Incentives in One-Sided Matching Problems With Ordinal Preferences

by

Hadi Hosseini

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

One of the core problems in multiagent systems is how to efficiently allocate a set of indivisible resources to a group of self-interested agents that compete over scarce and limited alternatives. In these settings, *mechanism design* approaches such as matching mechanisms and auctions are often applied to guarantee fairness and efficiency while preventing agents from manipulating the outcomes. In many multiagent resource allocation problems, the use of monetary transfers or explicit markets are forbidden because of ethical or legal issues. One-sided matching mechanisms exploit various randomization and algorithmic techniques to satisfy certain desirable properties, while incentivizing self-interested agents to report their private preferences truthfully.

In the first part of this thesis, we focus on deterministic and randomized matching mechanisms in one-shot settings. We investigate the class of deterministic matching mechanisms when there is a quota to be fulfilled. Building on past results in artificial intelligence and economics, we show that when preferences are lexicographic, *serial dictatorship* mechanisms (and their *sequential dictatorship* counterparts) characterize the set of all possible matching mechanisms with desirable economic properties, enabling social planners to remedy the inherent unfairness in deterministic allocation mechanisms by assigning quotas according to some fairness criteria (such as seniority or priority). Extending the quota mechanisms to randomized settings, we show that this class of mechanisms are envyfree, strategyproof, and ex post efficient for any number of agents and objects and any quota system, proving that the well-studied Random Serial Dictatorship (RSD) is also envyfree in this domain.

The next contribution of this thesis is providing a systemic empirical study of the two widely adopted randomized mechanisms, namely Random Serial Dictatorship (RSD) and the Probabilistic Serial Rule (PS). We investigate various properties of these two mechanisms such as efficiency, strategyproofness, and envyfreeness under various preference assumptions (*e.g.* general ordinal preferences, lexicographic preferences, and risk attitudes). The empirical findings in this thesis complement the theoretical guarantees of matching mechanisms, shedding light on practical implications of deploying each of the given mechanisms.

In the second part of this thesis, we address the issues of designing truthful matching mechanisms in dynamic settings. Many multiagent domains require reasoning over time and are inherently dynamic rather than static. We initiate the study of matching problems where agents' private preferences evolve stochastically over time, and decisions have to be made in each period. To adequately evaluate the quality of outcomes in dynamic settings, we propose a generic *stochastic decision process* and show that, in contrast to static settings, traditional mechanisms are easily manipulable. We introduce a number of properties that we argue are important for matching mechanisms in dynamic settings and propose a new mechanism that maintains a history of pairwise interactions between agents, and adapts the priority orderings of agents in each period based on this history. We show that our mechanism is globally strategyproof in certain settings (*e.g.* when there are 2 agents or when the planning horizon is bounded), and even when the mechanism is manipulable, the manipulative actions taken by an agent will often result in a Pareto improvement in general. Thus, we make the argument that while manipulative behavior may still be unavoidable, it is not necessarily at the cost to other agents.

To circumvent the issues of incentive design in dynamic settings, we formulate the dynamic matching problem as a Multiagent MDP where agents have particular underlying utility functions (*e.g.* linear positional utility functions), and show that the impossibility results still exist in this restricted setting. Nevertheless, we introduce a few classes of problems with restricted preference dynamics for which positive results exist. Finally, we propose an algorithmic solution for agents with *single-minded preferences* that satisfies strategyproofness, Pareto efficiency, and weak non-bossiness in one-shot settings, and show that even though this mechanism is manipulable in dynamic settings, any unilateral deviation would benefit all participating agents.

Acknowledgements

First, I would like to thank my supervisors, Kate Larson and Robin Cohen, for their continuous support, guidance, and patience throughout this journey. Kate has been a mentor and role model. Her professionalism and upbeat attitude towards research has truly influenced me as a researcher. Robin's dedication to achievement and growth of her students is exemplary. I feel extremely fortunate to have shared these years with them.

Second, I am grateful to my internal committee members, Peter van Beek and Pascal Poupart, who gave valuable feedback to further improve this thesis. Special thanks to my external-internal examiner, Stanko Dimitrov, for always showing interest in my research, having a fresh perspective on my work, and also giving invaluable suggestions into the composition and presentation of my work. I am also grateful to my external examiner, Makoto Yokoo, for his compelling questions and suggestions on the final manuscript.

I would like to also acknowledge the support and friendship of other researchers and faculty members at the University of Waterloo: Michael Godfrey, Khuzaima Daudjee, Jesse Hoey, Edith Law, Ian Goldberg, and S. Keshav. Particularly, a huge thanks to Mike who always shared his experience and expertise whenever I needed it the most. I am grateful to Svitlana Taraban-Gordon who has always been supportive of my academic activities. Thanks to the wonderful staff of the Cheriton School of Computer Science: Margaret Towell, Neoma Power, Helen Jardine, Paula Roser, and Wendy Rush for their tremendous support and assistance and occasional laughters. I am thankful for the fruitful discussions with the outstanding researchers throughout the past few years: David Parkes, Toby Walsh, Sven Sueken, Haris Aziz, Omer Lev, Kevin Leyton-Brown, Nick Mattei, and Yair Zick.

I am grateful to several wonderful collaborators and friends who I have had the opportunity to work with: John Doucette, Simina Brânzei, Alan Tsang, Olga Baysal, Tariq Elahi. A huge thanks to Arthur Carvalho, Mazen Melibari, Shehroz Khan, Marek Gzeres, Vijay Menon, Majid Karimi, Noel Sardana, John Champaign, Lachlan Dufton, Maxwell Hartt, Kristin Brown, Stephanie Chesser, Amy Hackney, Charis Enns, and Angela Nyhout, for sharing their journeys, challenges, and joys and making this an unforgettable experience.

And most importantly, I would like to thank my family for their support. My mom, dad, and sister, who have always showered me with endless love and inspiration and encouraged me to be ambitious. With special thanks to my lovely partner, Mehrnaz, for her love, friendship, and support during the past few years. She has always been the voice of wisdom and the source of my relaxations. Her input in my work and our intellectually nourishing discussions have shaped my personality and style of thinking. Her constant belief in me has been empowering and encouraging.

Dedication

To my family with love.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The study and design of autonomous agents that are capable of intelligent decision making is the cornerstone of artificial intelligence research. Multiagent systems research draws various techniques from mathematical theory to computational science aiming at creating autonomous agents that can intelligently behave towards achieving certain goals when interacting with (possibly) uncertain environments. Autonomous agents are often deployed in various applications such as expert systems, online systems, customer support systems, healthcare, robotics, transportation and traffic control, computer games, information management, and electronic commerce [128, 135, 142, 143]. These applications employ the autonomous agent paradigm to represent users of the system, reason over uncertainties, and automate their decision making processes. Within artificial intelligence, the subfield of multiagent systems studies how a group of intelligent agents make decisions when interacting with other agents. A central concept in modeling decision making, particularly in multiagent systems, is how to handle individual preferences [51, 64, 73]. Agents often deal with decision situations and require computational and mathematical models of preferences to reason over various choices. In multiagent systems, preferences may guide agent behavior when seeking collective decisions in social choice problems such as voting, or when making rational decisions while competing or cooperating with other intelligent agents.

One of the core problems in multiagent systems is how to efficiently allocate a set of indivisible resources to a group of self-interested agents that compete over scarce and limited alternatives. Individual preferences are key in measuring the quality of allocations and achieving desirable social outcomes. Multiagent resource allocation leverages techniques from game theory and economics to ensure various desirable properties based on agent preferences. In these settings, *mechanism design* approaches such as matching mechanisms and auctions are often applied to guarantee fairness and efficiency while preventing agents from manipulating the outcomes.

In many multiagent resource allocation problems the use of monetary transfers or explicit markets are forbidden because of ethical and legal issues [3,17,39,65,116,117,131]. For example, when students are assigned housing on campus based on their underlying private preferences over dormitory residences, it would be unethical (if not unjust) for academic institutions to ask for payment in lieu of better placements in dormitories. Other real-life application domains wherein the use of money is prohibited are assigning teaching load among faculty, college courses to students, scarce medical resources and organs to patients, and scientific equipment to researchers within the same organization [13,50,94,117,131,132,138].

