	Exp	BDC	BDC
Other axioms			Neutrality	Mon	

The relationship between the properties characterizing these sub-classes is presented below. Arrow from Property X to Y indicates that X is a strengthening of Y.

(I) Relationship between modified Contraction Consistency axioms

$$\begin{tikzcd}
& \text{WnCC [1(e)]} & & \\
\text{CC [1(a)]} \arrow[ur] \arrow[dr] & & & \\
& \text{IUA [1(d)]} \arrow[r] & \text{PCC [1(c)]} \arrow[r] & \text{BCC [1(b)]}
\end{tikzcd}$$

(II) Relationship between modified Expansion Consistency axioms

$$\text{RCC [2(d)]} \longrightarrow \text{S-Exp [2(c)]} \longrightarrow \text{Exp [2(a)]} \longrightarrow \text{BDC [2(b)]}$$

Chapter 2

Sequential Matching with Affirmative Action[1]

2.1 Introduction

Recently, there has been a great deal of discussion on multi-round assignment procedures in college admissions in countries such as India, Brazil, China, France, and Germany. Baswana et al. (2019) proposed a semi-centralized, multi-round matching mechanism for engineering college admissions in India. They claim that the welfare of candidates improves in every round of the mechanism.

> "New semi-centralized, multi-period matching mechanism enjoys monotonicity across runs. The options available to candidates are only enhanced in going from one period to the next...."

However, they did not provide a rigorous theoretical justification for their claim. In this chapter, we provide one in the context of a general *sequential* college admissions problem.

The admission process for engineering colleges in India matches approximately 1.3 million students to 34,000 university positions. The Indian Institute of Technologies (henceforth, IITs) and the non-IIT Centrally Funded Technical Institutes (henceforth, non-IITs) have implemented the mechanism developed by Baswana et al. (2019). Before that, both types of institutions conducted their admission processes separately and independently. Each applicant received an offer from both types of institutions, resulting in a vacancy in one set of institutions when the candidate chose a program. These vacant seats were either left unfilled or allocated in an ad-hoc and decentralized manner, which caused

[1] Co-authored with Bertan Turhan (Department of Economics, Iowa State University, USA).

inefficiency and/or unfairness. Under this new combined seat allocation procedure, students are required to rank all programs (including both IITs and non-IITs) according to their preferences and submit a single ranking. Both types of colleges, IITs, and non-IITs, independently run the individual-proposing deferred acceptance (DA) mechanism introduced by Gale and Shapley (1962) to find a match. The assignment of students is determined based on their complete ranking. In the subsequent rounds of the admissions process, students are given various options to update their preferences. These options include "withdraw", "reject", "freeze", "slide", and "float". Depending on the option chosen, the student may either exit the procedure with or without an assignment or choose to participate in the subsequent rounds. After the completion of a fixed number of rounds, each student is then assigned to their finalized program.

This process is subject to a comprehensive affirmative action program, which has been implemented via a reservation system. There are two types of reservations: *vertical* and *horizontal*. Each institution reserves a certain percentage of its slots for students from a vertical reserved category–Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Classes (OBC), and Economically Weaker Sections (EWS). Specifically, 15%, 7.5%, 27%, and 10% of the slots are reserved for SC, ST, OBC, and EWS students, respectively. Applicants who do not belong to any of these vertical reserved categories are referred to as General Category (GC) and positions that are not reserved are referred to as open-category positions. They are available to all applicants, including those from reserved categories who do not declare their membership. A minimum number of positions within each vertical category are earmarked for women students as horizontal reservations. Vertical reservations are implemented "over-and-above" by filling open-category positions before vertically reserved categories, while horizontal reservations are implemented as a minimum guarantees by filling horizontally reserved positions before unreserved positions.

Situations such as the one described earlier can be modeled using *gradual matching mechanisms*, introduced in Haeringer and Iehlé (2021) (henceforth, H&I). They introduce a multi-round college admissions problem where individuals are offered repeated opportunities to participate in the mechanism, using updated preferences. The final matching is constructed *gradually* over several rounds. An important assumption made in the H&I model is that each institution has a *responsive*[2] choice rule. However, the underlying

[2]An institution has a *responsive* choice rule if it can be generated by a strict preference order that always selects the $q-$best alternatives whenever available. Here q is the capacity of the institution (see

assumption of responsiveness does not accommodate affirmative action considerations.

Motivated by engineering college admissions in India, we adopt the *gradual matching problem* of H&I. We assume that institutions implement affirmative action policies via the *slot-specific priorities* approach Kominers and Sönmez (2016). The latter is a model of the individual-institution matching model where each institution has a set of positions (slots) that can be assigned to different individuals. Positions have their own (potentially independent) rankings for contracts (here individuals). Within each institution, a linear order – referred to as the precedence order – determines the sequence in which positions are filled.

In this chapter, we investigate the restrictions on the preferences of individuals across different rounds that result in *monotone* outcomes. We refer to a matching outcome as monotone when each individual is matched to an institution that is weakly higher than the match of previous rounds (see Theorem 2.1). Further, we introduce a "backward-looking" notion of stability for sequential matching mechanisms that take into consideration individual rationality, non-wastefulness and justified envy of individuals across different rounds. Theorem 2.2 establishes a relationship between this notion of stability, we refer to as sequential stability, and monotone outcomes. These results generalize the findings presented in H&I. However, we do not rely on the proof of H&I to validate our first result.

We apply our theoretical findings to analyze the *multi-run multi-stage DA mechanism* that has been implemented in engineering college admissions in India since 2016. We also relate our findings to the to the characterization result of Kojima and Manea (2010) for the DA mechanism.

The layout of the chapter is as follows: the next subsection provides a brief literature review. Section 2.2 formally introduces the framework of thematching problem at each stage. Section 2.3 introduces a multi-round matching problem and Section 2.4 considers a class of multi-round matching problems, referred to as sequential matching problems. Section 2.5 introduces the stability notion for sequential matching problems. Section 2.6 provides an application of the model.

Chambers and Yenmez (2018)).

2.1.1 Related Literature

There is a vast literature dedicated to the study of dynamic matching problems. Bo and Hakimov (2022) introduce a family of *iterative deferred acceptance* mechanism, where the students are asked to sequentially make choices or submit partial rankings from sets of colleges. These are used to produce a tentative allocation at each step. If a student is unacceptable to their previous choice in some round, she is asked to make another choice among colleges that would tentatively accept her. Kesten (2010) study dynamic structure of deferred-acceptance algorithm in the form of rejection-cyles. The closest sequential model to ours is that of Haeringer and Iehlé (2021). We extend their framework to a more generalized setting where institutions' choice rules that have slot-specific priorities (SSP) choice rules. They offer a comprehensive review of dynamic matching models and emphasize how their sequential matching problem[3] differs from other dynamic models. While our results are broader in scope, they align closely with the findings presented in their work.

The SSP framework of Kominers and Sönmez (2016) provides a tool for market designers to handle diversity and affirmative action constraints in two-sided matching models. Aygün and Bó (2021) design SSP choice rules for the Brazilian college admission problem. More recently, Pathak et al. (2021) use the SSP framework to design a triage protocol for ventilator rationing. Avataneo and Turhan (2021) extend the framework to a more general one by defining SSP choice rules that allow transfers as in many real-world applications.

Our paper also contributes to the recently active literature on affirmative action in India from a market design perspective. Aygün and Turhan (2017) and Aygün and Turhan (2020) focus on IIT admissions and transferring vacant OBC positions to open-category. Similarly, Aygün and Turhan (2022) introduce a new transfer scheme with superior theoretical and practical properties. Aygün and Turhan (2023) offers another choice rule to implement affirmative action constraints and transfer vacant seats. This chapter discusses the joint implementation of vertical and horizontal reservations in engineering college admissions in India via position-specific priorities choice rules in a setting where applicants can update their preferences for additional rounds. Another related paper is Sönmez and Yenmez (2022), in which the authors study the allocation of government jobs in India and relate matching theory to Indian law. Unlike their

[3] They refer to such matching problems as *gradual matching problem*.

work, our chapter considers engineering college admissions in India. None of the above papers study the sequential implementation of individual-proposing deferred acceptance in a setting where individuals can update their preferences. Other papers studying affirmative action implementations include Echenique and Yenmez (2015), Kamada and Kojima (2015), Correa et al. (2021) among others.

This chapter is also related to the characterization results for the DA mechanism of Kojima and Manea (2010), Morrill (2013). Their main axiom is built on individually rational monotonic transformations (i.r.m.t) of a preference relation. A preference profile R' is an i.r.m.t of a preference profile R at an allocation, if for every individual, any object that is acceptable and preferred to this allocation under R' is preferred to the allocation under R. An outcome satisfies IR monotonicity if every individual weakly prefers the new allocation with respect to R' over the earlier allocation whenever R' is i.r.m.t of R. They show that the DA mechanism satisfies *IR monotonicity* when the choice rules are *substitutable* and *acceptant*. We utilize this condition in the sequential framework by referring to i.r.m.t of preference profile of active individuals in the sequential framework as *proposal-adhering rules*[4]. We prove that when the mechanism in each round is DA, then the outcome will be monotone.

Our sequential stability notion is the generalization of gradual stability introduced in H&I. When institutions' choice rules are responsive, our sequential stability reduces to gradual stability. A related stability concept was introduced in Pereyra (2013) in the context of seniority-based allocation rules. Feigenbaum et al. (2020) study the two-stage dynamic matching problem where a main round of admission is followed by a reassignment stage to fill vacancies.

2.2 Model

There is a finite set of institutions $S = \{s_1, ..., s_m\}$ and a finite set of individuals $I = \{i_1, ..., i_n\}$. Each individual $i \in I$ has an asymmetric and transitive preference relation P_i over $S \cup \{\emptyset\}$, where $\emptyset$ denotes remaining unmatched (henceforth a preference order). We write $sP_i\emptyset$ to mean that institution s is *acceptable* for individual i. Similarly, $\emptyset P_i s$ means institution s is *unacceptable* for individual i. We denote the profile of individual preferences by $P = (P_i)_{i \in \mathcal{I}}$. We let $\mathcal{P}$ denote the set of all strict preferences over $S \cup \emptyset$. We

[4] H&I refers to this property of preference update rule as regularity. However, they do not mention this comparison with Kojima and Manea (2010).

denote by R_i the weak preference relation associated with the strict preference relation P_i and by $R = (R_i)_{i \in \mathcal{I}}$ the profile weak preferences.

Institution s has $\overline{q}_s$ positions, and its selection criterion is summarized by a choice rule C_s, which selects a subset from any given set of individuals. That is, $C_s(I) \subseteq I$.

We let $\Xi = (I, S, (P_i)_{i \in I}, (C_s, \overline{q}_s)_{s \in S})$ denote a **stage problem**. A **stage matching** in a stage problem Ξ is a mapping $\mu : I \cup S \to 2^I \cup S$ such that, for each $i \in I$ and $s \in S$, (i) $\mu(i) \in S \cup \{\emptyset\}$, (ii) $\mu(s) \subseteq I$, and (iii) $\mu(i) = s$ if and only if $i \in \mu(s)$. A stage matching is **feasible** if $|\mu(s)| \leq \overline{q}_s$ for all $s \in S$.

Definition 2.1. A feasible stage matching μ is **stage stable** if for all $i \in I$ and $s \in S$,

1. *Individual rationality for individuals:* $\mu(i) R_i \emptyset$,
2. *Individual rationality for institutions:* $C_s(\mu(s)) = \mu(s)$, and
3. *Unblockedness:* $s P_i \mu(i)$ implies $i \notin C_s(\mu(s) \cup \{i\})$.

The first condition, *individual rationality* for individuals, guarantees that no individual is assigned to an institution they find unacceptable. The second condition, *individual rationality* for institutions, ensures that institutions' selection procedures are respected. This condition guarantees the implementation of affirmative action constraints when they are encoded into institutions' choice rules.[5] The last condition is the standard *no blocking pair* condition.

A **stage matching mechanism** φ maps every stage problem Ξ to a feasible stage matching μ. The mechanism φ is stable if $\varphi(\Xi)$ is stable for every stage problem.

2.2.1 Institutions' Choice Rules

We model institutions' selection criterion to accommodate affirmative action considerations via choice rules that have *slot (position)-specific priorities* (SSP) structure (Kominers and Sönmez (2016)). Institution s has a set of $\overline{q}_s$ positions denoted by $\mathcal{B}_s \equiv \{p_s^1, ..., p_s^{\overline{q}_s}\}$. Each position $p_s^j \in \mathcal{B}_s$ has a linear priority order $\succ_s^j$ over elements of $I \cup \{\emptyset_s\}$, where $\emptyset_s$ represents remaining unassigned and can be assigned to at most one individual. We denote the positions' priority order profile by $\succ_s = (\succ_s^j)_{j=1}^{j=\overline{q}_s}$. The positions in $\mathcal{B}_s$ are ordered according to a linear order of precedence $\rhd_s$ with the interpretation that if $p \rhd_s p'$, whenever possible, institution s fills position p before filling p'.

[5]See Alva and Doğan (2021) for in depth discussion of this point.

Given the priority order profile $\succ_s$ and the precedence order $\rhd_s$, the choice of institution s from a given set of individuals $A \subseteq I$, $C_s(A, \succ_s, \rhd_s)$, is given as follows:

- First, position p_s^1 is assigned to the individual who is $\succ_s^1$–maximal among the individuals in A. Call this individual i_1.
- Then, position p_s^2 is assigned to the individual who is $\succ_s^2$–maximal among the remaining individuals in $A \setminus \{i_1\}$. Call this individual i_2.
- This process continues with each position p_s^k is being assigned to the individual who is $\succ_s^k$–maximal among the remaining individuals in $A \setminus \{i_1, ..., i_{k-1}\}$.

If no individual is assigned to a position $p_s^l \in \mathcal{B}_s$, then p_s^l is assigned $\emptyset_s$. We say that a choice rule has an SSP structure if it is "generated" by a $(\succ_s, \rhd_s)$. A special case when $\succ_s^i = \succ_s^j$ for all positions $p_s^i, p_s^j \in \mathcal{B}_s$ is referred to as *responsive* preferences in the literature (see Chambers and Yenmez (2018)). This structure accommodates a variety of other constraints, including admission criteria in engineering colleges in India (see Section 2.6).

We deal with choice rules of institutions that have SSP structures. A stage problem can now be alternatively represented as

$$\Xi = (I, S, (P_i)_{i \in I}, (\succ_s, \rhd_s, \overline{q}_s)_{s \in S})$$

where $(\succ_s, \rhd_s, \overline{q}_s)$ encapsulates the choice structure of an institution s.