In several of these settings, the mechanism assigns the set of resources (alternatives) to agents based on self-reported preferences, and self-interested agents do not necessarily report their private preferences truthfully. For example, in student placement markets for assigning students to public schools there is strong evidence that parents misreport their preferences in order to 'game the system' and manipulate the outcome in hopes of securing better placement for their children [4, 5]. The fields of mechanism design and matching theory precisely address this issue by designing algorithms to promote honest reporting of preferences without the use of transferable currencies such as money.

Truthful reporting of preferences is key in guaranteeing other desirable properties such as efficiency and fairness. Without ensuring truthfulness, a matching mechanism can only guarantee efficient and fair allocations with respect to the, perhaps untruthful, preferences, and thus, fails to satisfy the desirable economic properties with respect to the true underlying preferences. Hence, it is necessary to design allocation mechanisms that provide incentives to self-interested agents for truthful reporting of their preferences. We refer to such mechanisms as *strategyproof mechanisms* where no agent will prefer to strategize when reporting its preference.

This dissertation investigates mechanism design approaches for matching decisions in dynamic environments with repeated decisions as well as one-shot and static settings. In the first part of this thesis, we focus on deterministic and randomized matching mechanisms in one-shot settings. Matching mechanisms are often used for allocating multiple resources to each agent while satisfying certain desirable economic properties. In many real-life applications a social planner needs to assign alternatives to agents based on a given quota. One might decide to assign the resources to agents according to a predefined quota such that two or three of agents each receive at least two resources. Moreover, a social planner may be required to choose which mechanism to use in practice, depending on the population of agents and their attitude towards risk. In this vein, we address the following crucial questions:

1. What deterministic mechanisms are appropriate to assign multiple objects to agents based on a given quota?
2. How does randomization guarantee fairness and truthfulness of quota assignments?
3. Which randomized matching mechanism should a social planner adopt in practice to ensure a desirable level of social welfare under various risk attitudes?

In the second part of this thesis, we address the issues of designing truthful matching mechanisms in dynamic settings. Many multiagent domains require reasoning over time and are inherently dynamic rather than static. Agents' preferences may stochastically evolve over time, and the population may change according to arrival or departure of agents. Dynamic mechanism design has emerged as a new paradigm for addressing resource allocation problems where money is used to facilitate the exchange of resources. Nevertheless, little has been done in the space of dynamic mechanisms for allocation problems without monetary transfers (*e.g.* nurse scheduling, campus housing). The dynamic nature of the matching problems gives rise to the following key questions:

1. What would be an appropriate model to reason about various economic properties of dynamic allocations in the ordinal domain?
2. Do the desirable properties of traditional matching mechanisms carry over to dynamic settings with repeated allocations? and if not,
3. Can we design truthful mechanisms for dynamic matching problems?

1.1 Thesis Statement

This thesis studies the strategic behavior of agents in matching markets where the use of transferable currencies is prohibited. Matching mechanisms should be designed to provide truthful incentives to self-interested agents in order to ensure desirable properties such as efficiency and fairness. The thesis statements are:

- The choice of which matching mechanism to adopt in practice relies heavily on the type of allocation (single or multiple object), the comparative nature of preferences, and the risk attitudes of agents, and social planners should take these into consideration.

- The traditional truthful matching mechanisms are highly susceptible to manipulation in dynamic settings. Using our proposed framework for *stochastic matching under dynamic ordinal preferences*, it is possible to analyze sequential matching decisions and to design adaptable mechanisms that provide desirable outcomes in certain structured settings.

1.2 Contributions

In this thesis, we advance the state of the art by addressing the aforementioned questions. Our first contribution is in the domain of one-shot matching mechanisms. Within deterministic mechanisms, we characterize the set of strategyproof, non-bossy, and neutral quota mechanisms under lexicographic preferences. We show that under a mild Pareto efficiency condition, *serial dictatorship quota* mechanisms are the only mechanisms satisfying these properties. Dropping the neutrality requirement, this class of quota mechanisms further expands to *sequential dictatorship quota* mechanisms. Furthermore, we extend the quota mechanisms to randomized settings, and show that this class of quota mechanisms are envyfree, strategyproof, and ex post efficient for any number of agents and objects and any quota system, proving that the well-studied Random Serial Dictatorship (RSD) satisfies envyfreeness when preferences are lexicographic.

The second contribution of thesis is an empirical study of the two widely adopted random mechanisms, namely Random Serial Dictatorship (RSD) and the Probabilistic Serial Rule (PS). Both mechanisms require only that agents specify ordinal preferences and have a number of desirable, but orthogonal, economic and computational properties. However, the induced outcomes of the mechanisms are often incomparable, and thus, in multiagent settings there are challenges for social planners when it comes to deciding which mechanism to adopt in practice. In the space of general ordinal preferences, we provide empirical results on the (in)comparability of RSD and PS and analyze their respective economic properties. We then instantiate utility functions for agents, consistent with the ordinal preferences, with the goal of gaining insights on the manipulability, envyfreeness, and social welfare of the mechanisms under different risk attitude models.

Our third contribution is in the domain of matching with dynamic preferences. We initiate the study of matching problems where agents' private preferences evolve stochastically over time, and decisions have to be made in each period. To evaluate the quality of outcomes in dynamic settings, we propose a generic *stochastic decision process* and show that, in contrast to static settings, traditional mechanisms are easily manipulable. The lack of truthful incentives means that an agent may strategically misreport its true idiosyncratic preference, which consequently results in inefficient and unfair outcomes. Thus, we propose a history-dependent matching policy that guarantees some level of global truthfulness while sustaining the properties of fairness and efficiency in each period.

The final contribution of this thesis is investigating several classes of dynamic problems where each utility-maximizing agent's best response is conditioned upon the truthful revelation of other agents. We show that even under the Markovian assumption and under a linear positional utility function, no optimal matching policy satisfies truthfulness. Moreover, applying a widely-studied matching mechanism to such settings is still prone to manipulation. Furthermore, we provide an algorithmic solution for allocating resources to single-minded agents who only care about their top choices, and show that if agents' preferences exhibit certain structures then traditional matching mechanisms can still guarantee our desired truthfulness property.

Overall, this dissertation combines algorithmic and computational aspects of artificial intelligence and multiagent systems with insights from theoretical and empirical economics, spanning the fields of computer science and economic theory.

1.3 Thesis Outline

This thesis is organized in two parts; Part I investigates one-sided matching mechanisms in one-shot and static settings. Part II focuses on studying matching and allocation mechanisms in dynamic settings. Besides this introductory section, this dissertation has six more chapters:

- In Chapter 2, we introduce the *one-sided matching* problem and discuss the literature on deterministic and randomized matching mechanisms, providing a chronicle of the matching mechanisms in static settings.
- In Chapter 3, we characterize the set of deterministic matching mechanisms for any number of agents and objects and any quota system, and provide a randomized extension for fair assignment of resources.

- In Chapter 4, we empirically investigate various desirable properties of two seminal random matching models, providing insights into practicality of each of these two matching mechanisms under various risk attitude models.

- In Chapter 5, we initiate the study of repeated matching with possibly dynamic preferences and provide impossibility and possibility results in designing and adopting truthful matching mechanisms.

- In Chapter 6, we focus attention on dynamic matching problems under restricted settings and cardinal utilities, and propose a few classes of problems with restricted preference dynamics for which positive results exist.

- In Chapter 7, we conclude this dissertation by summarizing the main findings and outlining a number of intriguing research directions for future study.

Much of the work in this thesis has appeared in the following papers:

- Hadi Hosseini, Kate Larson, Strategyproof quota mechanisms for multiple assignment problems, *arXiv preprint (arXiv:1507.07064)*, 2015 [75].

- Hadi Hosseini, Kate Larson, Robin Cohen, Investigating the characteristics of one-sided matching mechanisms, In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, pp. 1443–1444, 2016 [80].

- Hadi Hosseini, Kate Larson, Robin Cohen, An empirical comparison of one-sided matching mechanisms, In *The 3rd Workshop on Exploring Beyond the Worst Case in Computational Social Choice (EXPLORE) at AAMAS 2016*, 2016 [79].