The SSP structure of the choice rule C_s enables us to define an associated matching of a feasible stage matching μ. The associated matching $\hat{\mu}$ maps each individual to an institution-position pair. For a given $(\succ_s, \rhd_s)$, the choice rule C_s assigns a position $p_s^j \in \mathcal{B}_s$ to each individual $i \in \mu(s)$. We denote this derived matching[6] as $\hat{\mu} : I \cup S \to 2^I \cup (S \times \mathcal{B})$ where $\mathcal{B} = \bigcup_{s \in S} \{\mathcal{B}_s\}$ is the collection of all the positions in the set of institutions S.

Definition 2.2. Slot-specific matching (SSM) of a stage matching μ is a mapping $\hat{\mu} : I \cup S \to 2^I \cup (S \times \mathcal{B})$ such that for all $s \in S$ and $i \in I$,

1. $i \in \hat{\mu}(s)$ if and only if $i \in \mu(s)$.
2. $\hat{\mu}(i) = (s, p)$ for some $p \in \mathcal{B}_s$ if and only if $\mu(i) = s$.

[6] The matching function $\hat{\mu}$ is different from the matching outcome $\tilde{\mu}$ of the associated one-to-one matching problem defined in Kominers and Sönmez (2016).

We illustrate with the example below that not every feasible matching μ can be associated with an SSM.

Example 2.1. Let $I = \{a, b, c, d\}$ and $S = \{s_1, s_2\}$ with capacities $\overline{q}_1 = \overline{q}_2 = 2$. Each institution with $\mathcal{B}_{s_1} = \{p_1^1, p_1^2\}$ and $\mathcal{B}_{s_2} = \{p_2^1, p_2^2\}$ has priority structure $\succ_1 = (\succ_1^1, \succ_1^2)$ and $\succ_2 = (\succ_2^1, \succ_2^2)$ respectively. The precedence order is $p_1^1 \rhd_{s_1} p_1^2$ and $p_2^1 \rhd_{s_2} p_2^2$.

$\succ_1^1$: $a - b - c - d$

$\succ_1^2$: $b - c - a - \emptyset_{s_1}$

$\succ_2^1$: $c - b - \emptyset_{s_2}$

$\succ_2^2$: $d - c - b - a$

Consider a feasible matching μ such that $\mu(s_1) = \{b, c\}$ and $\mu(s_2) = \{a, d\}$. For s_2, there is no associated $\hat{\mu}$ as the first position p_2^1 finds a and d unacceptable. Also, the second position p_2^2 prefers d over a. Thus, a cannot be associated with a position for this matching. However, the associated $\hat{\mu}$ for s_1 assigns b to p_1^1 and c to p_1^2.

Proposition 2.1. *Let μ be a stage matching of a stage problem $\Xi = (I, S, (P_i)_{i \in I}, (\succ_s, \rhd_s, \overline{q}_s)_{s \in S})$. Then μ has an associated SSM $\hat{\mu}$ if and only if μ is individually rational for institutions. Moreover, $\hat{\mu}$ is unique.*

For a matching μ such that $C_s(\mu(s)) = \mu_s$, each individual $i \in \mu(s)$ can be associated with a position in s. This is because each position finds the maximal individual of the surviving set acceptable. It can therefore be concluded that every stage stable matching μ has an SSM $\hat{\mu}$.

2.3 Multi-Period Matching with Preference and Priority Updates

We study a multi-period matching problem that consists of a sequence of stage problems $(\Xi^t)_{1 \leq t \leq T}$ where $\Xi^t = (I^t, S^t, (P_i^t)_{i \in I}, (\succ_s^t, \rhd_s^t, \overline{q}_s^t)_{s \in S})$ is the stage problem at stage t. The choice rule of an institution s has an SSP structure associated with the set of positions $\mathcal{B}_s^t$, profile of linear order $\succ_s^t$ and the precedence order $\rhd_s^t$.

Definition 2.3. A sequence of T stage problems Ξ^1, Ξ^2,..., Ξ^T is **nested** if for all $t = 1, ..., T-1$ and $s \in S$,

1. $I^{t+1} \subseteq I^t$

2. $S^t = S^{t+1} = S$
3. $\mathcal{B}_s^{t+1} \subseteq \mathcal{B}_s^t$

An individual $i \in I^t$ means that i is **active** at stage t. Individual $i \in I^t \setminus I^{t+1}$ $(t < T)$ means that i **finalize** her assignment at stage t before the final stage. We denote by $t_i := \arg\max_{1 \leq t \leq T}\{i \in I^t\}$, the stage at which individual i finalizes her assignment. In this round, the individual i is last active. Once the individual finalizes her assignment, she can not be active in further stages.

2.3.1 Updating Institutions' Choice Rules

For a feasible matching μ^t of stage problem Ξ^t at stage t, the institutions update their choice rule to accommodate the reduction in their capacity. The profile of linear order and precedence order $(\succ_s^t, \triangleright_s^t)$ is updated by removing "positions that are assigned to individuals" who finalize their assignments[7]. The idea is that when an individual finalizes her assignment, she leaves with the position she is assigned to. No other positions are added or removed in this process. The relative precedence order between two positions and the priority order over the set of active individuals remains unchanged.

Definition 2.4. A choice update rule is **consistent** if for all $s \in S$ and $t = 1, \dots, T-1$, given μ^t,

1. $\triangleright_s^{t+1} = \triangleright_s^t$.
2. For all $p \in \mathcal{B}_s^{t+1}$ and $i, j \in I^{t+1}$, $i(\succ_s^p)^t j$ implies $i(\succ_s^p)^{t+1} j$.
3. if $i \in I^t \setminus I^{t+1}$ and $\hat{\mu}^t(i) = (s, p)$ for some $p \in \mathcal{B}_s^t$, then $p \notin \mathcal{B}_s^{t+1}$.

Let us illustrate this choice update rule with an example.

Example 2.2. Consider an institution s with three positions p_1, p_2, and p_3 at stage $(t-1)$. That is, $\mathcal{B}_s^{t-1} = \{p_1, p_2, p_3\}$. The precedence order is such that $p_1 \triangleright_s p_2 \triangleright_s p_3$. The set of active individuals at $(t-1)$ stage is denoted by $I^{t-1} = \{i_1, i_2, i_3, i_4\}$. Priority

[7]Based on the discussion in Section 2.2, μ^t is individually rational for institutions to guarantee that each individual matched to an institution is assigned a position at the institution. When we are referring to a feasible matching generated by the SSP structure, this condition is assumed in the background.

orderings of the positions at stage $(t-1)$ are as follows:

$$(\succ_s^1)^{t-1} : \mathbf{i_1} - i_2 - i_3 - i_4 - \emptyset_s$$
$$(\succ_s^2)^{t-1} : \mathbf{i_2} - i_3 - \emptyset_s$$
$$(\succ_s^3)^{t-1} : i_1 - \mathbf{i_3} - \emptyset_s$$

Let the assigment at stage $(t-1)$ be the following: i_1, i_2, and i_3 are assigned to p_1, p_2, and p_3 respectively. Suppose that i_2 finalizes her assignment, but i_1 and i_3 decide to participate in the next stage. That is, $I^t = \{i_1, i_3, i_4\}$.

Then the choice rule is updated such that, at stage t, two positions are available: p_1 and p_3 with the precedence order $p_1 \rhd_s p_3$. The priority orders are updated as follows:

$$(\succ_s^1)^t : i_1 - i_3 - i_4 - \emptyset_s$$
$$(\succ_s^3)^t : i_1 - i_3 - \emptyset_s$$

It can be observed that if the choice rule of the institution is responsive, that is, the choice rule is generated by a preference order P_s, then updating the choice rule consistently reduces to the following: $P_s^t = P_s^1|_{I^t}$ where $= P_s^1|_{I^t}$ is the restriction of P_s^t to the set of active individuals I^t. As all the positions are homogenous, the updating rule is not dependent on the matching outcome of the previous period.

The second observation is that this rule implicitly puts a restriction on the relationship between the capacity of the institutions at every stage. The capacity of an institution is the number of unassigned seats plus the number of active individuals from the previous period. That is, for all $t \leq T-1$,

$$\overline{q}_s^{t+1} = (\overline{q}_s^t - |\mu^t(s)|) + |\{i \in I^{t+1} : \mu^t(i) = s\}|$$

This relationship between the capacities across stages is assumed explicitly in H&I. Thus, our choice update rule is more general than their preference-capacity update rule.

2.3.2 Updating Individual Preferences

A preference update rule maps a preference order P_i of individual i and an institution $s \in S$ that is assigned to her at stage t, to a set of permissible preference orders she can submit at stage $t+1$. In general, $\Gamma : \mathcal{P} \times (S \cup \{\emptyset\}) \to \mathcal{P}$ is a selection correspondence that maps a preference order and an institution s to a set of preference orders. Formally, for each $(P, s) \in \mathcal{P} \times (S \cup \{\emptyset\})$,

- If $s \in A_P$, then $A_{P'} \neq \emptyset$ for some $P' \in \Gamma(P, s)$;
- For each $P' \in \Gamma(P, s)$, we have $A_{P'} \subseteq A_P$.

Both conditions require that an institution unacceptable at previous stages can not be expressed as acceptable via preference update. There is a wider class of preference update rules that satisfy these mild conditions, including identity mapping[8], truncation mapping[9] and others. H&I provides a detailed discussion on this mapping, referred to as *refitting rules*. In the subsequent section, we discuss a practical application where the preference update rule of institutions satisfies these restrictions.

2.4 Sequential Matching Mechanisms

The multi-period matching problem that we are interested in has the following properties (i) the sequence of stage problems $(\Xi^t)_{1 \leq t \leq T}$ is nested, and (ii) choice rule of institutions is updated consistently. We refer to this class of multi-period matching problems as **sequential matching problems**[10].

Note that the nestedness of problems and consistency of institutions enables us to reduce this sequential problem to a simpler centralized framework. Instead of decentralized decisions of individuals to continue or finalize the match, it requires only the first stage problem Ξ^1 of the sequence $(\Xi^t)_{1 \leq t \leq T}$ and a list of preference orders $(P_i^1, P_i^2, \ldots, P_i^{t_i})$ for every $i \in I^1$ denoted by $\mathbf{P}_i = (P_i^t)_{t \leq t_i}$. We denote this reduced form by $\boldsymbol{\Xi} = (I^1, S, (\mathbf{P}_i)_{i \in I}, (\succ_s^1, \triangleright_s^1, \overline{q}_s^1)_{s \in S})$.

An outcome of a sequential matching problem is a sequence $(\Xi^t, \mu^t)_{t \leq T}$ that associates a feasible matching to every stage problem at every $t = 1, \ldots, T$. This outcome implicitly defines a matching $\nu : I \cup S \to 2^I \cup S$ such that $\nu(i) = \mu^{t_i}(i)$ for all $i \in I^1$. We refer to this sequence as a **sequential outcome**.

[8]An identity mapping is a singleton-valued selection correspondence such that for all $(P, s) \in \mathcal{P} \times (S \cup \{\emptyset\})$, $\Gamma(P, s) = \{P\}$.

[9]A truncation mapping is a singleton-valued selection correspondence discussed in Manjunath and Turhan (2016a) such that $\Gamma(P, s) = \{P'\}$ where $sPs' \implies \emptyset P' s'$ and $s'Ps''Ps \implies s'P's''P's$

[10]H&I refer to this class of multi-round matching problems as *gradual matching problems*.

For a given preference update rule Γ and a stage mechanism φ, a **sequential matching mechanism**, denoted by $\mathcal{M}_{\Gamma}^{\varphi}$, maps every sequential matching problem to a sequential outcome $\mathcal{M}_{\Gamma}^{\varphi}(\Xi) \equiv (\Xi^t, \mu^t)_{t \leq T}$ such that

- $\mu^t = \varphi(\Xi^t)$ for all $t = 1, \ldots, T$
- $P_i^t = \Gamma(P_i^{t-1}, \mu^{t-1}(i))$ for all $i \in I^1$ and $t = 2, \ldots, T$

We restrict our attention to stable stage mechanisms, that play a key role in the matching literature.

A desirable property of a sequential matching outcome is that for each period when an individual is active, she is assigned to an institution that is ranked weakly higher than the previous assignment. In effect, the current assignment is acceptable for individuals with respect to the *outside option*—assignment proposed in the previous period. We refer to such outcomes as *monotone* outcomes.

Definition 2.5. A sequential outcomes $(\Xi^t, \mu^t)_{t \leq T}$ is **monotone** if, for each $2 \leq t \leq T$ and $i \in I^t$,

$$\mu^t(i) R_i^t \mu^{t-1}(i).$$

A mechanism $\mathcal{M}_{\Gamma}^{\varphi}$ is monotone if for every sequential matching problem Ξ, the sequential outcome $(\Xi^t, \mu^t)_{t \leq T}$ is monotone. We illustrate the monotonicity of sequential outcome by Example 2.3 below.

Example 2.3. Consider a sequential matching problem with $T = 2$. For $t = 1$, let $I^1 = \{1, 2, 3, 4, 5, 6\}$ and $S = \{s_1, s_2\}$. Capacity of institutions is $q_{s_1}^1 = 3$, $q_{s_2}^1 = 2$. Preference profile of individuals, P^1 is as follows:

P_1^1	P_2^1	P_3^1	P_4^1	P_5^1	P_6^1
s_1	$\underline{s_1}$	$\underline{s_1}$	s_2	$\underline{s_2}$	s_2
$\underline{s_2}$	$\emptyset$	s_2	$\underline{s_1}$	$\emptyset$	s_1
$\emptyset$		$\emptyset$	$\emptyset$		$\underline{\emptyset}$

Suppose running a mechanism φ at stage one results in a matching μ^1 such that $\mu^1(s_1) = \{2, 3, 4\}$ and $\mu^1(s_2) = \{1, 5\}$. The assignment is depicted in the preference profile above.

Now, let $i = 2$ finalize its allocation at $t = 1$ and other individuals participate in the next period. The choice rule of institution s_1 is updated consistently. Thus,

$I^2 = \{1,3,4,5,6\}$ and $q^2_{s_1} = 2$, $q^2_{s_2} = 2$. Preference profile is updated to $P^2 = \Gamma(P^1, \mu^1)$ as follows:

P^2_1	P^2_3	P^2_4	P^2_5	P^2_6
s_1	s_2	s_2	$\underline{s_2}$	$\underline{s_1}$
$\underline{s_2}$	s_1	$\underline{s_1}$	$\emptyset$	s_2
$\emptyset$	$\underline{\emptyset}$	$\emptyset$		$\emptyset$

Now suppose running the mechanism φ at stage two results in the matching μ^2 such that $\mu^2(s_1) = \{4,6\}$ and $\mu^2(s_2) = \{1,5\}$ as underlined above. Note that for $i = 3$, outcome is not monotone as $s_1 = \mu^1(3)P^2_3\mu^2(3) = \emptyset$. However, for $i = 6$, the assignment is strictly better in the second period with respect to P^2_6.

The notion of monotonicity of the outcome defined above has a parallel interpretation in the static setting where the preference profile in the current period is a transformation of the previous period's preference profile[11]. Kojima and Manea (2010) introduce a notion of *individually rational monotonic transformation (i.r.m.t)* of a preference profile that guarantees the outcome to be monotone when the stage mechanism φ is the DA algorithm. A preference profile R is an *i.r.m.t* of R at $s \in S \cup \{\emptyset\}$ if for every individual i, any institution that is ranked above both s and $\emptyset$ under R'_i is ranked above s under R_i. In other words, each individual's set of acceptable institutions that are preferred to institution s shrinks. In the sequential matching setting, we refer to this property of preference update rule for an arbitrary proposal v as *proposal-adhering*. Formally,

Definition 2.6. The preference update rule Γ is **proposal-adhering** if for all $(P, v) \in \mathcal{P} \times (S \cup \{\emptyset\})$ and $s \in S \setminus \{v\}$, if $P' \in \Gamma(P, v)$ then,

$$sP'v \text{ and } sP'\emptyset \implies sPv$$

Example 2.4. In Example 2.3, Γ is not proposal-adhering with respect to assignment at stage one. That is because for $i = 3$, $s_2P^2\mu^1(3) = s_1$ and $s_2P^2\emptyset$. But, $s_1P^2s_2$. A proposal adhering preference profile P^2 is presented below.

[11] The monotonicity of sequential outcome is not identical to the static interpretation as the underlying stage problems in both the periods is not the same.

P_1^2	P_3^2	P_4^2	P_5^2	P_6^2
s_1	s_1	s_2	s_2	s_1
s_2	$\emptyset$	s_1	$\emptyset$	s_2
$\emptyset$		$\emptyset$		$\emptyset$

Our first main result of this chapter is the following

Theorem 2.1. *Let $\mathcal{M}_\Gamma^\varphi$ be the sequential mechanism where φ is individually-optimal stage stable mechanism (IOSSM) and Γ is preference-update rule. Then the following statements are equivalent:*

1. *$\mathcal{M}_\Gamma^\varphi$ is monotone.*

2. *Γ is proposal-adhering.*

Theorem 2.1 generalizes Theorem 1 of H&I. However, our proof does not depend on their result. Instead, we establish a connection of our result with the characterization result of the DA mechanism in Kojima and Manea (2010). This provides an alternative proof for Theorem 1 of H&I (see Appendix 2.A.1). A Corollary of Theorem 1 in Kojima and Manea (2010) is stated below.

Proposition 2.2. *Let φ be the deferred acceptance mechanism and C_s be an acceptant, substitutable choice rule. Then, φ satisfies **IR monotonicity**. That is, for all Ξ and proposal-adhering Γ,*

$$P_i' \in \Gamma(P_i, \mu(i)) \quad \Longrightarrow \quad \mu'(i) P_i' \mu(i) \quad \forall\ i \in I$$

Note that the existence of an IOSSM is not guaranteed in the case of a general choice rule. To address this, we rely on the results presented in Hatfield and Milgrom (2005) and Aygün and Sönmez (2013) to show the existence of such mechanisms when the choice rule has an SSP structure. Refer to Appendix 2.A.1 for the proof.

2.5 Sequential Stability

In this section, we introduce a stability notion for sequential matching problems. Our definition of sequential stability generalizes the notion of *gradual stability* introduced in H&I. When institutions' choice rules are responsive, our definition reduces to theirs.

DEFINITION 2.7. A sequential outcome $(\Xi^t, \mu^t)_{t \leq T}$ is **sequentially stable** if for all $i \in I^1$, $t' \leq t \leq t_i$,

1. *Individual Rationality:* $\mu^t(i) R_i^t \emptyset$ and $\mu^t(i) R_i^t \mu^{t'}(i)$,

2. *Non-wastefulness:* if $|\mu^{t'}(s)| < \overline{q}_s^{t'}$ for some $s \in S$, then $s P_i^t \mu^t(i)$ implies $\emptyset_s \succ_s^p i$ where $p \in \mathcal{B}_s^{t'}$ is unassigned at t',

3. *No justified envy:* for all $j \in I^{t'} \setminus I^{t'+1}$, if $\hat{\mu}^{t'}(j) = (s,p)$ for some $s \in S$, $p \in \mathcal{B}_s^{t'}$, then $s P_i^t \mu^t(i)$ implies $j \succ_s^p i$.

The stability of the stage matching mechanism considers only the final assignments of the individuals and institutions. However, in the sequential matching mechanism, individuals finalize their matchings at different stages. As a result, the definition above extends the notion of stage stability to also include the claims by individuals across stages as long as they are active.

The first condition requires that each individual's assignment must be *individually rational* for each individual at every stage. Here, individual rationality is defined by comparing an assignment with the "outside option". In the case of a stage problem, this outside option corresponds to being unmatched or having an empty matching. In the sequential problem, proposals from previous periods also serve as an outside option.

The second condition is a sequential version of *non-wastefulness.* If an individual i prefers institution s to her stage t assignment, this condition ensures that all positions in s that are deemed acceptable by individual i, must be assigned to someone else in the current period and all of the previous periods.

Finally, the last condition is a *no justified envy* condition adapted to our sequential environment with an SSP structure. Consider two individuals i and j such that j finalizes her assignment before i. Then, justified envy by i against j is checked for all periods $t_j \leq t \leq t_i$ using the preference order of the position that is assigned to j when she finalizes her assignment.

A sequential mechanism is sequentially stable if the outcome for each sequential problem is stable. The theorem we propose in this section shows a relationship between sequential stability, stage stability, and the monotonicity of sequential outcomes.

Theorem 2.2. *Let Γ be a proposal-adhering preference update rule and φ be a stage mechanism. Then the following statements are equivalent:*

1. $\mathcal{M}^{\varphi}_{\Gamma}$ *is sequentially stable.*

2. $\mathcal{M}^{\varphi}_{\Gamma}$ *is monotone and* φ *is stage stable.*

This result generalizes Theorem 2 in H&I to a more general setup where the choice rules have an SSP structure. The proof of this theorem can be found in Appendix 2.A.2.

2.6 Admissions to Engineering Colleges in India

The admission process in engineering colleges in India is subject to a comprehensive affirmative action program, which has been implemented via a reservation system. The reservation scheme at an institution s partitions the set of positions $\mathcal{B}_s$ in various categories $\mathcal{R}$ and the set of individuals in categories $\mathcal{C}$. It consists of the following key components:

- $\mathcal{R} = \{SC, ST, OBC, EWS\}$ denote the set of reserved categories. The students that belong to no reserved category are in General Category (GC).

- $\mathcal{C} = \{o, SC, ST, OBC, EWS\}$ denote the set of all position categories (o is the open category).

- The vector $q_s = \left(q_s^o, q_s^{SC}, q_s^{ST}, q_s^{OBC}, q_s^{EWS}\right)$ describes the *initial distribution* of positions over reserved categories where $q_s^o = \overline{q}_s - q_s^{SC} - q_s^{ST} - q_s^{OBC} - q_s^{EWS}$. The profile of vectors for the initial distribution of positions over categories at institutions is denoted by $\mathbf{q} = (q_s)_{s \in \mathcal{S}}$.

- The function $t : I \to \mathcal{R} \cup \{GC\}$ denotes the category membership of individuals. For every individual $i \in I$, $t(i)$, or t_i, denotes the category individual i belongs to. We denote a profile of reserved category membership by $T = (t_i)_{i \in \mathcal{I}}$, and let $\mathcal{T}$ be the set of all possible reserved category membership profiles.

Merit scores induce strict meritorious ranking of individuals at each institution s, denoted by $\succ_s$, which is a linear order over $\mathcal{I} \cup \{\emptyset\}$. $i \succ_s j$ means that applicant i has a higher priority (higher merit score) than applicant j at institution s. We write $i \succ_s \emptyset$ to say that applicant i is acceptable for institution s. Similarly, we write $\emptyset \succ_s i$ to say that applicant i is unacceptable for institution s. The profile of institutions' priorities is denoted $\succ = (\succ_{s_1}, ..., \succ_{s_m})$.

For each institution $s \in \mathcal{S}$, the merit ordering for individuals of type $r \in \mathcal{R}$, denoted by $\succ_s^r$, is obtained from $\succ_s$ in a straightforward manner as follows:

- for $i, j \in \mathcal{I}$ such that $t_i = r$, $t_j \neq r$, $i \succ_s \emptyset$, and $j \succ_s \emptyset$, we have $i \succ_s^r \emptyset \succ_s^r j$, where $\emptyset \succ_t^r j$ means individual j is unacceptable for category r at institution s.
- for any other $i, j \in \mathcal{I}$, $i \succ_s^r j$ if and only if $i \succ_s j$.

The *over-and-above* implementation requires filling open-category positions before the reserved categories. Formally, given an initial distribution of positions q_s, a set of applicants $A \subseteq I$, and a category membership profile $T \in \mathcal{T}$ for the members of A, the set of chosen applicants $C_s^{\text{Res}}(A, q_s)$, is computed as follows:

Step 1: Unreserved positions are considered first. Individuals are chosen one at a time following $\succ_s$ up to the capacity q_s^o. Let us call the set of chosen applicants $C_s^o(A, q_s^o)$.

Step 2: Among the remaining applicants $A' = A \backslash C_s^o(A, q_s^o)$, for each reserve category $r \in \mathcal{R}$, applicants are chosen one at a time following $\succ_s^r$ up to the capacity q_s^r. Let us call the set of chosen applicants for reserve category r as $C_s^r(A', q_s^r)$.

Then, $C_s^{Res}(A, q_s)$ is defined as the union of the set of applicants chosen in Steps 1 and Step 2. That is,

$$C_s^{Res}(A, q_s) = C_s^o(A, q_s^o) \cup \bigcup_{t \in \mathcal{R}} C_s^t\left(A', q_s^t\right)$$

This leads to our first lemma of this section, which we state formally below.

Lemma **2.1.** *The selection rule of engineering colleges in India can be modeled via choice rule, C_s^{Res} that admits SSP structure.*

2.6.1 The Multi-round Deferred Acceptance Mechanism

Admissions to engineering colleges in India implement a multi-round matching procedure. Each individual submits a preference list over all programs, including IITs and non-IITs. Each program provides the number of available positions and a merit list of eligible individuals. After collecting this information, the individual-proposing Deferred Acceptance mechanism is run in each round for IITs and non-IITs separately[12]. Some individuals may finalize their assignments, while others may want to participate in future rounds. In each round, the inputs of individual-proposing DA are modified. The options available to individuals at the end of each round are *freeze*, *float*, *slide*, *reject*, and *withdraw* (see Baswana et al. (2019)).

LEMMA **2.2.** *The permitted preference update rule is proposal-adhering.*

This lemma is validated by the description of the options available to the individuals.

Reject. If a candidate rejects an assigned program, then the candidate is completely removed from the assignment process by setting his/her rank-ordered list to an empty set.

$$\Gamma(P,s) = P', \text{ where } \emptyset P' s' \text{ for all } s' \in S.$$

Freeze. Individuals who choose this option accept the assigned program. Their preferences are modified so the assigned program and all other programs ranked below it are kept while the rest are removed. Formally,

$$\Gamma(P,s) = P'$$

where $\emptyset P' s$ for all $s' \in S$ such that $s' P s$, and P and P' agree for the rest.

Float. Preferences of the individuals who choose the float option remain unchanged. These candidates are willing to participate in future rounds in the hopes of getting assigned to a better program in their rank-ordered lists.

$$\Gamma(P,s) = P.$$

Slide. Individuals who choose the slide option want to participate the future rounds but be considered only for the programs in the same university as the assigned program. In this case, programs in all other universities above the assigned program are removed. The assigned program and other programs below it remain unchanged. Let $U(s)$ be the set of programs that are in the same university as program s.

$$\Gamma(P,s) = P'$$

[12]akin to the algorithm proposed by Manjunath and Turhan (2016a).

where $\emptyset P's$ for all $s' \notin U(s)$ such that $s'Ps$, and P and P' agree for the rest.

Withdraw. Individuals who choose this option are removed from the problem. The preferences of these applicants are set to an empty set.

$$\Gamma(P, s) = P'$$

where $\emptyset P's$ for all $s \in S$.

Note that the 'reject' and 'withdraw' options have the same effect on the rank-ordered lists. Their difference is about the timing. A candidate who previously accepts an offered program may withdraw in later stages.

Lemmas 2.1 and 2.2 suggest that we can use Theorems 2.1 and 2.2 to state our final result.

Proposition 2.3. *Multi-round deferred acceptance mechanism implemented in engineering colleges in India is monotone and sequentially stable.*

2.7 Conclusion

This chapter studies a special class of multi-round matching mechanisms. By generalizing the framework of H&I, we can explain a wider range of applications, including the college admissions process in engineering colleges in India. The French college admission system (studied in H&I) is another application that can be explained by our model.

One possible approach to understanding our results is to consider an associated one-to-one matching market that corresponds to the original many-to-one matching market. This technique is utilized in Kominers and Sönmez (2016), where positions, rather than institutions, compete for individuals. In the one-to-one market with unit capacity, the priority structure of institutions is responsive. Therefore, the results of H&I can be employed to understand the one-to-one market. However, these results do not straightforwardly extend to the original many-to-one market. In our study, we take an alternate approach to comprehend our results. Our first result to a great extent relies on the proof of the first characterization result of DA by Kojima and Manea (2010). This enables us to draw a comparison of their static problem with our sequential model.

Appendix 2.A

Before we begin the proof of Theorem 2.1, let us establish some properties of choice rules with an SSP structure.

Definition 2.8. A choice rule C satisfies *Irrelevance of Rejected Alternatives* if for every set $A, B \subset I$ such that $C(B) \subseteq A \subseteq B$, $C(A) = C(B)$.

Definition 2.9. A choice rule C satisfies *substitutability* if for every set $A, B \subset I$, $a \in A \subseteq B$, $a \in C(B) \implies a \in C(A)$.

Lemma **2.3.** *If C_s^t has an SSP structure, then it satisfies substitutability and IRA.*

The bilateral substitutes condition of Hatfield and Milgrom (2005) is equivalent to the substitutability condition defined above. As Kominers and Sönmez (2016) prove that SSP choice rules satisfy bilateral substitutes condition, this implies C_s^t satisfies substitutability. IRA is proved in Aygün and Sönmez (2013) for contracts setting.