- Hadi Hosseini, Kate Larson, Robin Cohen, On manipulability of random serial dictatorship in sequential matching with dynamic preferences, Student Abstract, In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*, pp. 4168–4169, 2015 [77].

- Hadi Hosseini, One-sided matching with dynamic preferences, Doctoral Consortium, In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, pp. 2005–2006, 2015 [74].

- Hadi Hosseini, Kate Larson, Robin Cohen, Matching with dynamic ordinal preferences, In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*, pp. 936–943, 2015 [76].

Chapter 2

Background

In this chapter, we describe deterministic and randomized allocation mechanisms for *one-sided matching* in one-shot settings, explain the current theoretical and computational findings, and provide a systematic review of the literature. We also briefly describe the state of the art on dynamic mechanism design for markets with transferable utilities and sketch out its connections to matching mechanisms in dynamic and uncertain domains.

2.1 Basic Definitions and Properties

We first provide a general definition for matching problems and explain our desired properties for matching and resource allocation in multiagent settings, in words. We relegate the formal description of matching models and the detailed formulations of the desired properties to the respective chapters.

The *one-sided matching problem* is a fundamental, yet widely applicable, problem in economics and computer science [16, 30, 82, 116].[1] The goal in a matching problem is to allocate a set of discrete and indivisible objects to a set of self-interested agents. Consider a set of M objects, with $|M| = m$, and N agents, where $|N| = n$, with private ordinal preferences over the objects. A matching (or an assignment) is a mapping between agents and objects that indicates which objects would be allocated to each of the agents. A

[1] In the literature, these problems are also known as *assignment problems* or *house allocation* problems. In this thesis, we use these terms interchangeably and sometimes refer to such settings as matching problems, for short.

matching mechanism is a procedure that determines the allocation of objects to agents according to agents' preferences.

For clarity, we define some of the main properties that are generally used in multiagent settings to assess the desirability of matching mechanisms. In the next chapters, we will introduce our formal model and provide the necessary formalisms to define each of these properties:

- **Pareto efficiency**: an allocation is Pareto efficient (or Pareto optimal) if there exists no other allocation that makes at least one agent better off without making any of the agents worse off.
- **Non-bossiness**: a matching mechanism is non-bossy if no agent can change the allocation of another agent without making its own allocation worse off.
- **Strategyproofness**: a matching mechanism is said to be strategyproof if no agent can benefit from misreporting its preference.
- **Neutrality**: a matching mechanism is neutral if it does not depend on the name of the objects, *i.e.* changing the name of some objects results in a one-to-one identical change in the final allocation.
- **Proportionality**: a matching mechanism is proportional (or symmetric) if all equal agents (with equal preferences) are treated equally.
- **Envyfreeness**: a matching mechanism is envyfree if no agent prefers the allocation of another agent to its own allocation.

In the next section, we elaborate on some of the main results in the area of matching and discuss how some of the above properties remain incompatible in multiagent domains.

2.2 Deterministic Mechanisms

In one-shot settings, the problem of allocating indivisible objects to agents can be categorized into two main settings: standard assignment problems and multiple assignment problems.

In the *standard assignment* problem (sometimes known as the house allocation problem), each agent is entitled to receive exactly one object from the market. Svensson [133, 134] formulated the *standard assignment problem* (first proposed by Shapley and Scarf [126]) where each agent receives exactly one item, and showed that Serial Dictatorship mechanisms are the only social choice rules that satisfy strategyproofness, non-bossiness, and neutrality. Pápai [108] extended the standard model of Svensson [133, 134] to settings where there are potentially more objects than agents (each agent receiving at most one object) with a hierarchy of endowments, generalizing Gale's top trading cycle procedure. This result showed that the hierarchical exchange rules characterize the set of all Pareto efficient, group-strategyproof, and reallocation-proof mechanisms.

In the *multiple-assignment* problem, agents may receive sets of objects, and thus, might have various interesting preferences over the bundles of objects. Pápai [110] studied this problem on the domain of strict preferences allowing for complements and substitutes, and showed that sequential dictatorships are the only strategyproof, Pareto optimal, and non-bossy mechanisms. Ehlers and Klause [55] restricted attention to responsive and separable preferences and essentially proved that the same result persists even in a more restrictive setting. Furthermore, Ehlers and Klause showed that considering resource monotonic allocation rules, where changing the available resources (objects) affects all agents similarly, limits the allocation mechanisms to serial dictatorships.

Pápai [109] and Hatfield [72] studied the multiple assignment problem where objects are assigned to agents subject to a quota. Pápai [109] showed that under quantity-monotonic preferences every strategyproof, non-bossy, and Pareto efficient social choice mechanism is sequential; while generalizing to monotonic preferences, the class of such social choice functions gets restricted to *quasi-dictatorial* mechanisms where every agent except the first dictator is limited to pick at most one object. Pápai's characterization is essentially a negative result and rules out the possibility of designing neutral, non-bossy, strategyproof, and Pareto efficient mechanisms that are not strongly dictatorial. Restricting the preferences to responsive preferences only and when agents have precisely fixed and equal quotas, Hatfield [72] showed that the only strategyproof, Pareto efficient, non-bossy, and neutral mechanisms are serial dictatorships. Figure 2.1 illustrates some of the most important results in the assignment problems.

In the domain of cardinal utilities, the leading fairness concept of Competitive Equilibrium from Equal Incomes (CEEI) efficiently assigns objects to agents and is always guaranteed to exist for divisible objects [139]. However, a CEEI solution may not always exist for indivisible objects. Moreover, the CEEI solution concept and its randomized counterpart based on pseudo-markets by Hylland and Zeckhauser [82], are highly susceptible to manipulation. Zhou [145], based on Gale's conjecture [61], proved that there do not exist (randomized) allocation rules that satisfy symmetry, Pareto efficiency, and strategyproofness.

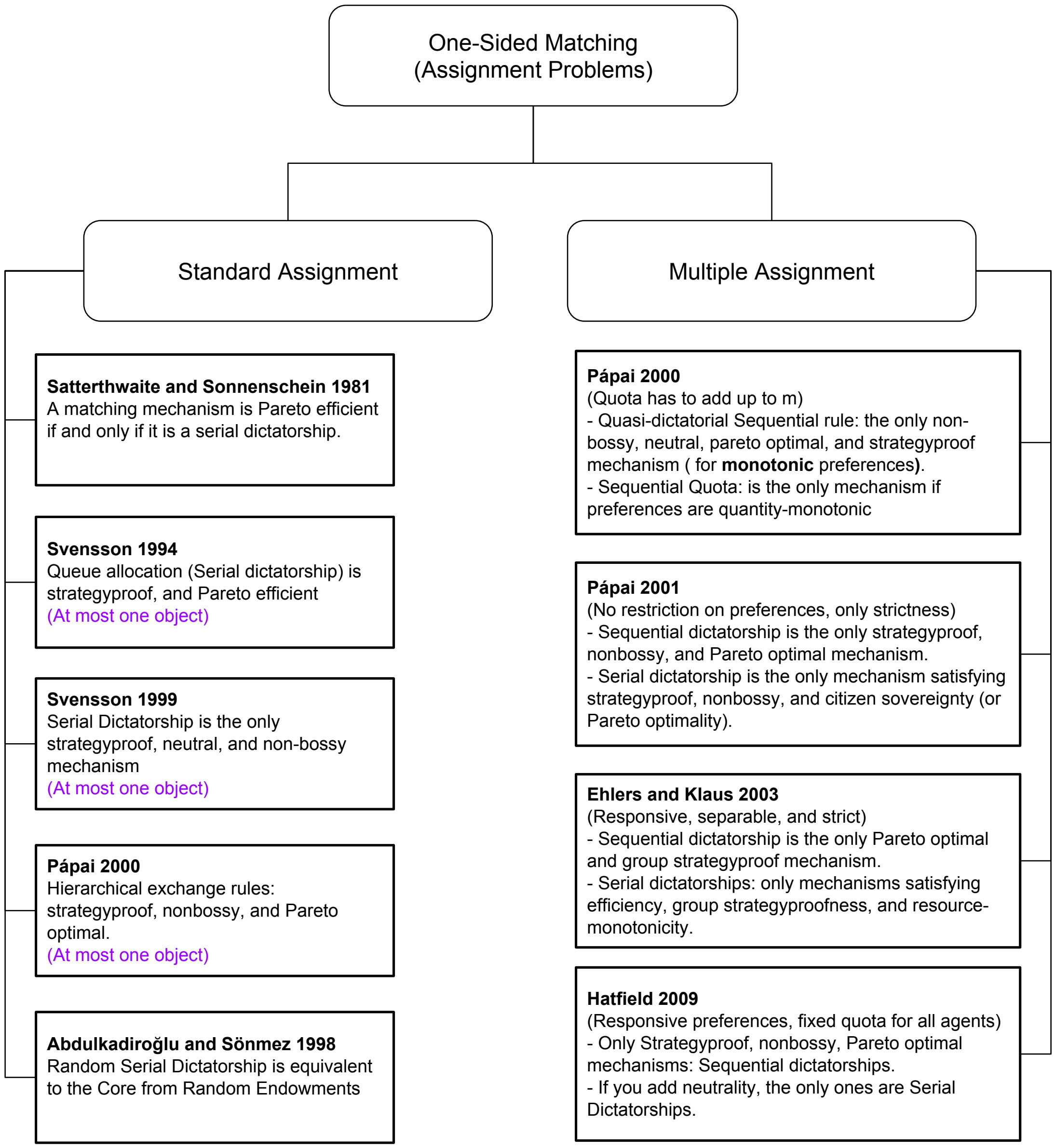


Figure 2.1: Characterization of standard and multiple assignment mechanisms.

2.3 Randomized Ordinal Mechanisms

Randomized matching mechanisms provide a feasible way to fairly assign scarce resources to agents. In such settings, Random Serial Dictatorship (RSD) and Probabilistic Serial Rule (PS) are well-known for their prominent economic properties. RSD assigns the alternatives to agents by choosing a priority ordering of agents uniformly at random, and then running a serial dictatorship mechanism where each agent picks its desired object from the pool of available objects according to its preference ordering. PS, on the other hand, simulates a simultaneous eating algorithm by assuming that resources are divisible. Agents start consuming from their most desired objects, and move to their next preferred objects upon exhausting each object. The final fractional allocations can be seen as probabilities for assigning indivisible objects to each agent.

RSD satisfies strategyproofness, ex post efficiency, and equal treatment of equals [6], while PS is ordinally efficient and envyfree but not strategyproof [30]. For divisible objects, Schulman and Vazirani [124] showed that if agents have lexicographic preferences, the Probabilistic Serial rule is strategyproof under strict conditions over the minimum available quantity of objects and the maximum demand request of agents. Under indivisible objects, these strict requirements translate to situations where the number of agents is greater than the number of objects and each agent receives at most one object. When allocating multiple objects to agents, Kojima [92] obtained negative results on (weak) strategyproofness of PS in the general domain of preferences

Random Serial Dictatorship (RSD) satisfies strategyproofness, proportionality, and ex post efficiency [6]. Bogomolnaia and Moulin noted the inefficiency of RSD from the ex ante perspective, and characterized the matching mechanisms based on *first-order stochastic dominance* [30]. They proposed the probabilistic serial mechanism as an efficient and envyfree mechanism with regards to ordinal preferences. While PS is not strategyproof, it satisfies weak strategyproofness for problems with equal number of agents and objects. However, PS is strictly manipulable (not weakly strategyproof) when there are more objects than agents [92]. Kojima and Manea, showed that in large assignment problems with sufficiently many copies of each object, truth-telling is a weakly dominant strategy in PS [93]. In fact PS and RSD mechanisms become equivalent [44], that is, the inefficiency of RSD and manipulability of PS vanishes when the number of copies of each object approaches infinity.

The practical implications of deploying RSD and PS have been the center of attention in many one-sided matching problems [3,104]. In the school choice setting with multi-capacity alternatives, Pathak observed that many students obtained a more desirable random assignment through PS in public schools of New York City [113]; however, the efficiency difference was quite small. These equivalence results and their extensions to all random mechanisms [97], do not hold when the quantities of each object is limited to one.

The incompatibility of efficiency and strategyproofness gave rise to a more relaxed notion of strategyproofness for assignment mechanisms [102,105]. These mechanisms provide partial incentives by adjusting a parameter that sets the degree of strategyproofness and can be extended to hybrid mechanisms that facilitate the design of partially efficient and strategyproof random mechanisms [104].

The utilitarian and egalitarian welfare guarantees of RSD have been studied under ordinal and linear utility assumptions [18, 26]. For arbitrary utilities, RSD provides the best approximation ratio for utilitarian social welfare when $m = n$ among all mechanisms that rely only on ordinal preferences [57].

2.4 Computation

Due to the practical implications of economics models in real life, the goal in multiagent settings is to design allocation mechanisms that are simple and easy to implement. Most matching mechanisms that are developed in the literature such as RSD, PS, and Draft mechanisms are simple to deploy. However, computing allocation probabilities (also known as fractional assignments) in some randomized mechanisms can be computationally hard. For example, probabilities that are assigned by PS can be easily computed in polynomial time via the well-known "Simultaneous Eating" algorithm while computing RSD probabilities have shown to be #P-hard [16, 121].

The equilibrium notions and best-response strategies under PS have recently been studied [20, 56]. Aziz et al. showed that computing expected utility best response is NP-hard, but for the case of two agents or when preferences are lexicographic, manipulation can be done in polynomial time [20]. Moreover, they proved the existence of pure Nash equilibria, but showed that computing an equilibrium is NP-hard [19]. Nevertheless, Mennle et al. [106] showed that agents can easily find near-optimal strategies by simple local and greedy search. In the absence of truthful incentives, the outcome of PS is no longer guaranteed to be efficient or envyfree with respect to agents' true underlying preferences, and this inefficiency may result in outcomes that are worse than RSD, especially in 'small' markets [56].

2.5 Dynamic Mechanisms

The temporal nature of the decisions in most allocation settings gives rise to two types of uncertainties: (1) uncertainty about the population of agents and (2) uncertainty about the types and preferences of the agents. Dynamic mechanism design [111] is a compelling research area that has attracted attention in recent years and addresses these type of uncertainties in dynamic settings.

Markets with dynamic population of agents have been studied for their incentive and efficiency properties in various contexts such as campus housing and organ transplant, which assume time-invariant preferences [27, 96, 138]. In these markets, agents arrive and depart stochastically and the outcome of matching mechanisms depends heavily on the reported preferences of agents as well as private information about agents' departure and arrival times [2,8,68,96,132]. In the context of organ donation and kidney exchange, several studies focused on dynamic arrival or departure of patients with multiple (but fixed) types. Dynamic kidney matching, therefore, has been studied extensively for its efficiency, fairness, and computational properties [13, 14, 50, 117, 138]. Bade studied matching problems with endogenous information acquisition, and showed that simple serial dictatorship is the only ex ante Pareto optimal, strategyproof, and non-bossy mechanism when agents reveal their preferences only after acquiring information about those with higher priorities [23].

There are numerous investigations on the incentive compatibility and efficiency of allocation mechanisms in settings where agents' types or preferences are subject to uncertainties. In these settings, agents act to improve their outcomes over time, and decisions both in the present and in the past influence how the world and preferences look in the future. The dynamic pivot mechanism for dynamic auctions [25], dynamic Groves mechanisms [41, 111] and many others [15, 140] are a few of myriad examples of mechanisms in dynamic settings that consider agents with private dynamic preferences. However, almost all of these works (excluding a recent study on dynamic social choice [112]) assume an underlying utility function with possible utility transfers. Despite the interest in matching problems, little has been done in dynamic settings where agents' preferences evolve over time.