Lemma **2.4.** *If C_s^t has an SSP structure and the choice update rule is consistent, then for all $A, B \subset I^t$,*

$$A \subseteq C_s^t(B) \implies A \cap I^{t+1} \subseteq C_s^{t+1}(B \cap I^{t+1})$$

Proof. Let C_s^t be a choice rule that is generated by $(\succ_s^t, \rhd_s^t)$ and C_s^{t+1} is updated as per Definition 2.4. Let $i \in A \cap I^{t+1}$ such that $\hat{\mu}^t(i) = (s, p_s^j)$ for some $p_s^j \in \mathcal{B}_s^t$. As i is active at stage $t+1$, $p_s^j \in \mathcal{B}_s^{t+1}$. Since i is $\succ_s^j$ maximal at stage t of the surviving set of individuals at Step j, it must be that i is $\succ_s^j$ maximal at stage $t+1$ as no individual is added to the set. Thus, $\hat{\mu}^{t+1}(i) = (s, p_s^k)$ for some $k \leq j$. This completes the proof. □

2.A.1 Proof of Theorem 2.1

Proof. Let $\boldsymbol{\Xi} = (I^1, S, (\mathbf{P}_i)_{i\in I}, (\succ_s^1, \rhd_s^1, \overline{q}_s^1)_{s\in S})$ be a sequential matching problem with the outcome $\mathcal{M}_\Gamma^{IOSSM}(\boldsymbol{\Xi}) \equiv (\Xi^t, \mu^t)_{1\leq t\leq T}$. We first prove that IOSSM exists when the choice rule has an SSP structure.

Lemma **2.5** (Hatfield and Milgrom (2005), Aygün and Sönmez (2013))**.** *When C_s satisfies substitutability and IRA, then IOSSM exists, and it is unique. Moreover, it is the outcome of the deferred acceptance algorithm.*

Lemma 2.3 and 2.5 prove the existence of IOSSM.

To prove that (1.) $\implies$ (2.), let $\mathcal{M}_\Gamma^{IOSSM}$ be monotone and Γ be not proposal-adhering. That is, there exists a Ξ such that for some $t \leq T-1$, and $i \in I^t$, $P_i^{t+1} \notin \Gamma(P_i^t, \mu^t(i))$. The counter-example provided in Proposition 1 of H&I suffices as responsive choice rules are a special case of choice rules with SSP structure where each institution is decomposed to identical multiple copies with unit demand.

Lemma **2.6** (Proposition 1, H&I)**.** *Let $\mathcal{M}_\Gamma^\varphi$ be a sequential matching mechanism such that $\succ_s^p = \succ_s^{p'}$ for all $s \in S$ and $p, p' \in \mathcal{B}_s$. If $\mathcal{M}_\Gamma^\varphi$ is monotone, then Γ is proposal-adhering.*

Following the methodology used in Kojima and Manea (2010), we now show (2.) $\implies$ (1.). Let Γ be a proposal-adhering preference update rule. Consider Ξ^t for some $t \geq 2$. Let $\mu^{t-1} = \varphi(\Xi^{t-1})$ and $\mu^t = \varphi(\Xi^t)$ be stage matchings at period $t-1$ and t respectively. We need to prove that for all $i \in I^t$, $\mu^t(i) R_i^t \mu^{t-1}(i)$. We prove this with the steps below.

Step 1: Define x_0 as the allocation μ^{t-1} restricted to I^t. That is, $x_0(s) = \mu^{t-1}(s) \cap I^t$.

Define for all $i \in I^t$,

$$x_1(i) = \begin{cases} x_0(i), & \text{if } x_0(i) P_i^t \emptyset, \\ \emptyset, & \text{otherwise.} \end{cases}$$

If x_1 is a stable matching at $(P_{I^t}^t, (C_s^t)_{s\in S})$, then using the fact that φ generates individual optimal stable matching, $\mu^t(i) R_i^t x_1(i)$ for all $i \in I^t$ holds and we are done.

Let us define a sequence $(x_k)_{k\geq 1}$ as follows:

Definition 2.10 (Step-wise unblocking process). Define for all $k \geq 1$ and $i \in I^t$,

$$x_{k+1}(i) = \begin{cases} s_k, & \text{if } i \in C_{s_k}^t \left(x_k(i) \cup \left\{ j \in I^t \mid s_k P_j^t x_k(j) \right\} \right), \\ x_k(i), & \text{otherwise.} \end{cases}$$

where s_k is an arbitrary institution that is part of a blocking pair if x_k can be blocked at $(P_{I^t}^t, (C_s^t)_{s\in S})$. If x_k cannot be blocked, then $x_{k+1} = x_k$.

We now prove the following lemma.

Lemma **2.7.** *The sequence $(x_k)_{k\geq 0}$ satisfies for every $k \geq 1$:*

(I) x_k is a feasible stage matching.

(II) $x_k(i) R_i^t x_{k-1}(i)$ *for all* $i \in I^t$*.*

(III) $x_k(s) \subseteq C_s^t \left(x_k(s) \cup \left\{ j \in I^t \mid s P_j^t x_k(j) \right\} \right)$ *for all* $s \in S$*.*

As $x_{k+1}(i) R_i^t x_k(i)$ for all $i \in I^t$, the step-wise unblocking process generates a sequence $(x_k)_{k \geq 0}$ that converges to a matching x_K in finite number of steps K. Each iteration results in a different allocation if the initial matching within the iteration is not stable. Hence, x_K is stable at $(P_{I^t}^t, (C_s^t)_{s \in S})$.

Because $x_K(i) R_i^t x_1(i) R_i^t \emptyset$ for all $i \in I^t$, the matching x_K is individually rational for agents. Also, as the outcome $\mathcal{M}_F^{SOSM}(\Xi)$ at time period t is the individual optimal among all the stable outcomes at t, we get for all $i \in I^t$,

$$\mu^t(i) R_i^t x_K(i) R_i^t \mu^{t-1}(i)$$

It remains to prove Lemma 2.7.

Proof. We prove the lemma by induction with the base case $k = 1$. (I) and (II) hold for $k = 1$ by definition of x_1. Consider any $s \in S$. We now prove that $x_1(s) \subseteq C_s^t \left(x_1(s) \cup \left\{ j \in I^t \mid s P_j^t x_1(j) \right\} \right)$. By definition of $x_1(i)$, we have

$$x_1(s) \subseteq x_0(s) \subseteq \mu^{t-1}(s)$$

As Γ is proposal-adhering,

$$\{j \in I^t \mid s P_j^t x_1(j)\} \subseteq \{j \in I^t \mid s P_j^{t-1} x_0(j)\}.$$

Together we get

$$x_1(s) \cup \{j \in I^t \mid s P_j^t x_1(j)\} \subseteq x_0(s) \cup \{j \in I^t \mid s P_j^{t-1} x_0(j)\} \tag{2.1}$$

As the period $t-1$ outcome is stable at $(P_{I^{t-1}}^{t-1}, (C_s^{t-1})_{s \in S})$, we have

1. $C_s^{t-1}(\mu^{t-1}(s)) = \mu^{t-1}(s)$ for all $s \in S$, and
2. $i \notin C_s^{t-1}(\mu^{t-1}(s) \cup \{i\})$, for all $i \in \{j \in I^t \mid s P_j^{t-1} x_0(j)\}$.

By consistency of C_s^{t-1}, we have

$$C_s^{t-1}(\mu^{t-1}(s) \cup \{j \in I^t \mid s P_j^{t-1} x_0(j)\}) = \mu^{t-1}(s).$$

In order to state the set-inclusion property at stage t, we state our next lemma, which is straightforward to prove. Using Lemma 2.4, we have

$$x_0(s) \subseteq C_s^t(x_0(s) \cup \{j \in I^t \mid sP_j^{t-1}x_0(j)\})$$

By substitutability of choice fuction C_s^t derived by updating structure of choice function,

$$x_1(s) \subseteq C_s^t\left(x_1(s) \cup \left\{j \in I^t \mid sP_j^t x_1(j)\right\}\right).$$

This concludes our proof for the base case.

Assuming the conclusions of step $k \geq 1$ hold, we now prove it for $k+1$ (the only case to prove is when $x_k \neq x_{k+1}$).

Let us prove (I) first. Consider $s \neq s_k$. Observe that $x_{k+1}(s) \subseteq x_k(s)$ by construction. As x_k is an allocation, by the inductive hypothesis we get $|x_{k+1}(s)| \leq |x_k(s)| \leq q_s^t$.

For institution s_k, $x_k(s_k) \subseteq C_{s_k}^t\left(x_k(s_k) \cup \left\{j \in I^t \mid s_k P_j^t x_k(j)\right\}\right)$ holds by inductive hypothesis at k. Then using definition of $x_{k+1}(i)$,

$$x_{k+1}(s_k) = C_{s_k}^t\left(x_k(s_k) \cup \left\{j \in I^t \mid s_k P_j^t x_k(j)\right\}\right)$$

Feasibility of choice rule $C_{s_k}^t$ thus guarantees that $|x_{k+1}(s_k)| \leq q_{s_k}^t$.

We now prove (II). Observe that

$$x_{k+1}(s_k) \backslash x_k(s_k) \subseteq \left\{j \in I^t \mid s_k P_j^t x_k(j)\right\} \tag{2.2}$$

Thus, for $j \in x_{k+1}(s_k) \backslash x_k(s_k)$, we get

$$s_k = x_{k+1}(j) P_j^t x_k(j) \tag{2.3}$$

Each agent outside of $x_{k+1}(s_k) \backslash x_k(s_k)$ is assigned the same institution under x_{k+1} and x_k. Therefore, $x_{k+1}(i) R_i^t x_k(i)$ for all $i \in I^t$.

We now show (III) for all $s \neq s_k$. By construction, we have $x_{k+1}(s) \subseteq x_k(s)$. By Equation 2.3, we have $\left\{j \in I^t \mid sP_j^t x_{k+1}(j)\right\} \subseteq \left\{j \in I^t \mid sP_j^t x_k(j)\right\}$. Therefore,

$$x_{k+1}(s) \cup \{j \in I^t \mid sP_j^t x_{k+1}(j)\} \subseteq x_k(s) \cup \{j \in I^t \mid sP_j^t x_k(j)\} \tag{2.4}$$

Now, substitutability of C_s^t, inductive hypothesis for k (condition (II)), and Equation 2.4 implies

$$x_{k+1}(s) \subseteq C_s^t \left(x_{k+1}(s) \cup \left\{ j \in I^t \mid s P_j^t x_{k+1}(j) \right\} \right).$$

Let us now consider institution s_k. By Equation 2.2, agents in $x_{k+1}(s_k) \backslash x_k(s_k)$ prefer s_k over their allocation in x_k. Individuals who are not chosen from this set in this iteration are those who still prefer s_k over their allocation in x_{k+1}. This is because $x_{k+1}(i) R_i^t x_k(i)$ for all $i \in I^t$. This implies

$$x_{k+1}(s_k) \backslash x_k(s_k) = \left\{ j \in I^t \mid s_k P_j^t x_k(j) \right\} \setminus \left\{ j \in I^t \mid s_k P_j^t x_{k+1}(j) \right\}.$$

Or equivalently,

$$x_{k+1}(s_k) \cup \left\{ j \in I^t \mid s_k P_j^t x_k(j) \right\} = x_k(s_k) \cup \left\{ j \in I^t \mid s_k P_j^t x_{k+1}(j) \right\} \tag{2.5}$$

Using substitutability of C_s^t, (III) for k and Equation 2.5, we obtain

$$x_{k+1}(s_k) \subseteq C_{s_k}^t \left(x_{k+1}(s_k) \cup \left\{ j \in I^t \mid s_k P_j^t x_{k+1}(j) \right\} \right).$$

This concludes our proof of Lemma 2.7. □

□

2.A.2 Proof of Theorem 2.2

We employ the technique utilized by H&I in the proof of Theorem 2. The notions of stage stability and sequential stability introduced in this chapter are generalized versions of spot stability and gradual stability, respectively, as defined by H&I. Thus, it remains to show the equivalence between Definition 2.1 and Definition 2.7 for an arbitrary stage t and $t = t'$ for the stage mechanism φ. We refer to the conditions of Definition 2.1 as C1, C2, and C3, and the conditions of Definition 2.7 as C1', C2', and C3', respectively.

We first show that Definition 2.7 implies Definition 2.1. Then, C1 directly follows from C1'. If possible, assume that C2 is not true. That is, there exists an $s \in S$ such that $\mu^t(s) \subsetneq C_s(\mu^t(s))$. As $|\mu^t(s)| \leq \overline{q}_s$, this implies $|C_s(\mu^t(s))| < \overline{q}_s$. Suppose $i \notin C_s(\mu^t(s))$. Since C_s is an SSP choice rule, at each position $p_k \in \left\{ p_1, \ldots, p_{\overline{q}_s} \right\}$, either (i) $\emptyset_s \succ_s^{p_k} i$ or (ii) there exists some other individual j such that $\hat{\mu}(j) = (s, p_k)$ and $j \succ_s^{p_k} i$. Both cases contradicts with our supposition that $(\Xi^t, \mu^t)_t$ is sequentially stable. Thus C2 is true.

If possible, assume that $C3$ is not true. That is, there exists an institution-individual pair (s, i) such that $sP_i^t\mu^t(i)$ and $i \in C_s^t(\mu^t(s) \cup \{i\})$. Let i be assigned to the position $p_s^k \in \mathcal{B}_s$. If p_s^k is unassigned at μ, then C2' is violated and if $\hat{\mu}(j) = (s, p_s^k)$, then C3' is violated. Thus, φ is stage stable.

We now prove that Definition 2.1 implies Definition 2.7. first, C1' follows from C1. If possible, C2' is violated. That is, there exists an institution s and individual i such that $sP_i^t\mu^t(i)$. Also, for some unassigned position $p \in \mathcal{B}_s$, $i \succ_s^p \emptyset_s$. This contradicts C3 as this implies $i \in C_s^t(\mu^t(s) \cup \{i\})$. We now prove C3' by contradiction. Consider $i, j \in I^t$ such that $\hat{\mu}^t(j) = (s, p)$, $\mu^t(i) \neq s$ and $i \succ_s^p j$ for some $p \in \mathcal{B}_s^t$. This implies $i \in C_s^t(\mu^t(s) \cup \{i\})$. If $sP_i^t\mu^t(i)$, C3 is violated. Thus, Definition 2.7 is true for $t = t'$.