2.6 Two-Sided Matching

Lastly, we briefly explain some of the findings in two-sided matching, a closely related matching market without monetary transfers. The desirable properties of matching mechanisms in two-sided markets have been extensively studied [62, 118]. Two-sided matching problems deal with mapping a set of agents to another disjoint set, both having private ordinal preferences over the members of the other set. Although the deferred acceptance algorithm proposed by Gale and Shapley provides a stable and optimal solution for two-sided matching [62], several variations such as the school choice problem with indifferences and college admission with quotas require new mechanisms to satisfy efficiency, strategyproofness, fairness, and stability of the matching markets [1, 60, 66, 95].

The deferred acceptance mechanism have also been studied in dynamic matching problems where preferences of one side evolves over time [88]. Although incentives for manipulation in such dynamic settings approaches to zero under some restrictions in large markets, these settings remain highly manipulable in general settings. In fact, Kennes and Dur prove that there exists no strategyproof and stable mechanism in dynamic matching models [54, 89].

Part I

One-Shot Matching Mechanisms

Chapter 3

Strategyproof Quota Mechanisms for Multiple Assignment Problems

The goal in most multiagent systems is to design mechanisms that efficiently allocate resources and objects to agents based on self-reported preferences. Intelligent agents often behave strategically by reasoning over other agents' preferences and the allocation mechanisms. Hence, it is necessary for mechanism designers of multiagent systems to ensure the quality of outcomes by carefully crafting mechanisms that satisfy certain desirable properties.

In this chapter, we focus attention on the problem of allocating indivisible objects to agents in one-shot, static settings without any explicit market. In many real-life domains such as course assignment, room assignment, school choice, medical resource allocation, etc. the use of monetary transfers or explicit markets are forbidden because of ethical and legal issues [116–118]. Much of the literature in this domain is concerned with designing incentive compatible mechanisms that incentivizes agents to reveal their preferences truthfully. Moreover, the efficiency criterion of Pareto efficiency along with strategyproofness provide stable solutions to such allocation problems.

We are interested in allocation problems where each agent may receive a set of objects and thus we search for mechanisms that satisfy some core axiomatic properties of strategyproofness, Pareto efficiency, and non-bossiness. Examples of such allocation problems include distributing inheritance among heirs[1], allocating multiple tasks to employees, assigning scientific equipment to researchers, assigning teaching assistants to different

[1]Here we only consider non-liquid assets that cannot be quickly or easily converted to transferable assets such as money.

courses, and allocating players to sports teams. The common solution for allocating players to teams or allocating courses to students in the course assignment problem is the Draft mechanism [35], where agents choose one item in each picking round. However, allocation mechanisms, such as the Draft mechanism, have been shown to be highly manipulable in practice and fail to guarantee Pareto optimality [38].

Svensson [133, 134] formulated the *standard assignment problem* (first proposed by Shapley and Scarf [126]) where each agent receives exactly one item, and showed that Serial Dictatorship mechanisms are the only social choice rules that satisfy Pareto efficiency, strategyproofness, non-bossiness, and neutrality. In contrast to the standard assignment problem, in the *multiple assignment problem* agents may require bundles or sets of objects according to a predefined quota and might have various interesting preferences (*e.g.* complements or substitutes) over these sets. However, the class of sequential dictatorships mechanisms no longer characterizes all non-bossy, Pareto efficient, and strategyproof social choice mechanisms. In fact, even with monotonic preferences, the class of such social choice mechanisms gets restricted to *quasi-dictatorial* mechanisms where every agent except the first dictator is limited to pick at most one object [109]. Assuming a fixed quota size for all agents, serial dictatorships are the only mechanisms satisfying strategyproofness, Pareto efficiency, non-bossiness, and neutrality [72].

Our work generalizes these results, for a subclass of preferences, by allowing any number of agents or objects, and assuming that individual agents' quotas can vary and be agent specific, imposing no restrictions on the problem size nor quota structures. Instead, we are interested in expanding the possible quota mechanisms to a larger class, essentially enabling a social planner to choose any type of quota system. Our main focus is on the *lexicographic preference* domain, where agents have idiosyncratic private preferences.

Lexicographic preferences [59] have recently attracted attention among researchers in economics and computer science [91,120,124]. There is plenty of evidence for the presence of lexicographic preferences among individuals for breaking ties among equally valued alternatives [52,130], making purchasing decisions by consumers [47,84,144], and examining public policies, job candidates, etc. [137].

In this vein, allocating indivisible alternatives entails lexicographic choices in many settings. Examples of such include assigning scientific resources or labs to research groups, assigning teaching assistants to instructors, etc. Take the example of assigning teaching assistants to instructors. An instructor requiring three assistants may plan to utilize her team by delegating the most important tasks (let's say teaching tutorials) to Alice (her top choice), perhaps because of past interactions. Thus, she would consider any subset that includes Alice superior to all those that do not assign Alice to her course.

Our main results in the lexicographic preference domain are the following:

- We characterize the set of strategyproof, non-bossy, and neutral allocation mechanisms when there is a quota system. We show that serial dictatorships are the only mechanisms satisfying our required properties of strategyproofness, non-bossiness, Pareto efficiency, and neutrality. Allowing any quota system enables the social planner to remedy the inherent unfairness in deterministic allocation mechanisms by assigning quotas according to some fairness criteria (such as seniority, priority, etc.).

- We generalize our findings to *randomized mechanisms* and show that random serial dictatorship quota mechanisms (RSDQ) satisfy strategyproofness, ex post efficiency, and envyfreeness in the domain of lexicographic preferences. Consequently, we prove that the well-known Random Serial Dictatorship (RSD) mechanism in standard assignment settings satisfies envyfreeness when preferences are lexicographic. Thus, random quota mechanisms provide a rich and extended class for object allocation with no restriction on the market size nor quota structure while providing envyfreeness in the lexicographic domains, justifying the use of such mechanisms in many practical applications.

3.1 Related Work

In the *standard assignment* problem (sometimes known as the house allocation problem), each agent is entitled to receive exactly one object from the market. Pápai [108] extended the standard model of Svensson [133, 134] to settings where there are potentially more objects than agents (each agent receiving at most one object) with a hierarchy of endowments, generalizing Gale's top trading cycle procedure. This result showed that the hierarchical exchange rules characterize the set of all Pareto efficient, group-strategyproof, and reallocation proof mechanisms.

In the *multiple-assignment* problem, agents may receive sets of objects, and thus, might have various interesting preferences over the bundles of objects. Pápai [110] studied this problem on the domain of strict preferences allowing for complements and substitutes, and showed that sequential dictatorships are the only strategyproof, Pareto optimal, and non-bossy mechanisms. Ehlers and Klaus [55] restricted attention to responsive and separable preferences and essentially proved that the same result persists even in a more restrictive setting. Furthermore, Ehlers and Klaus showed that considering resource monotonic allocation rules, where changing the available resources (objects) affects all agents similarly, limits the allocation mechanisms to serial dictatorships.

Pápai [109] and Hatfield [72] studied the multiple assignment problem where objects are assigned to agents subject to a quota. Pápai [109] showed that under quantity-monotonic preferences every strategyproof, non-bossy, and Pareto efficient social choice mechanism is sequential; while generalizing to monotonic preferences, the class of such social choice functions gets restricted to *quasi-dictatorial* mechanisms where every agent except the first dictator is limited to pick at most one object. Pápai's characterization is essentially a negative result and rules out the possibility of designing neutral, non-bossy, strategyproof, and Pareto efficient mechanisms that are not strongly dictatorial. Restricting the preferences to responsive preferences only and when agents have precisely fixed and equal quotas, Hatfield [72] showed that the only strategyproof, Pareto efficient, non-bossy, and neutral mechanisms are serial dictatorships.

Hylland and Zeckhauser's pseudo-market design based on eliciting cardinal utilities [82] and its deterministic counterpart based on competitive equilibrium from equal incomes (CEEI) [139] provide efficient and envyfree solutions but are highly susceptible to manipulation. Zhou [145], based on Gale's conjecture [61], proved that there do not exist (randomized) allocation rules that satisfy symmetry, Pareto efficiency, and strategyproofness.