Chapter 3

On the Integration of Shapley–Scarf Markets[1]

3.1 Introduction

Shapley–Scarf markets, in which agents own one house each which they can exchange among themselves without using monetary transfers, have been helpful to analyze several real-life allocation problems, such as the assignment of campus housing to students (Chen and Sönmez, 2002), house allocation with existing tenants (Abdulkadiroğlu and Sönmez, 1999) and kidney exchanges involving incompatible donor-patient pairs (Roth et al., 2004). A common complication in these allocation problems is that a big market is fragmented into several small and disjoint ones, causing inefficiencies. For example, house swaps in Australia are restricted to tenants within the same constituencies and community housing provider, blocking potentially beneficial exchanges (Powell et al., 2019). Similarly, most kidney exchanges in the US are conducted locally, despite the existence of centralized clearinghouses, which if used could increase the number of transplants by up to 63 percent (Agarwal et al., 2019).

Motivated by these observations, we investigate theoretically the welfare effects of integrating disjoint Shapley–Scarf markets. In our model, there are k Shapley–Scarf markets with n_j agents each (n_j is potentially different for each market) and n agents in total. The segregated allocation is obtained by treating each community separately and calculating the core allocation for each of them. The integrated allocation is the core allocation for the entire economy.

[1] Co-authored with Josué Ortega and Rajnish Kumar (Queen's Management School, Queen's University Belfast, UK). A version of this chapter has been previously published in the *Journal of Mathematical Economics* (2022).

Our first result (Proposition 3.1) states that up to, but not more than, $n-k$ agents may be harmed by integration, that is, they receive a house they prefer more when trade is only allowed within their own disjoint markets. This upper bound holds for any choice of n and k. It shows that Shapley–Scarf markets may fail to integrate because doing so could generate significantly more losers than winners.

Our second result (Proposition 3.2) concerns the size of the gains from integration in terms of house rank. For example, if an agent receives her 3rd best house before integration, but her 1st best after integration, the size of her gains from integration is $3-1=2$. Even if most agents are harmed by the merger of disjoint markets, integration may still be justified if the size of the gains from integration experienced by a few is substantially larger than the size of the losses from many. We show that, in the worst-case scenario, the size of the average gains from integration may be down to, but not less than, $\frac{-n^2+n+k^2+k}{2n^2}$. This lower bound can be achieved for any choice of n and k, and shows that, asymptotically, integration may increase the average house rank by 50% of the size of agents' preference lists.

Taken together, our first two results show that there are real obstacles to the integration of Shapley–Scarf markets. For example, if we have three small markets that merge into one with 60, 30, and 10 agents respectively, up to 97 agents may obtain a worse house after integration occurs, and on average (across all agents) each agent may receive a house 50 positions down on her preference list, equivalent to going from her top choice to her 51st choice.

However, these results are obtained in worst-case scenarios, which occur only when preferences are very specific. Consequently, studying the expected gains from integration across all possible preference profiles may be more informative. Therefore, our third result studies the size of the expected gains from integration in random Shapley–Scarf markets, in which agents' preferences over houses are drawn uniformly and independently.

In Proposition 3.3, we compute the exact expected gains from integration, which equal $\frac{(n+1)[(n_j+1)H_{n_j}-n_j]}{n_j(n_j+1)n} - \frac{(n+1)H_n-n}{n^2}$ (where H_n is the n-th harmonic number). This result shows that the expected welfare gains from integration are positive for all agents and larger for agents belonging to smaller markets. Going back to our example of three markets integrating with 60, 30 and 10 agents, the agents of the market with size ten go up 16 positions in their expected house rank, whereas those in the market of size sixty also increase their expected allocated house rank, but only by 2 rank positions. Our third result gives some context to our first two propositions, and shows that on average we should expect an overall positive effect from integration in Shapley–Scarf markets for agents from all disjoint markets.

Our fourth result (Proposition 3.4) establishes a connection between the number of trading cycles that occur in the top trading cycles algorithm and the expected number of agents harmed by integration. We use this connection to show that the expected number of agents harmed by integration in each economy is less than $n_j - \sqrt{2\pi n_j} - O(\log n_j)$, and consequently the expected number of agents harmed by integration in the entire economy is smaller than $n - \sqrt{2\pi}(\sum_{j=1}^{k} \sqrt{n_j}) - O(\log \prod_{j=1}^{k} n_j)$. In our example regarding the integration of markets with sizes 60, 30 and 10, our result implies that the expected number of agents harmed by integration is less than 44, 19, and 4 for each respective market. A consequence of our result is that, when all markets are of the same size, the expected fraction of agents harmed by integration is less than 50% whenever each market has less than $8\pi \approx 25.13$ agents.

A different approach to ensure that integration does not harm a majority of agents is to focus on specific preference domains. We find a preference domain that achieves this purpose, called sequential dual dictatorship, which enforces a particular correlation among agents' preferences. When preferences satisfy this property, we can guarantee that no more than 50% of agents in any individual market are harmed by integration (Proposition 3.5). The sequential dual dictator property is equivalent to assigning the title of a dictator to at most two agents at each step of the top trading cycle algorithm, therefore bounding the length of cycles that can occur.

Structure of the paper Next subsection 3.1.1 discusses the literature. Section 3.2 presents our model. Section 3.3 introduces a running example. Section 3.4 presents worst-case results. Section 3.5 discusses average-case results. Section 3.6 studies preference domains. Section 3.7 concludes.

3.1.1 Related Literature

A few other papers study the effects of integration on variations of Shapley–Scarf markets. For example, Ashlagi and Roth (2014) study the incentives for hospitals to fully reveal their patient–donor pairs to a centralized clearinghouse. In their model, agents do not have preferences but only dichotomous compatibility restrictions. Thus, welfare is measured by the size of the matching. They obtain worst- and average-case results that have a similar flavour to ours: the average-case cost for hospitals to fully integrate into a centralized clearinghouse is small, but the worst-case cost is high. In the same framework as them, Toulis and Parkes (2015) propose a mechanism that is efficient and asymptotically individually rational for hospitals. Our paper differs from the aforementioned articles in that we measure welfare in terms of how desirable the integrated allocation is with respect to the segregated one, rather than by the number of total exchanges (which is constant in the canonical Shapley–Scarf market that we consider where preferences are strict). Both welfare measures are relevant in different real-world settings and therefore we think of these two research strands as complementary.

Our work is also related to a series of recent articles that have studied the integration of other types of markets without money, in particular for Gale–Shapley one-to-one matching markets (Ortega, 2018, 2019), Gale–Shapley many-to-one matching markets with applications to school choice (Manjunath and Turhan, 2016b; Doğan and Yenmez, 2019; Ekmekci and Yenmez, 2019; Turhan, 2019; Aue et al., 2020), exchange economies (Chambers and Hayashi, 2017, 2020) and networking markets (Gersbach and Haller, 2021). Among these, the closest to ours are Ortega (2018, 2019). He shows that, in Gale–Shapley marriage markets, market integration never harms more agents than it benefits, even though the average rank of an agent's spouse can decrease by 37.5% of the length of the agents' preference list. He also provides an approximation for the gains from integration in random markets. Some of our results parallel his for Gale–Shapley marriage markets, although ours are more general as: i) they apply to the integration of markets of different sizes, ii) they provide tight bounds on the welfare losses, and iii) in the case of the gains from integration in random markets, our results are exact rather than approximations.

Our average-case results rely on two seminal papers from the computer science literature regarding random Shapley–Scarf markets with uniform and independent preferences. The first of these, by Frieze and Pittel (1995), computes the expected number of iterations that the top trading cycles algorithm takes to find the unique core allocation and the number of cycles created in the process. The second paper, by Knuth (1996), finds the expected sum of ranks of obtained houses and establishes the equivalence between the core allocation obtained from random endowments and the random serial dictatorship mechanism with no property rights.[2] Che and Tercieux (2019) use a similar random

[2]The latter result was also independently discovered by Abdulkadiroğlu and Sönmez (1998).

market approach to show that, in a related two-sided model, the top trading cycles algorithm achieves efficiency and stability asymptotically when agents' preferences are independent.

3.2 Model

Preliminary definitions We study the housing market proposed by Shapley and Scarf (1974), where there are n agents, each of them owning an indivisible good (say a house). The agents have strict ordinal preferences over all houses, including their own, and no agent has any use for more than one house.[3]

Formally, let $N \coloneqq \{1, \ldots, n\}$ be the set of agents and let $\omega \coloneqq \{\omega_1, \ldots, \omega_n\}$ be the initial endowment of the market. Let $\succ_i$ denote the strict preference of agent i and let $\succ \coloneqq (\succ_i)_{i \in N}$. The weak preference corresponding to $\succ_i$ is denoted by $\succcurlyeq_i$. A *housing market (HM)* is a pair $(N, \succ)$. An *allocation* $x = \{x_1, \ldots, x_n\}$ is any permutation of the initial endowment. That is, ω_i (resp. x_i) denotes the house endowed (resp. allocated) to agent i.

An allocation x is *individually rational* if $x_i \succcurlyeq_i \omega_i$ for all $i \in N$. An allocation x is a *core allocation* if there does not exist a coalition $S \subseteq N$ and an allocation y such that $\{y_i : i \in S\} = \{\omega_i : i \in S\}$ and $y_i \succ_i x_i$ for all $i \in S$. An allocation x is *Pareto optimal* if, for every alternative allocation x' such that $x'_i \succ_i x_i$ for some $i \in N$, there exists some $j \in N$ for which $x_j \succ_j x'_j$. A *matching mechanism* $\mathcal{M}$ is a map from HMs to allocations, and is said to be a core one (resp. individually rational, Pareto optimal) if it produces a core (resp. individually rational, Pareto optimal) allocation for every HM. The mechanism $\mathcal{M}$ is *strategy-proof* if, for every $i, \succ'_i, \succ$, $\mathcal{M}_i(N, \succ) \succcurlyeq_i \mathcal{M}_i(N, (\succ'_i, \succ_{-i}))$.

There is a unique core allocation (henceforth denoted by x^*) in every housing market. The unique core allocation can be found with an algorithm known as *top trading cycles* (TTC) (Shapley and Scarf, 1974; Roth and Postlewaite, 1977), which works by repeating the following two steps until all agents have been assigned a house.

1. Construct a graph with one vertex per agent. Each agent points to the owner of his top-ranked house among the remaining ones. At least one cycle exists and no

[3] We only consider the case where agents have strict preferences; for an analysis of housing markets with weak preferences, see Quint and Wako (2004); Alcalde-Unzu and Molis (2011); Aziz and De Keijzer (2012); Jaramillo and Manjunath (2012); Saban and Sethuraman (2013) and Aslan and Lainé (2020).

two cycles overlap. Select the cycles in this graph.

2. Permanently assign to each agent in a cycle the object owned by the agent he points to. Remove all agents and objects involved in a cycle from the problem.

TTC is the only mechanism satisfying individual rationality, Pareto-efficiency and strategy-proofness on the strict preference domain (Ma, 1994).

New definitions We study *extended housing markets* (EHM), which consist of an HM and a partition of the set of agents into k disjoint communities $C_1, \ldots, C_k$. That is, an EHM is a triple $(N, \succ, C)$, where $C \coloneqq \{C_1, \ldots, C_k\}$. An *integrated allocation* is any allocation for the HM $(N, \succ)$, whereas a *segregated allocation* is an allocation for $(N, \succ)$ in which every agent receives a house owned by an agent in her own community. That is, a segregated allocation x is such that $\{x_i : i \in S\} = \{\omega_i : i \in S\}\ \forall S \in C$. A matching scheme σ is a map from EHMs into a segregated and integrated allocation, denoted by $\sigma(\cdot, C)$ and $\sigma(\cdot, N)$, respectively.[4]

For agent's $i \in C_j$ preference $\succ_i$, we denote its restriction to C_j by $\widetilde{\succ}_i$. In other words, $\widetilde{\succ}_i$ is the strict ranking of agent i on all the houses belonging to agents in the community C_j (including his own) that is consistent with $\succ_i$. The matching scheme σ^* is the *core matching scheme* if $\sigma^*(\cdot, N)$ is the core matching for the HM $(N, \succ)$ and, for every community C_j, $\sigma^*(\cdot, C_j)$ is the core matching for the HM $(C_j, \widetilde{\succ}_{C_j})$, where $\widetilde{\succ}_{C_j} \coloneqq (\widetilde{\succ}_i)_{i \in C_j}$.

The *rank* of house ω_h in the preference order of agent i is defined by $\mathrm{rk}_i(\omega_h) \coloneqq |\{j \in N : \omega_j \succcurlyeq_i \omega_h\}|$. The *gains from integration* for agent i under the matching scheme σ are defined as $\gamma_i(\sigma) \coloneqq \mathrm{rk}_i(\sigma(i, C)) - \mathrm{rk}_i(\sigma(i, N))$. The *total gains from integration* is given by $\Gamma(\sigma) \coloneqq \sum_{i \in N} \gamma_i$. If these are negative, we speak of the total losses from integration. The *average percentile gains from integration* are denoted by $\overline{\Gamma}(\sigma) \coloneqq \frac{\Gamma(\sigma)}{n^2}$. We divide by n^2 to account for both the number of agents (n) and the length of an agent's preference list (which is also n). Thus, $\overline{\Gamma}(\sigma) \in (-1, 1)$, where $\overline{\Gamma}(\sigma) = -1$ means that everybody was harmed by integration and moved from their best possible house to the worst possible one.

We use $N^+(\sigma) \coloneqq \{i \in N : \sigma(i, N) \succ_i \sigma(i, C)\}$ to denote the set of agents who benefit from integration. Similarly, $N^0(\sigma) \coloneqq \{i \in N : \sigma(i, N) = \sigma(i, C)\}$ and $N^-(\sigma) \coloneqq \{i \in$

[4] Matching schemes are similar to the concept of assignment schemes in cooperative game theory (Sprumont, 1990).

$N : \sigma(i,C) \succ_i \sigma(i,N)\}$ denote the set of agents that are unaffected and harmed by integration, respectively. For all $j \in \{1,\ldots,k\}$, we define $N^+_{C_j}(\sigma) \coloneqq \{i \in C_j : \sigma(i,N) \succ_i \sigma(i,C)\}$ to be the set of agents in community C_j who benefit from integration. The sets $N^0_{C_j}(\sigma)$ and $N^-_{C_j}(\sigma)$ are defined analogously.

Henceforth we focus on σ^*, that is, we study the gains from integration that occur when the allocation obtained before and after integration occurs is the unique core allocation.

3.3 Running Example

Example 3.1 presents an EHM that we will use throughout the paper to illustrate how market integration may harm the majority of agents, and the welfare losses of such agents can be significant. In this EHM, $n = 7$ and $k = 2$ with $C_1 = \{a,b,c\}$ and $C_2 = \{d,e,f,g\}$. The integrated (resp. segregated) core allocation appears in a diamond (resp. circle).

Table 3.1: An EHM with $C_1 = \{a,b,c\}$ and $C_2 = \{d,e,f,g\}$.

a	b	c	d	e	f	g
d	a	b	a	d	e	f
c	d	a	g	a	a	a
⋮	b	d	⋮	b	d	d
	⋮	c		c	b	b
				e	c	c
				⋮	f	e
					g	g

In Example 3.1, there are two communities with three and four agents each, such that one agent from each community (in this case a and d) is assigned to their second best house in the segregated core allocation, whereas all remaining agents are assigned to their most preferred house. However, when both communities integrate, a and d exchange their houses, each obtaining their most preferred house, thus making that all other five agents are assigned to their own house, which they prefer less than the segregated core allocation.

The two agents who experience welfare gains (a and d) go from their second to their first best after integration occurs, obtaining a rank gain of +1. However, agent c goes from his first to his third best (a change of -2 in rank), agent g goes from his first to his fourth best (a change of -3 in rank), and so on, until agent e who goes from his best to his worst option (a change of -6 in rank). When we add the total welfare losses $(+1+1-2-3-4-5-6)$, we obtain $-\frac{1}{2}(n^2-n-k^2-k) = -18$. Dividing -18 by $n^2 = 49$, we find an average welfare reduction of 36.7% of the length of agents' preferences.

In the next section, we generalize these findings, providing upper bounds for i) the number of agents harmed by integration, and ii) the size of average welfare losses.

3.4 Worst-case Results

Unfortunately, the integration of housing markets may harm the vast majority of agents. In the worst-case scenario, up to $n-k$ agents are harmed by integration, and this upper bound is tight.

Proposition 3.1. *For any pair (n,k), there exists an EHM in which $|N^-(\sigma^*)| = n-k$; whereas there is no EHM in which $|N^-(\sigma^*)| > n-k$.*

Proof. The EHM in Example 3.1 illustrates an EHM showing that the $n-k$ bound is attainable. We can extend the construction of this example to arbitrary values of n and k as follows:

1. Enumerate agents arbitrarily so that agents from community C_1 are first, then those in C_2, and so on. Separate agents into two sets, namely X and $N \setminus X$. The set X contains the first agent from each community only. The agent from community C_j in X is denoted by j^*. The last agent in community C_j (which is in $N \setminus X$) is denoted by $\underline{j}$.
2. The preferences for any agent $i^* \in X$ are such that:
   (a) $\text{rk}_{\succ_{i^*}}((i+1)^*) = 1$ (modulo k) and
   (b) $\text{rk}_{\succ_{i^*}}(\underline{i}) = 2$.
3. The preferences for any agent $i \in N \setminus X$ are such that:
   (a) $\text{rk}_{\succ_i}(i-1) = 1$,

(b) For any $j^* \in X$ and $h \in N \setminus X$ (with $h \neq i+1$), $\text{rk}_{\succ_i}(j^*) < \text{rk}_{\succ_i}(h)$, and

(c) For any two $h, h' \in N \setminus X$ and $h, h' \neq i+1$, $\text{rk}_{\succ_i}(h) < \text{rk}_{\succ_i}(h')$ if $h < h'$.

Constructing the preferences in such a way guarantees that, in the segregated core allocation, every agent in $N \setminus X$ obtains their first choice, whereas every agent in X gets their second choice. In contrast, in the integrated core allocation, every agent in X obtains their first choice, whereas everybody in $N \setminus X$ obtains an object ranked from $k+1$ to n. Example 3.1 was constructed in this fashion.

To see that the $n-k$ upper bound is tight, assume by contradiction that more than $n-k$ agents are harmed by integration, which implies that there is one community in which all agents are harmed by integration, say C_j. But then $\sigma^*(\cdot, N)$ is not a core allocation for $(N, \succ)$, because any alternative allocation x such that $x_i = \sigma^*(i, C)\ \forall i \in C_j$ dominates it (since C_j is effective for allocation x and every agent in C_j prefers the segregated over the integrated allocation). That the integrated core allocation is not a core allocation is a contradiction, which terminates the proof. □

Proposition 3.1 implies that integration may harm the majority of agents in Shapley–Scarf markets. This is a striking observation, since the integration of Gale–Shapley marriage markets (in which two sets of agents are matched to each other) always benefits more agents than those it harms (see Proposition 2 in Ortega (2018), also Gale and Shapley (1962); Gärdenfors (1975)).[5]

Given the negative result in Proposition 3.1, we may think that integration can still be justified if the size of the welfare gains experienced by a minority are much larger than the size of the welfare losses suffered by a majority. Unfortunately, in the worst-case scenario, the size of the losses from integration is much larger than the size of the gains from integration. In particular, we show below that the agents' average welfare loss may be negative and asymptotically equivalent to an increase in ranking of 50% of the length of the agents' preference list. We provide a tight lower bound on the size of agents' average welfare loss.

Proposition 3.2. *For any pair (n, k), there exists an EHM in which $\overline{\Gamma}(\sigma^*) = \frac{-n^2+n+k^2+k}{2n^2}$; whereas there is no EHM in which $\overline{\Gamma}(\sigma^*) < \frac{-n^2+n+k^2+k}{2n^2}$*

[5]One may consider the opposite scenario, in which an integrated market of size n breaks into k disjoint communities. Simple examples show that all agents can become worse off after markets disintegrate, irrespective of the value of k.

Proof. Example 3.1 shows that our lower bound for $\overline{\Gamma}(\sigma^*)$ is attainable. We constructed the EHM in Example 3.1 in such a way that the minimum possible number of agents gain from integration (that is, k, per Proposition 3.1), and that the size of such gains is as small as possible (+1). On the other side, the welfare losses of the remaining $n-k$ individuals go from -2 to $-n+1$ (the largest possible welfare loss). We can replicate such construction for EHMs with arbitrary values of n and k as described in the proof of Proposition 3.1 to obtain:

$$\begin{aligned}
\overline{\Gamma}(\sigma^*) &= \frac{1}{n^2}\left(k*1-\sum_{i=1}^{n-k} n-i\right) && (3.1)\\
&= \frac{1}{n^2}\left(k-n(n-k)+\sum_{i=1}^{n-k} i\right) && (3.2)\\
&= \frac{1}{n^2}\left(k-n^2+nk+\frac{(n-k)(n-k+1)}{2}\right) && (3.3)\\
&= -\frac{1}{2n^2}(n^2-n-k^2-k) && (3.4)
\end{aligned}$$

This establishes that our lower bound can be attained for arbitrary values of n and k. Interestingly, our lower bound does not depend on the size of each community relative to the size of the whole society. Note that when n grows and k remains constant, $\overline{\Gamma}(\sigma^*) \sim -1/2$.

We now show that our lower bound for $\overline{\Gamma}(\sigma^*)$ is tight, with the help of some additional definitions and two auxiliary lemmas. Given a core allocation x^* for a HM $(N, \succ)$ and an integer r such that $1 \le r \le n$, let $m(r, x^*) := |\{i \in N : \mathrm{rk}_i(x_i^*)\}| = r$. Similarly, let $M(r, x^*) := |\{i \in N : \mathrm{rk}_i(x_i^*)\}| \ge r$.

Lemma **3.1.** *In any core allocation x^*, $rk_i(x_i^*) \le rk_i(\omega_i)$.*

Proof. This is a well-known fact due to any core allocation being individually rational. □

Lemma **3.2.** *In any core allocation x^*, $m(r, x^*) \le n-r+1$.*

Proof. For $r = n$, our lemma says $m(n, x^*) \le 1$. Note that if $\mathrm{rk}_i(x_i^*) = n$, then $x_i{}^* = \omega_i$ because of Lemma 3.1. Therefore, we cannot have $m(n, x^*) > 1$, as otherwise two agents are assigned their own house but they would like to exchange their house with each other, and thus x^* is not a core allocation.

For $r = n-1$, suppose by contradiction that $m(n-1, x^*) > 2$. Then there exists three agents j, l, h for which $\text{rk}(x_i^*) = n-1$ for all $i \in \{j, l, h\}$. But for each of those agents, there exists a house $\omega_i' \in \{\omega_j, \omega_l, \omega_h\}$ such that $\omega_i' \succ_i x_i^*$ and $\omega_i' \succ_i \omega_i$ for all $i \in \{j, l, h\}$. Therefore, x^* is not a core allocation, since there is a reallocation of houses among j, l, h that is effective for such coalition and that is strictly preferred.

The same argument applies to any other values of $r < n-1$. Suppose by contradiction that there exists some $r' \leq n-1$ such that $m(r', x^*) > n - r' + 2$. Then there are $n - r' + 2$ agents for which $\text{rk}(x_i^*) = r'$. But for each of these agents i, there exists a house ω_j belonging to one of these $n - r' + 2$ agents such that $\omega_j \succ_i x_i^*$ and $\omega_j \succ_i \omega_i$. Therefore, x^* is not a core allocation, since there is a reallocation of houses among those $n - r' + 2$ agents that is effective for such coalition and that is strictly preferred. Hence, the argument holds for all r.

□

Lemma **3.3.** *In any core allocation x^*, $M(r, x^*) \leq n - r + 1$.*

Proof. For $r = n$, the statement in Lemma 3.3 is the same as in Lemma 3.2. For $r = n-1$, assume by contradiction that $M(n-1, x^*) > 2$. By Lemma 2 we cannot have that two agents are allocated a house ranked n for both, or that three agents are allocated a house ranked $n-1$. Thus, it must be that one agent gets a house ranked n (agent j) and two agents get a house ranked $n-1$ (agents h and l). Then we have $x_j = \omega_j$ by Lemma 3.1. Furthermore, for $i \in \{h, l\}$, there are two houses $x_i', x_i'' \in \{\omega_j, \omega_h, \omega_l\}$ such that $x_i' \succ_i x_i$ and $x_i \succ_i x_i''$, where $x_i' \neq \omega_i$ per Lemma 3.1. If, for either agent h or l, $x_i' = \omega_j$, then j and such agent would like to exchange their endowments and would be strictly better off, and thus $\text{rk}_h(\omega_j) = \text{rk}_l(\omega_j) = n$. But because $\text{rk}_h(x_h) = \text{rk}_l(x_l) = n-1$, they must be getting their own houses, that is, $x_h = \omega_h$ and $x_l = \omega_l$. But then, agents h and l are better of by trading their endowments, and thus x^* is not a core allocation, a contradiction. The same argument applies for all other values of $r < n-1$. □

Armed with these three auxiliary lemmas, we are ready to prove that $\overline{\Gamma}(\sigma^*) \geq \frac{-n^2+n+k^2+k}{2n^2}$. By Proposition 3.1, at most $n-k$ people may experience negative gains from integration. These are defined, for each agent i, as $\gamma_i(\sigma^*) := \text{rk}_i(\sigma_i^*(i, C)) - \text{rk}_i(\sigma_i^*(i, N))$. To make $\gamma_i(\sigma^*)$ as small as possible, we need to fix $\text{rk}_i(\sigma_i^*(i, C)) = 1$ and make $\text{rk}_i(\sigma_i^*(i, N))$ as large as possible. But Lemma 3.3 shows that $\text{rk}(\sigma_i^*(i, N)) = n$ for at most one agent, $\text{rk}_i(\sigma_i^*(i, N)) \geq n-1$ for at most two agents, and so on. Thus, in the worst-case scenario, the sum of the welfare gains from integration among those $n-k$ agents equals

$$-\sum_{i=1}^{n-k}(n-i) = \frac{-n^2+n+k^2-k}{2} \tag{3.5}$$

Similarly, the smallest positive gains from integration for the remaining k agents (which must exist by Proposition 3.1) are equal to 1. Thus, the smallest possible value for $\overline{\Gamma}(\sigma^*)$ is

$$\overline{\Gamma}(\sigma^*) = -\frac{1}{2n^2}(n^2-n-k^2-k) \tag{3.6}$$

□

Proposition 3.2 can be compared to an analogous result in Gale–Shapley marriage markets. The average welfare gains may also be negative in Gale–Shapley markets, but only up to 37.5% of the length of preference lists (Ortega, 2019).[6] Taken together, Propositions 3.1 and 3.2 show that the integration of Shapley–Scarf markets can be hard to achieve, and in particular is more difficult to obtain (in the worst-case scenario) than in Gale–Shapley marriage markets.

3.5 Average-case Results

In the previous section, we found two negative results regarding the integration of Shapley–Scarf markets; however, both results are about worst-case scenarios. While these results are interesting on their own, one may argue that these are knife-edge scenarios, and wonder whether market integration would generate welfare gains on average.

To answer this question, we study *random housing markets* (RHM). Given a set of agents, an RHM is generated by drawing a complete preference list for each agent independently and uniformly at random. Similarly, a *random extended housing market* (REHM) is an RHM where the set of agents is partitioned into disjoint communities $C_1, \ldots, C_k$, each of size $n_1, \ldots, n_k$ (where $n = n_1 + \ldots + n_k$). We emphasize that the randomness refers to agents' preferences and not to the partition C, which is deterministic. Random housing markets were first studied by Frieze and Pittel (1995) and Knuth (1996). The latter proved the following seminal result.

[6] This lower bound is not proven to be tight but is the best bound available.

Lemma **3.4** (Knuth, 1996). *In a RHM,* $\mathbb{E}(\sum_{i=1}^{n} rk_i(x_i^*)) = (n+1)H_n - n$, *where* H_n *is the* n*-th harmonic number, that is,* $H_n \coloneqq \sum_{i=1}^{n} \frac{1}{i}$.

We can use Knuth's theorem to find the expected size of the average welfare gains in REHMs. Let us define the *total gains from integration for community* C_j as $\Gamma_{C_j}(\sigma) \coloneqq \sum_{i \in C_j} \gamma_i$.[7] The *average percentile gains from integration for community* C_j are denoted by $\overline{\Gamma}_{C_j}(\sigma) \coloneqq \frac{\Gamma(\sigma)}{n\, n_j}$. We divide by n_j to take the average across all agents in the community C_j, and by n to normalize by the length of agents' preference lists. Equipped with these new definitions, we can compute average gains from integration, which are positive for agents belonging to any community.

Proposition 3.3. $\mathbb{E}[\overline{\Gamma}_{C_j}(\sigma^*)] = \frac{(n+1)[(n_j+1)H_{n_j} - n_j]}{n_j(n_j+1)n} - \frac{(n+1)H_n - n}{n^2}$.