In the randomized settings, Random Serial Dictatorship (RSD) and Probabilistic Serial Rule (PS) are well-known for their prominent economic properties. RSD satisfies strategyproofness, ex post efficiency, and equal treatment of equals [6], while PS is ordinally efficient and envyfree but not strategyproof [30]. For divisible objects, Schulman and Vazirani [124] showed that if agents have lexicographic preferences, the Probabilistic Serial rule is strategyproof under strict conditions over the minimum available quantity of objects and the maximum demand request of agents. Under indivisible objects, these strict requirements translate to situations where the number of agents is greater than the number of objects and each agent receives at most one object. When allocating multiple objects to agents, Kojima [92] obtained negative results on (weak) strategyproofness of PS in the general domain of preferences, which was later confirmed for lexicographic preferences by Hosseini et al. [78]. In contrast, we seek to find strategyproof and envyfree mechanisms with no restriction on the number of agents or objects under the lexicographic preference domain, addressing the open questions in [109] and in [124] about the existence of a mechanism with more favorable fairness and strategyproofness properties.

3.2 The Model

There is a set of m indivisible objects $M = \{1, \ldots, m\}$ and a set of n agents $N = \{1, \ldots, n\}$. There is only one copy of each object available, and an agent may receive more than one object. Let $\mathbb{M} = \mathbb{P}(M)$ denote the power set of M.

Agents have private preferences over sets of objects. Let $\mathcal{P}$ denote the set of all complete and strict preferences over $\mathbb{M}$. Each agent's preference is assumed to be a strict relation $\succ_i \in \mathcal{P}$.[2] A *preference profile* denotes a preference ordering for each agent and is written as $\succ = (\succ_1, \ldots, \succ_n) \in \mathcal{P}^n$. Following the convention, $\succ_{-i} = (\succ_1, \ldots, \succ_{i-1}, \succ_{i+1}, \ldots, \succ_n) \in \mathcal{P}^n$, and thus $\succ = (\succ_i, \succ_{-i})$.

An allocation is a $n \times m$ matrix $A \in \mathcal{A}$ that specifies a (possibly probabilistic) allocation of objects to agents. The vector $A_i = (A_{i,1}, \ldots, A_{i,m})$ denotes the allocation of agent i, that is,

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,m} \\ A_{2,1} & A_{2,2} & \ldots & A_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \ldots & A_{n,m} \end{pmatrix}$$

We sometimes abuse the notation and use A_i to refer to the set of objects allocated to agent i. Let $\mathcal{A}$ refer to the set of possible allocations. Allocation $A \in \mathcal{A}$ is said to be *feasible* if and only if $\forall j \in M, \sum_{i \in N} A_{i,j} = \{0, 1\}$, no single object is assigned to more than one agent, that is, $A_i \bigcap A_j = \emptyset$ for $i \neq j$, while some objects may not be assigned. Note that we allow *free disposal*, and therefore, $\bigcup_{i \in N} A_i \subseteq M$. For two allocations we write $A_i \succ_i B_i$ if agent i with preferences $\succ_i$ strictly prefers A_i to B_i. Thus, $A_i \succeq_i B_i$ and $B_i \succeq_i A_i$ implies $A_i = B_i$.

Preference $\succ_i$ is *lexicographic* if there exists an ordering of objects, $(a, b, c, \ldots)$, such that for all $A, B \in \mathcal{A}$ if $a \in A_i$ and $a \notin B_i$ then $A_i \succ_i B_i$; if $b \in A_i$ and $a, b \notin B_i$ then $A_i \succ_i B_i$; and so on. That is, the ranking of objects determines the ordering of the sets of objects in a lexicographic manner. Note that lexicographic preferences are *responsive* and *strongly monotonic*. A preference relation is *responsive* if $A_i \bigcup B_i \succ_i A_i \bigcup B'_i$ if and only if $B_i \succ_i B'_i$. Strong monotonicity means that any set of objects is strictly preferred to all of its proper subsets. We make no further assumption over preference relations.

An allocation mechanism is a function $\pi : \mathcal{P}^n \to \mathcal{A}$, which assigns a feasible allocation to every preference profile. Thus, agent i's allocation A_i can also be represented as π_i. An allocation mechanism assigns objects to agents according to a *quota system* q, where

[2] Preference relations are assumed to be complete, transitive, and antisymmetric.

q_i is the quota of the ith dictator such that $\sum_{i=1}^{n} q_i \leq m$. Since in our model not all agents need to be assigned an object, we use the size of quota $|q|$ to denote the number of agents that are assigned at least one object. Thus, we have $|q| \leq n$. From the revelation principle [48], we can restrict our analysis to direct mechanisms that ask agents to report their preferences to the mechanism directly.

3.2.1 Properties

In the context of deterministic assignments, an allocation A *Pareto dominates* another allocation B at $\succ$ if $\exists i \in N$ such that $A_i \succ_i B_i$ and $\forall j \in N$ $A_j \succeq_j B_j$, where $A_i \succ_i B_i$ denotes that agent i strictly prefers allocation A_i over B_i. An allocation is *Pareto efficient* at $\succ$ if no other allocation exists that Pareto dominates it at $\succ$.

Since, a social planner may decide to only assign a $C \leq m$ number of objects, we need to slightly modify our efficiency definition. We say that an allocation that assigns $C = \sum_{i=1}^{n} q_i$ objects is Pareto C-efficient if there exists no other allocation that assigns an equal number of objects, C, that makes at least one agent strictly better off without making any other agent worse off. A Pareto C-efficient allocation is also Pareto efficient when $\sum_{i=1}^{n} q_i = m$.

Definition 3.1 (Pareto C-efficiency)**.** *A mechanism π with quota q, where $C = \sum_i q_i$, is Pareto C-efficient if for all $\succ \in \mathcal{P}^n$, there does not exist $A \in \mathcal{A}$ which assigns C objects such that for all $i \in N$, $A_i \succeq_i \pi_i(\succ)$, and $A_j \succ_j \pi_j(\succ)$ for some $j \in N$.*

A mechanism is strategyproof if there exists no non-truthful preference ordering $\succ_i' \neq \succ_i$ that improves agent i's allocation. More formally,

Definition 3.2 (Strategyproofness)**.** *Mechanism π is strategyproof if for all $\succ \in \mathcal{P}^n$, $i \in N$, and for any misreport $\succ_i' \in \mathcal{P}$, we have $\pi_i(\succ) \succeq_i \pi_i(\succ_i', \succ_{-i})$.*

Although strategyproofness ensures that no agent can benefit from misreporting preferences, it does not prevent an agent from reporting a preference that changes the prescribed allocation for some other agents while keeping her allocation unchanged. This property was first proposed by Satterthwaite and Sonnenschein [122]. A mechanism is *non-bossy* if an agent cannot change the allocation without changing the allocation for herself.

Definition 3.3 (Non-bossiness)**.** *A mechanism is non-bossy if for all $\succ \in \mathcal{P}^n$ and agent $i \in N$, for all $\succ_i'$ such that $\pi_i(\succ) = \pi_i(\succ_i', \succ_{-i})$ we have $\pi(\succ) = \pi(\succ_i', \succ_{-i})$.*

Non-bossiness and strategyproofness only prevent certain types of manipulation; changing another agent's allocation or individually benefiting from a strategic report. However, it may still be possible for two or more agents to form a coalition and affect the final outcome, so that at least one of them improves her allocation ex post.

Definition 3.4 (Group Strategyproofness). *A mechanism π is group-strategyproof if for all $\succ$, there does not exist a subset of agents $N' \subseteq N$ with $\succ'_{N'}$ such that for all $i \in N'$, $\pi_i(\succ'_{N'}, \succ_{N \setminus N'}) \succeq_i \pi_i(\succ)$ and for some $j \in N'$, $\pi_j(\succ'_{N'}, \succ_{N \setminus N'}) \succ_j \pi_j(\succ)$.*

The requirement of group-strategyproofness precludes group manipulation as well as individual agent manipulation.

Our last requirement is neutrality. Let $\phi : M \rightarrow M$ be a permutation of the objects. For all $A \in \mathcal{A}$, let $\phi(A)$ be the set of objects in A renamed according to ϕ. Thus, $\phi(A) = (\phi(A_1), \dots, \phi(A_n))$. For each $\succ \in \mathcal{P}^n$ we also define $\phi(\succ) = (\phi(\succ_1), \dots, \phi(\succ_n))$ as the preference profile where all objects are renamed according to ϕ.

Definition 3.5 (Neutrality). *A mechanism π is neutral if for any permutation function ϕ and for all preference profiles $\succ \in \mathcal{P}^n$, $\phi(\pi(\succ)) = \pi(\phi(\succ))$.*

In other words, a mechanism is *neutral* if it does not depend on the name of the objects, that is, changing the name of some objects results in a one-to-one identical change in the outcome. It is clear that above conditions reduce the set of possible mechanisms drastically.

3.3 Allocation Mechanisms

Several plausible multiple allocation mechanisms exploit interleaving picking orders to incorporate some level of fairness, where agents can take turns each time picking one or more objects [34,36,90]. An interleaving mechanism alternates between agents, allowing a single agent to pick objects in various turns. The *interleaving mechanisms* have been widely used in many everyday life activities such as assigning students to courses, members to teams, and in allocating resources or moving turns in boardgames or sport games. To name a few, *strict alternation* where agents pick objects in alternation (*e.g.* 1212 and 123123) and *balanced alternation* where the picking orders are mirrored (*e.g.* agent orderings 1221 and 123321), and the well-known *Draft mechanism* [35, 36, 38] that randomly chooses a priority ordering over n agents and then alternates over the drawn priority ordering and its reverse sequence are the examples of such mechanisms. However, all these interleaving mechanisms are highly manipulable in theory; computing optimal manipulations under interleaving mechanisms is shown to be easy for a single agent under additive and separable preferences [33]. Extending to non-separable preferences, deciding a strategic picking strategy is NP-complete, even for two agents [33]. Kalinowski et al. [85] studied interleaving mechanisms (alternating policies) from a game-theoretical perspective and showed that under linear order preferences the underlying equilibrium in a two-person picking game is incentive compatible [86]. Nonetheless, such interleaving mechanisms have been shown to be heavily manipulated in practice [38].

We generalize such allocation procedures to any mechanism with an interleaving order of agents with general preferences where at least one agent gets to choose twice, once before and once after one (or more) agents.

Theorem 3.1. *There exists no interleaving mechanism that satisfies Pareto C-efficiency, non-bossiness, and strategyproofness.*

Proof. The proof follows by constructing a manipulable preference profile. Given any Pareto C-efficient and non-bossy interleaving mechanism, we show that we can construct an instance (preference profile) at which at least one agent can manipulate the outcome.

Suppose there is a non-bossy and Pareto C-efficient mechanism π with at least one alternation between agents i and j. Note that the alternation could be through a fixed ordering or through a picking process. Since we are constructing an instance, we can assume that all other agents $k \in N \setminus \{i, j\}$ will receive their objects after agents i and j (or have already received their non-conflicting objects before the two). We now construct a preference profile such that $\succ = (\succ_i, \succ_j, \succ_{N\setminus\{i,j\}})$.

Let f_k denote the agent in the kth picking order, that is, $f_2 = i$ indicates that the agent in the second picking order is agent i. Consider the ordering such that for agents 1 and 2 we have $f_1 = f_3 = 1$ and $f_2 = 2$. Assume there are 3 objects available and construct a preference profile as follows: $\succ_1 = a \succ b \succ c$ and $\succ_2 = o_1 \succ o_2 \succ o_3$, where o_k represent the kth ranked object in $\succ_2$. By Pareto C-efficiency and non-bossiness of π, agents final allocations must preclude any further exchange between the two agents, and no agent can change the allocation of the other while its own allocation remains unchanged.

Since agent 1 is going to pick first and last according to ordering f, agent 1 can pick her first choice either at stage 1 or 3 as long as agent 2's top choice is not equal to agent 1's top choice, *i.e.* $o_1 \in \{b, c\}$. If $o_1 = c$ then there is no conflict between agent 1 and 2 and playing truthfully has the best outcome for agent 1. Thus, it follows that $o_1 = b$ and $o_2 \in \{a, c\}$. Now we only need to construct the rest of agent 2's ordering such that agent 1's top choice, object a, remains in the pool of objects until the last stage. Thus, for the following preference profile $\succ_1 = a \succ b \succ c$ and $\succ_2 = b \succ c \succ a$, the interleaving mechanism is manipulable. This implies that no Pareto C-efficient and non-bossy interleaving mechanism guarantees strategyproofness. □

Clearly, an *imposed mechanism* that assigns a fixed allocation to every preference profile is strategyproof and non-bossy but does not satisfy Pareto C-efficiency [110].[3]

With these essentially negative results for interleaving mechanisms, we restrict our attention to the class of sequential dictatorship mechanisms, where each agent only gets one chance to pick (possibly more than one) objects.

3.3.1 Sequential Mechanisms

Let q denote a quota system such that $\sum_i q_i \leq m$. In a sequential dictatorship mechanism with quota q, the first dictator chooses q_1 of her most preferred objects; the second dictator is chosen depending on the set of objects allocated to the first dictator. The second dictator then chooses q_2 objects of her most preferred objects among the remaining objects. This procedure continues, where the choice of the next dictator may be determined depending on the earlier allocations, until no object or no agent is left.

Let f be a function that, given a partial allocation of objects to some agents, returns the next dictator. Then, $f_i(\cdot) = j$ means that agent j is ranked ith in the ordering of dictators. There exists an agent f_1 (first dictator) for each preference profile $\succ \in \mathcal{M}$, and an ordering of the remaining dictators such that the ith dictator is identified recursively by

$$f_i(\pi_{f_1}(\succ), \ldots, \pi_{f_{i-1}}(\succ))$$

In other words, the choice of the next dictator only depends on the previous dictators and their allocation sets and *does not* depend on the preferences of the previous dictators. The following example shows why the choice of dictator should not depend on previous dictators' preferences.

Example 3.1. Assume three agents and four objects with $q = (2, 1, 1)$ and consider the following rule for identifying the order of the dictators: if the first dictator's preference is $a \succ b \succ c \succ d$ then the ordering of other agents is (2,3), otherwise the order is (3,2). Now if agent 2 and 3 have identical preferences as agent 1, then agent 1 can simply change agent 2 and 3's allocations by misrepresenting her preference as $\hat{\succ}_1 : b \succ a \succ c \succ d$ while her allocation remains unchanged. Thus, this sequential dictatorship mechanism is bossy even though it satisfies Pareto efficiency and strategyproofness.

[3] An imposed mechanism does not take agents' preferences into account and prescribes the same allocation to every preference profile.

Definition 3.6 (Sequential Dictatorship)**.** *Let $\mathbb{M}_k = \mathbb{P}_{\leq k}(M)$ be the set of subsets of M of cardinality less than or equal k. An allocation mechanism $\pi : \mathcal{P}^n \to \mathcal{A}$ is a sequential dictatorship quota mechanism if there exists a quota system q and an ordering f such that for all $\succ \in \mathcal{P}^n$,*

$$\pi_{f_1}(\succ) = \{Z \in \mathbb{M}_{q_1} | Z \succ_1 Z' \text{ for all } Z' \in \mathbb{M}_{q_1}\}$$

$$\pi_{f_i(\pi_{f_1},\ldots,\pi_{f_{i-1}})}(\succ) = \{Z \in \mathbb{M}_{q_i} \setminus \bigcup_{j=1}^{j=i-1} \pi_{f_j}(\succ) | Z \succ_{f_i} Z' \text{ for all } |Z'| = |q_i|\}$$

A **serial dictatorship mechanism** is an example of a sequential mechanism where the ordering is a permutation of the agents, determined a priori, that is, for all $\succ \in \mathcal{P}^n$, $\pi_{f(\cdot)}(\succ) = \pi_f(\succ)$. Such mechanisms satisfy neutrality. From now on, we simply use the vector f instead of $f(\cdot)$ when the ordering is predefined independent of the choice of objects.

3.4 Serial Dictatorship Quota Mechanisms

In this section, we study serial dictatorship mechanisms for quota allocations and characterize the set of strategyproof, non-bossy, neutral, and Pareto efficient mechanisms subject to various quota systems.