Proof. For any $i, j \in C_j$ and any community C_j, define the *relative rank* of house ω_h in the preference order of agent i by $\hat{\mathrm{rk}}_i(\omega_h) \coloneqq |\{l \in C_j : \omega_l \succcurlyeq_i \omega_h\}|$. This is, the relative rank indicates the position of a house in an agent's preference ranking compared *only* to houses owned by other agents belonging to the same community. Knuth's result directly implies that

$$\mathbb{E}[\sum_{i=1}^{n} \mathrm{rk}_i(\sigma^*(i, N))] = (n+1)H_n - n, \text{and} \tag{3.7}$$

$$\mathbb{E}[\sum_{i=1}^{n_j} \hat{\mathrm{rk}}_i(\sigma^*(i, C_j))] = (n_j+1)H_{n_j} - n_j,\ \forall j \in \{1, \ldots, k\} \tag{3.8}$$

So that before integration, agents are assigned to a house relatively ranked $(n_j + 1)H_{n_j} - n_j$. To complete the proof, we need to figure out which position is such a house in the absolute rank of all houses (that is, convert the relative rank into the full rank). To do so, suppose that a house assigned to an agent in a segregated allocation has a relative rank q. A randomly chosen house, belonging to an agent from another community, could be better ranked than house 1, between houses 1 and 2, ..., between houses $q-1$ and q, and so on. Therefore, a random house belonging to another agent is in any of those gaps with probability $\frac{1}{n_j+1}$ and thus has $\frac{q}{n_j+1}$ chances of being more highly ranked than the house with relative ranking q. There are $(n - n_j)$ houses from other communities. On average, $\frac{q(n-n_j)}{n_j+1}$ houses will be better ranked. Furthermore, there were already q houses in his own community ranked better than it. This implies that its expected ranking is

[7] Recall that $\gamma_i(\sigma) \coloneqq \mathrm{rk}_i(\sigma(i, C)) - \mathrm{rk}_i(\sigma(i, N))$.

$q + \frac{q(n-n_j)}{n_j+1} = \frac{q(n+1)}{n_j+1}$. Substituting q for the expression obtained in equation (3.8), we obtain

$$\mathbb{E}[\overline{\Gamma}_{C_j}(\sigma^*))] = \frac{(n+1)[(n_j+1)H_{n_j} - n_j]}{n_j(n_j+1)n} - \frac{(n+1)H_n - n}{n^2} \tag{3.9}$$

□

Proposition 3.3, which is interesting per se, provides valuable comparative statics, which we present in the following Corollary.

Corollary 3.1. *The expected welfare gains from integration are positive for all agents and higher for agents in smaller communities.*

For example, if we merge three Shapley–Scarf markets of size 60, 30, and 10, the corresponding welfare gains in terms of house rank are 1.98, 5.95, and 16.16, that is, agents from the market with only 10 agents improve the ranking of their assigned house by 16 positions, whereas those in the market with 60 agents only improve theirs by 2 positions. In percentile terms, agents from the smallest market improve the rank of their assigned house by 16% of the length of their preference list, whereas agents from the largest market increase their corresponding rank only by 2% of the length of their preference list. Our theoretical predictions match very accurately the gains from integration observed in simulated random markets. Averaging the results of a thousand random markets (with three markets each of sizes 60, 30 and 10), we obtain that the realized gains from integration are 2.07, 5.93, and 16.12 (with standard deviations of 1.31, 2.27, and 5.82, respectively).[8]

We now turn to study the expected number of agents who are harmed by integration in each community, that is, $|N^-_{C_j}(\sigma^*)|$. To do so, we relate the number of trading cycles in TTC for the segregated markets to the number of agents harmed by integration via two auxiliary Lemmas. For any community, C_j, let t_j be the number of cycles obtained by TTC when computing the segregated core allocation $\sigma^*(\cdot, C_j)$, and let $t \coloneqq \sum_{j=1}^{k} t_j$.[9]

Our first auxiliary Lemma relates $|N^-_{C_j}(\sigma^*)|$ to t_j.

Lemma **3.5.** *In any EHM,* $|N^-_{C_j}(\sigma^*)| \leq n_j - t_j$.

[8] The corresponding code is available from www.josueortega.com.

[9] To clarify, t is the number of cycles, not of iterations. One iteration in TTC may generate more than one cycle.

Proof. In any cycle obtained by TTC when computing the segregated core allocation $\sigma^*(\cdot, C_j)$, we must either have that all agents in the cycle are in $N^0_{C_j}(\sigma^*)$ or that at least one agent is in $N^+_{C_j}(\sigma^*)$. Otherwise, there is a cycle (involving a set of agents S) with at least one agent in $N^-_{C_j}(\sigma^*)$ and with no agent in $N^+_{C_j}(\sigma^*)$. Such a combination cannot occur. If all agents in the cycle are in $N^-_{C_j}(\sigma^*)$, then those agents are clearly a blocking coalition to the integrated core allocation. If some agents are in $N^-_{C_j}(\sigma^*)$ and some in $N^0_{C_j}(\sigma^*)$, then when we run TTC to find the integrated core allocation, there is an agent $i \in N^-_{C_j}(\sigma^*)$ who is pointed by an agent $h \in N^0_{C_j}(\sigma^*)$, that is, h's assignment does not change (it is ω_i before and after integration) but the one of i becomes worse. But when we run TTC, i points to the agent owning the best house available. Now, if $\sigma^*(i, C)$ is no longer available, it means that its owner exited in an earlier cycle during TTC, and thus she must have received a better house, and thus there is an agent in $N^+_{C_j}(\sigma^*)$, a contradiction. □

Our second auxiliary lemma computes the expected number of cycles in random housing markets. It appears as Theorem 2 in Frieze and Pittel (1995). Let t' denote the number of cycles formed during the execution of TTC in an RHM with n' agents. Then,

Lemma **3.6** (Frieze and Pittel, 1995). *$\mathbb{E}[t'] = \sqrt{2\pi n'} + O(\log n')$.*

Note that Lemma 3.6 implies that, in a REHM:

$$\mathbb{E}[t_j] \quad = \quad \sqrt{2\pi n_j} + O(\log n_j) \tag{3.10}$$

Combining Lemmas 3.5 and 3.6, we obtain an upper bound on the expected number of agents harmed by integration in each community. Proposition 3.4 below presents this upper bound.

Proposition 3.4. *$\mathbb{E}[|N^-_{C_j}(\sigma^*)|] \leq n_j - \sqrt{2\pi n_j} - O(\log n_j)$.*

Proof. Substituting t_j in Lemma 3.6 for its value in equation (3.10) together, we directly obtain the proof of our result. □

Proposition 3.4 provides, as a Corollary, a bound on the expected total number of agents in the whole economy that are harmed by integration.

Corollary 3.2. *$\mathbb{E}[N^-(\sigma^*)] \leq n - \sqrt{2\pi}(\sum_{j=1}^k \sqrt{n_j}) - O(\log \prod_{j=1}^k n_j)$.*

Proposition 3.4 is our only bound that is not tight, but is nevertheless informative. Returning to our example of an EHM divided into three communities of sizes 60, 30, and 10, Proposition 3.4 tells us that, on average, the TTC algorithm generates around 30 trading cycles when computing the integrated core allocation. In each of those cycles, at least one person is not harmed by integration. Consequently, at most 70 agents can be harmed by integration. But Proposition 3.4 says more: it tells us the distribution of agents harmed by integration across communities. Thus, in the market of size 60, the expected number of agents harmed by integration is smaller than 44. Similarly, for the markets of sizes 30 and 10, the expected number of agents harmed by integration is smaller than 19 and 4, respectively.

Another Corollary of Proposition 3.4 is that whenever all communities have the same number of agents n_1, market integration never harms more than half of the total population if n_1 is sufficiently small.

Corollary 3.3. *If $n_1 = \ldots = n_k$, then $\mathbb{E}[|N^-(\sigma^*)|] \leq \frac{n}{2}$ if $n_1 \leq 8\pi \approx 25.13$.*

Proof. From Corollary 3.2, we have that:

$$\mathbb{E}[N^-(\sigma^*)] \quad = \quad kn_1 - k\sqrt{2\pi n_1} - O(\log n_1^k) \tag{3.11}$$

and therefore $\mathbb{E}[N^-(\sigma^*)]$ is less than $n/2$ when

$$\begin{aligned} k(n_1 - \sqrt{2\pi n_1}) - O(\log n_1^k) \quad &\leq \quad \frac{kn_1}{2} &(3.12)\\ n_1 \quad &\leq \quad 2[\sqrt{2\pi n_1} + O(\log n_1^k)/k] &(3.13)\end{aligned}$$

In particular, condition (13) is satisfied whenever:

$$n_1 \quad \leq \quad 2[\sqrt{2\pi n_1}] = 8\pi \approx 25.13 \tag{3.14}$$

□

For several sensible combinations of parameters, we never observe that the number of agents harmed by integration was over 50%. The fraction of agents harmed by integration was between 14% to 22%, and becomes smaller as k increases and as n decreases (see Table 3.2). The intuition behind these changes is that as k grows, integrations offer more opportunities for trade; whereas when n grows (and k remains constant) the probability that the integrated and segregated matchings are the same becomes smaller, and therefore more agents benefit and are harmed by market integration (because fewer agents are unaffected by integration). Our simulations show that the bound in Proposition 3.4 regarding the number of agents harmed by integration can be improved. We leave this interesting question for future research.

Table 3.2: Fraction of agents affected by integration.
Average over a thousand simulations with preferences drawn uniformly at random. Standard errors in parenthesis.

n	k					
	2		3		5	
	Benefit	Harmed	Benefit	Harmed	Benefit	Harmed
25	53.52	19.95	64.68	17.63	75.05	14.08
	(6.8)	(4.66)	(5)	(3.54)	(3.59)	(2.55)
50	54.64	21.45	65.57	18.47	75.54	14.67
	(4.47)	(3.47)	(3.52)	(2.66)	(2.41)	(1.84)
100	55.4	22.17	66.06	18.99	75.88	14.88
	(3.28)	(2.34)	(2.42)	(1.8)	(1.78)	(1.34)

Other Interpretations As discussed in the related literature section, the core from random endowments is equivalent to the allocation obtained with random serial dictatorship in a market with no property rights, that is, assigning a random order among agents and letting them choose their most preferred object that remains available according to such order (Knuth, 1996; Abdulkadiroğlu and Sönmez, 1998). Therefore, the results obtained for random markets in this section also apply to the integration of markets with no endowments in which random serial dictatorship is used.

3.6 Specific Preference Domains

Although uniform and independent preferences are the most natural and simple preferences to consider in random markets, it is well-known that in real-life applications such as kidney exchange, agents' preferences are strongly correlated, with some "houses" being particularly desired by most agents. In this section, we show that if preferences satisfy a particular type of correlation structure, we can guarantee that no more than half of the total population of agents is harmed by integration.

To do so, let $q(r, \widetilde{\succ}_{C_j})$ be the set of agents in the community C_j placed at rank r by any agent in their own community (including themselves) in preference profile $\widetilde{\succ}_{C_j}$. This is, for any positive integer r and any $j \in \{1, \ldots, k\}$, $q(r, \widetilde{\succ}_{C_j}) \coloneqq \{i \in C_j : \exists h \in C_j : \mathrm{rk}_j(\omega_i) = r\}$. Similarly, let $Q(r, \widetilde{\succ}_{C_j}) \coloneqq \bigcup\limits_{t=1}^{r} q(t, \widetilde{\succ}_{C_j})$ be the set of agents in community j placed at rank r and above.

Now we introduce the property that will ensure that market integration does not harm a majority of agents, which we call *sequential dual dictator*. This property was recently introduced by Troyan (2019) in a two-sided extension of a Shapley–Scarf market, which he used to characterize the obvious strategy-proof implementation of TTC.

Definition 1 (Sequential dual dictator property)**.** A preference profile $\succ$ satisfies the sequential dual dictator property if, for any positive integer r and $\forall j \in \{1, \ldots, k\}$, each of their corresponding preference restriction $\succ_{C_j}$ satisfies

$$|Q(r, \widetilde{\succ}_{C_j})| \leq r + 1$$

In Example 3.3, we show that the preference profile in Example 3.1 does not satisfy the sequential dual dictator property and provides a preference profile that does. In Example 3.1, $|Q(1, \widetilde{\succ}_{C_1})| = |\{b, c, a\}| > 2$, violating the sequential dual dictator property. Similarly, $|Q(1, \widetilde{\succ}_{C_2})| = |\{e, f, g, d\}| > 2$. In contrast, in the profile on the right in Example 3.3, $|Q(1, \widetilde{\succ}_{C_1})| = |\{c, a\}| \leq 2$, $|Q(1, \widetilde{\succ}_{C_2})| = |\{e, f\}| \leq 2$ and $|Q(2, \widetilde{\succ}_{C_2})| = |\{e, f, d\}| \leq 3$. Whenever preferences satisfy the sequential dual dictator property, we can guarantee that no more than half of the agents in each community are harmed by integration. Note that, in contrast to Proposition 3.4, here we bound the number of agents harmed by integration in every EHM, instead of the expected number of agents harmed by integration across all REHMs.

Example 3.3: The preference profile on the right satisfies the sequential dual dictator property, unlike the one on the left.

a	b	c	d	e	f	g
b	c	a	e	f	g	d
⋮	⋮	⋮	⋮	⋮	⋮	⋮

a	b	c	d	e	f	g
c	c	a	e	f	f	e
b	b	c	f	d	e	d
a	c	b	d	g	g	f
			g	e	d	g

Proposition 3.5. *If $\succ$ satisfies the sequential dual dictator property, then* $|N^-_{C_j}(\sigma^*)| \leq \frac{n_j}{2}$.

Proof. To complete the proof, we examine the number and length of trading cycles generated by the TTC algorithm when computing the segregated core allocation $\sigma^*(\cdot, C_j)$ for community C_j. At the first iteration, all agents point to the owner of their most preferred house, and if the sequential dual dictator property is satisfied, there are only two vertices with a positive in-degree. A trading cycle is created, either of those agents pointing to themselves or pointing at each other, and therefore each cycle created in the first iteration of TTC has length at most 2. In the second iteration, at most two agents have positive in-degree (because at least one agent was removed in the first iteration). Either one or two cycles are formed in iteration 2, and they have length of at most 2. The argument repeats for each iteration: each trading cycle has length at most 2.

Now we invoke an argument that we used in the proof of Lemma 5, showing that in any cycle, we must either have that all agents are in $N^0(\sigma^*)$ or that at least one agent is in $N^+_{C_j}(\sigma^*)$. We have shown that there are at least $n_j/2$ cycles in each community. Therefore, $|N^-_{C_j}(\sigma^*)| \leq \frac{n_j}{2}$. □

One particular case of preference profiles satisfying the sequential dual dictator property are those in which all agents have the same preferences. Such preferences have been extensively studied in Gale–Shapley marriage markets because they guarantee the uniqueness of the core allocation and ensure that truth-telling is a Nash equilibrium of the revelation game induced by any stable mechanism (Gusfield and Irving, 1989). The sequential dual dictator preference domain is larger than this classical domain of equal preferences. The sequential dual dictatorship only imposes a particular structure on the preferences of each community over its own houses and is therefore substantially less restrictive than identical preferences.

3.7 Conclusion

Market integration leads to more efficient outcomes and yet real-life offers plenty of examples of markets that fail to integrate and operate disjointly. In this paper, we have provided results that shed light on why this might be the case for a specific type of market where monetary transfers are not permitted.

Our explanation lies in the fact that market integration may have negative consequences for most traders. These negative consequences are so dire that they vastly outweigh the welfare benefits of those who become better off with market integration. Somewhat surprisingly, the average effect on the economy can be so bad that the average trader ends up with an allocation in the lower half of their preference list.

These negative consequences of market integration, however, are the exception rather than the rule, as we have shown formally. Two interesting open problems for further research are: i) to obtain conditions that fully characterize which types of Shapley–Scarf markets benefit from integration, and more generally, ii) to provide a comprehensive and unified discussion of the institutional features that prevent markets from integrating.

Bibliography

Abdulkadiroğlu, A. and T. Sönmez (1998): "Random serial dictatorship and the core from random endowments in house allocation problems," *Econometrica*, 66, 689–701.

——— (1999): "House allocation with existing tenants," *Journal of Economic Theory*, 88, 233–260.

Agarwal, N., I. Ashlagi, E. Azevedo, C. R. Featherstone, and Ö. Karaduman (2019): "Market failure in kidney exchange," *American Economic Review*, 109, 4026–70.

Alcalde-Unzu, J. and E. Molis (2011): "Exchange of indivisible goods and indifferences: The top trading absorbing sets mechanisms," *Games and Economic Behavior*, 73, 1–16.

Alva, S. and B. Doğan (2021): "Choice and market design," *arXiv preprint arXiv:2110.15446*.

Arrow, K. J. (1959): "Rational choice functions and orderings," *Economica*, 26, 121–127.

Ashlagi, I. and A. E. Roth (2014): "Free riding and participation in large scale, multi-hospital kidney exchange," *Theoretical Economics*, 9, 817–863.

Aslan, F. and J. Lainé (2020): "Competitive equilibria in Shapley–Scarf markets with couples," *Journal of Mathematical Economics*, 89, 66–78.

Aue, R., T. Klein, and J. Ortega (2020): "What happens when several separate and unequal school districts merge?" *ArXiv*.

Avataneo, M. and B. Turhan (2021): "Slot-specific priorities with capacity transfers," *Games and Economic Behavior*, 129, 536–548.

Aygün, O. and I. Bó (2021): "College admission with multidimensional privileges: The Brazilian affirmative action case," *American Economic Journal: Microeconomics*, 13, 1–28.

Aygün, O. and T. Sönmez (2013): "Matching with contracts: Comment," *American Economic Review*, 103, 2050–2051.

Aygün, O. and B. Turhan (2017): "Large-scale affirmative action in school choice: Admissions to IITs in India," *American Economic Review*, 107, 210–213.

——— (2020): "Dynamic reserves in matching markets," *Journal of Economic Theory*, 188.

——— (2022): "How to de-reserve reserves: Admissions to technical colleges in India," *Management Science*, 0.

——— (2023): "Priority design for engineering colleges in India," *Indian Economic Review*, 1–11.

Aziz, H. and B. De Keijzer (2012): "Housing markets with indifferences: A tale of two mechanisms," in *Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Bajraj, G. and L. Ülkü (2015): "Choosing two finalists and the winner," *Social Choice and Welfare*, 45, 729–744.

Barberà, S., G. De Clippel, A. Neme, and K. Rozen (2019): "Order-k rationality," *Economic Theory*, 1–19.

Baswana, S., P. P. Chakrabarti, S. Chandran, Y. Kanoria, and U. Patange (2019): "Centralized admissions for engineering colleges in India," *INFORMS Journal on Applied Analytics*, 45, 338–354.

Bhardwaj, B. and K. Manocha (2021): "Choice by Rejection," *arXiv preprint arXiv:2108.07424*.

——— (2023): "Choice by Rejection," *Working Paper*.

Bo, I. and R. Hakimov (2022): "The iterative deferred acceptance mechanism," *Games and Economic Behavior*, 135, 411–433.

Bossert, W., Y. Sprumont, and K. Suzumura (2005): "Maximal-element rationalizability," *Theory and Decision*, 58, 325–350.

Chambers, C. P. and T. Hayashi (2017): "Gains from trade," *International Economic Review*, 58, 923–942.

——— (2020): "Can everyone benefit from economic integration?" *Journal of Public Economic Theory*.

Chambers, C. P. and M. B. Yenmez (2018): "A simple characterization of responsive choice," *Games and Economic Behavior*, 111, 217–221.

Che, Y.-K. and O. Tercieux (2019): "Efficiency and stability in large matching markets," *Journal of Political Economy*, 127, 2301–2342.

Chen, Y. and T. Sönmez (2002): "Improving efficiency of on-campus housing: An experimental study," *American Economic Review*, 92, 1669–1686.

Cherepanov, V., T. Feddersen, and A. Sandroni (2013): "Rationalization," *Theoretical Economics*, 8, 775–800.

Chernoff, H. (1954): "Rational selection of decision functions," *Econometrica: journal of the Econometric Society*, 422–443.

Correa, J., R. Epstein, J. Escobar, I. Rios, B. Bahamondes, C. Bonet, N. Epstein, N. Aramayo, M. Castillo, A. Cristi, et al. (2021): "School choice in Chile," *Operations Research*, 70, 1066–1087.

Doğan, B. and M. B. Yenmez (2019): "Unified versus divided enrollment in school choice: Improving student welfare in Chicago," *Games and Economic Behavior*, 118, 366–373.

Echenique, F., D. Golovin, and A. Wierman (2011): "A revealed preference approach to computational complexity in economics," in *Proceedings of the 12th ACM conference on Electronic commerce*, 101–110.

Echenique, F. and K. Saito (2019): "General luce model," *Economic Theory*, 68, 811–826.

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

Ehlers, L. and Y. Sprumont (2008): "Weakened WARP and top-cycle choice rules," *Journal of Mathematical Economics*, 44, 87–94.

Ekmekci, M. and M. B. Yenmez (2019): "Common enrollment in school choice," *Theoretical Economics*, 14, 1237–1270.

Eliaz, K. and E. A. Ok (2006): "Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences," *Games and economic behavior*, 56, 61–86.

Feigenbaum, I., Y. Kanoria, I. Lo, and J. Sethuraman (2020): "Dynamic matching in school choice: Efficient seat reassignment after late cancellations," *Management Science*, 66, 5341–5361.

Frick, M. (2016): "Monotone threshold representations," *Theoretical Economics*, 11, 757–772.

Frieze, A. and B. G. Pittel (1995): "Probabilistic analysis of an algorithm in the theory of markets in indivisible goods," *The Annals of Applied Probability*, 768–808.

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

Gärdenfors, P. (1975): "Match making: assignments based on bilateral preferences," *Behavioral Science*, 20, 166–173.

Gerasimou, G. (2016): "Partially dominant choice," *Economic Theory*, 61, 127–145.

Gersbach, H. and H. Haller (2021): "Gainers and losers from market integration," *Mathematical Social Sciences*.

Gusfield, D. and R. W. Irving (1989): *The stable marriage problem: structure and algorithms*, MIT press.

Haeringer, G. and V. Iehlé (2021): "Gradual college admission," *Journal of Economic Theory*, 198, 105378.

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

Houthakker, H. S. (1950): "Revealed preference and the utility function," *Economica*, 17, 159–174.

Huber, J., J. W. Payne, and C. Puto (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of consumer research*, 9, 90–98.

Huber, V. L., M. A. Neale, and G. B. Northcraft (1987): "Decision bias and personnel selection strategies," *Organizational Behavior and Human Decision Processes*, 40, 136–147.

Jaramillo, P. and V. Manjunath (2012): "The difference indifference makes in strategy-proof allocation of objects," *Journal of Economic Theory*, 147, 1913–1946.

Kamada, Y. and F. Kojima (2015): "Efficient matching under distributional constraints: Theory and applications," *American Economic Review*, 105, 67–99.

Kesten, O. (2010): "School choice with consent," *Quarterly Journal of Economics*, 125.

Knuth, D. E. (1996): "An exact analysis of stable allocation," *Journal of Algorithms*, 20, 431–442.

Kojima, F. and M. Manea (2010): "Axioms for deferred acceptance," *Econometrica*, 78, 633–653.

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

Kovach, M. and L. Ülkü (2020): "Satisficing with a variable threshold," *Journal of Mathematical Economics*, 87, 67–76.

Lahiri, S. (2001): "The top cycle and uncovered solutions for weak tournaments," *Control and Cybernetics*, 30, 439–450.

Laran, J. and K. Wilcox (2011): "Choice, rejection, and elaboration on preference-inconsistent alternatives," *Journal of Consumer Research*, 38, 229–241.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2017): "When more is less: Limited consideration," *Journal of Economic Theory*, 170, 70–85.

Ma, J. (1994): "Strategy-proofness and the strict core in a market with indivisibilities," *International Journal of Game Theory*, 23, 75–83.

Mandler, M., P. Manzini, and M. Mariotti (2012): "A million answers to twenty questions: Choosing by checklist," *Journal of Economic Theory*, 147, 71–92.

Manjunath, V. and B. Turhan (2016a): "Two school systems, one district: What to do when a unified admissions process is impossible," *Games and Economic Behavior*, 95, 25–40.

——— (2016b): "Two school systems, one district: What to do when a unified admissions process is impossible," *Games and Economic Behavior*, 95, 25–40.

Manzini, P. and M. Mariotti (2007): "Sequentially rationalizable choice," *American Economic Review*, 97, 1824–1839.

——— (2010): "Revealed preferences and boundedly rational choice procedures: an experiment," *University of St. Andrews and IZA Working Paper*, 141–167.

——— (2012): "Categorize then choose: Boundedly rational choice and welfare," *Journal of the European Economic Association*, 10, 1141–1165.

Manzini, P., M. Mariotti, and L. Ulku (2019): "Sequential Approval: A Model of "Likes", Paper Downloads and Other Forms of Click Behaviour." .

Masatlioglu, Y. and D. Nakajima (2007): "A theory of choice by elimination," *Unpublished paper.[778, 785, 788].*

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed attention," *American Economic Review*, 102, 2183–2205.

——— (2016): "Revealed attention," *Behavioral Economics of Preferences, Choices, and Happiness*, 495–522.

Morrill, T. (2013): "An alternative characterization of the deferred acceptance algorithm," *International Journal of Game Theory*, 42, 19–28.

Moulin, H. (1985): "Choice functions over a finite set: a summary," *Social Choice and Welfare*, 2, 147–160.

——— (1986): "Choosing from a tournament," *Social Choice and Welfare*, 3, 271–291.

Ortega, J. (2018): "Social integration in two-sided matching markets," *Journal of Mathematical Economics*, 78, 119–126.

——— (2019): "The losses from integration in matching markets can be large," *Economics Letters*, 174, 48–51.

Pathak, P. A., T. Sönmez, M. U. Ünver, and M. B. Yenmez (2021): "Fair allocation of vaccines, ventilators and antiviral treatments: leaving no ethical value behind in health care rationing," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 785–786.

Pereyra, J. S. (2013): "A dynamic school choice model," *Games and Economic Behavior*, 80, 100–114.

Powell, A., A. Meltzer, C. Martin, W. Stone, E. Liu, K. Flanagan, K. Muir, and S. Tually (2019): "The construction of social housing pathways across Australia," *Australian Housing and Urban Research Institute*.

Quint, T. and J. Wako (2004): "On houseswapping, the strict core, segmentation, and linear programming," *Mathematics of Operations Research*, 29, 861–877.

Roth, A. E. and A. Postlewaite (1977): "Weak versus strong domination in a market with indivisible goods," *Journal of Mathematical Economics*, 4, 131 – 137.

Roth, A. E., T. Sönmez, and M. U. Ünver (2004): "Kidney exchange," *Quarterly Journal of Economics*, 119, 457–488.

Rubinstein, A. and Y. Salant (2006): "A model of choice from lists," *Theoretical Economics*, 1, 3–17.

Saban, D. and J. Sethuraman (2013): "House allocation with indifferences: a generalization and a unified view," in *Proceedings of the ACM Conference on Electronic Commerce*, 803–820.

Salant, Y. and A. Rubinstein (2008): "(A, f): choice with frames," *The Review of Economic Studies*, 75, 1287–1296.

Samuelson, P. A. (1938): "A note on the pure theory of consumer's behaviour," *Economica*, 5, 61–71.

Sen, A. K. (1971): "Choice functions and revealed preference," *The Review of Economic Studies*, 38, 307–317.

Shafir, E. (1993): "Choosing versus rejecting: Why some options are both better and worse than others," *Memory & cognition*, 21, 546–556.

Shapley, L. and H. Scarf (1974): "On cores and indivisibility," *Journal of Mathematical Economics*, 1, 23–37.

Simon, H. A. (1955): "A behavioral model of rational choice," *The quarterly journal of economics*, 69, 99–118.

Simonson, I. (1989): "Choice based on reasons: The case of attraction and compromise effects," *Journal of consumer research*, 16, 158–174.

Sokolova, T. and A. Krishna (2016): "Take it or leave it: How choosing versus rejecting alternatives affects information processing," *Journal of Consumer Research*, 43, 614–635.

Sönmez, T. and M. B. Yenmez (2022): "Affirmative action in India via vertical, horizontal, and overlapping reservations," *Econometrica*, 90, 1143–1176.

Sprumont, Y. (1990): "Population monotonic allocation schemes for cooperative games with transferable utility," *Games and Economic Behavior*, 2, 378–394.

Teppan, E. C. and A. Felfernig (2009): "Minimization of product utility estimation errors in recommender result set evaluations," in *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, IEEE, vol. 1, 20–27.

Toulis, P. and D. Parkes (2015): "Design and analysis of multi-hospital kidney exchange mechanisms using random graphs," *Games and Economic Behavior*, 91, 360–382.

Troyan, P. (2019): "Obviously strategy-proof implementation of top trading cycles," *International Economic Review*, 60, 1249–1261.

Tserenjigmid, G. (2019): "Choosing with the worst in mind: A reference-dependent model," *Journal of Economic Behavior & Organization*, 157, 631–652.

Turhan, B. (2019): "Welfare and incentives in partitioned school choice markets," *Games and Economic Behavior*, 113, 199–208.

Tversky, A. (1972): "Choice by elimination," *Journal of mathematical psychology*, 9, 341–367.

Yaniv, I. and Y. Schul (2000): "Acceptance and elimination procedures in choice: Noncomplementarity and the role of implied status quo," *Organizational behavior and human decision processes*, 82, 293–313.

Kriti