When allocating objects sequentially via a quota system q, Pareto C-efficiency requires that no two agents be envious of each others' allocations since then they can simply exchange objects ex post, implying that the initial allocation is dominated by the new allocation after the exchange. For example, take a serial dictatorship with $q_1 = 1$ and $q_2 = 2$ and three objects. Agent 1 will receive her top choice object $\{a\}$ (since $\{a\} \succ_1 \{b\} \succ_1 \{c\}$) according to her preference and agent 2 receives $\{b, c\}$. However, it may be the case that $\{b, c\} \succ_1 \{a\}$ while $\{a\} \succ_2 \{b, c\}$ and both agents may be better off exchanging their allocations. Thus, we have the following proposition for general preferences.

Proposition 3.1. *For general preferences, sequential (and serial) dictatorship quota mechanisms do not guarantee Pareto C-efficiency.*

In the absence of Pareto C-efficiency in the domain of general preferences, a social planner is restricted to use only one type of quota system; either assigning at most one object to all agents except the first dictator (who receives the remaining objects), or setting equal quotas for all agents [72, 109].

Due to the impossibility shown in Proposition 3.1, in the rest of this chapter we restrict ourselves to the interesting class of lexicographic preferences. We show that if preferences are lexicographic, regardless of the selected quota system, any serial dictatorship mechanism guarantees Pareto C-efficiency. We first provide the following lemma in the lexicographic domain.

Lemma 3.1. *The following statements hold for two sets of objects when preferences are lexicographic:*

- *If $B_i \subset A_i$ then $A_i \succ_i B_i$.*
- *For all X such that $X \cap A_i = \emptyset$, we have $A_i \succ_i B_i$ iff $A_i \cup X \succ_i B_i \cup X$.*
- *If $B_i \not\subset A_i$ and $A_i \succ_i B_i$ then there exists an object $x \in A_i$ such that $x \succ_i X$ for all $X \in \mathbb{P}(B_i - A_i)$.*

Proof. We provide proof for each of the statements in the lexicographic domain.

- Since $B_i \subset A_i$ then all objects in B_i are also in A_i, and there exists an object $x \in A_i$ such that $x \notin B_i$. By the definition of lexicographic preferences, having an object is preferred to not having the object (*i.e.* objects are goods). Therefore, $A_i \succ_i B_i$.
- It is easy to see that adding a set of object $X \cap A_i = \emptyset$ to two sets such that $A_i \succ_i B_i$ maintains the preference over the two sets. This is because elements in X are added to both sets and by assumption there is still an element $x \in A_i$ and $x \notin X$ that is preferred to all objects in B_i. We should prove the converse that if $A_i \cup X \succ_i B_i \cup X$ then $A_i \succ_i B_i$. Suppose not, that is $B_i \succeq_i A_i$. By adding $X = B_i - A_i$ to both sides we have $B_i \cup X \succeq_i A_i \cup X$, that is, $B_i \succeq_i A_i \cup B_i$, which contradicts the strong monotonicity of lexicographic preferences when A_i is nonempty.
- Suppose that there does not exist an object $x \in A_i$ such that $x \succ_i X$ for all $X \in \mathbb{P}(B_i - A_i)$. The set X can be any power set of $B_i - A_i$, and for the sake of this proof we assume that $X = B_i - A_i$. By the second statement in this lemma, for $A_i \succ_i B_i$, we can add any X such that $X \cap A_i = \emptyset$ to the both sides and write $A_i \cup X \succ_i B_i \cup X$, which holds since $X = B_i - A_i$. This states that for any object $x \in B_i$, x is also a member of $A_i \cup X$, implying that $B_i \subset A_i \cup X$. Note that $B_i \neq A_i \cup X$ because A_i is considered to be nonempty. Using the first statement in this lemma, if $B_i \subset A_i \cup X$ then $A_i \cup X \succ_i B_i$. Replacing X with $B_i - A_i$ and subtracting it from both sides, we have $A_i \succ_i \emptyset$, which implies that there exists an object $x \in A_i$ such that $x \notin B_i$ and $x \succ_i B_i - A_i$, contradicting the initial assumption.

The above items conclude our proof for the statements in this lemma. □

Proposition 3.2. *If preferences are lexicographic, the serial dictatorship quota mechanism is Pareto C-efficient.*

Proof. Consider a mechanism π with quota q, that assigns $C = \sum_i q_i$ objects. Suppose for contradiction that there exists an allocation B with arbitrary quota q', where $C' = \sum_i q'_i$, that Pareto dominates $A = \pi(\succ)$. We assume $C' = C$ to ensure that both allocations assign equal number of objects (Otherwise by strong monotonicity of lexicographic preferences and Lemma 3.1 one can assign more objects to strictly improve some agents' allocations.).

Thus, for all agents $j \in N$, $B_j \succeq_j A_j$, and there exist some agent i where $B_i \succ_i A_i$. If for all $j \in N$, $|B_j| \geq |A_j|$ then $q'_j \geq q_j$. Now suppose for some i, $|B_i| > |A_i|$. This implies that $q'_i > q_i$. By adding these inequalities for all agents we have $\sum_i q'_i > \sum_i q_i$, contradicting the initial assumption of equal quota sizes ($C' = C$).

For the rest of the proof, we consider two cases; one where the size of B_i is greater than that of A_i, *i.e.*, $|B_i| > |A_i|$, and one where $|B_i| \leq |A_i|$.

Case I: Consider $|B_i| \leq |A_i|$ and $B_i \succ_i A_i$. If $B_i \subset A_i$ then monotonicity of lexicographic preferences in Lemma 3.1 implies that $A_i \succ_i B_i$ contradicting the assumption. On the other hand, if $B_i \not\subset A_i$ by Lemma 3.1 there exists an object $x \in B_i$ such that for all $X \in \mathbb{P}(B_i - A_i)$ agent i ranks it higher than any other subset, that is, $x \succ_i X$. In this case, serial dictatorship must also assign x to agent i in A_i, which is a contradiction.

Case II: Consider $|B_i| > |A_i|$ and $B_i \succ_i A_i$. The proof of this case heavily relies on the lexicographic nature of preferences (as opposed to Case I that held valid for the class of monotonic, and not necessarily lexicographic, preferences). The inequality $|B_i| > |A_i|$ indicates that $q'_i > q_i$. We construct a preference profile $\succ'$ as follows: for each $j \in N$, if $B_j = A_j$ then $\succ'_j = \succ_j$, otherwise if $B_j \neq A_j$ rank the set B_j higher than A_j in $\succ'_j$ ($\succ'_j = B_j \succ A_j \succ \ldots$). Now run the serial dictatorship on $\succ'$ with quota q. Suppose that $B' = \pi(\succ')$. For agent i, B'_i is the top q_i objects of B_i where $B'_i \subsetneq B_i$ and because q_i is fixed, then $|B'_i| = |A_i|$. Given $\succ'$ we have $B_i \neq A_i$, which implies that $B'_i \neq A_i$. By strong monotonicity for agent i we have $B_i \succ_i B'_i \succ_i A_i$. However, according to the constructed quotas we have $|B_i| > |B'_i|$ but $|B'_i| = |A_i|$, where $B'_i \neq A_i$. By Lemma 3.1 there exists an object $x \in B'_i$ which is preferred to all proper subsets of $A_i - B_i$. However, if such object exists it should have been picked by agent i in the first place, which is in contradiction with agent i's preference. This concludes our proof. □

Under lexicographic preferences a social planner can use variants of quotas based on different needs (fairness, seniority, etc.). The ability to use a variety of quotas is crucial in many multiagent systems. Suppose that a planner wants to assign a set of services to a group of agents while ensuring that those with lower priority receive more services or goods to (partially) compensate for undesirable nature of not being ordered first. In this case, a social planner may set a quota such that those who are ordered last, at least receive more objects.

We state a few preliminary lemmas before proving our main result in characterizing the set of non-bossy, Pareto C-efficient, neutral, and strategyproof mechanisms. Given a non-bossy and strategyproof mechanism, an agent's allocation is only affected by her predecessor dictators. Thus, an agent's allocation may only change if the preferences of one (or more) agent with higher priority changes.

Lemma 3.2. *Take any non-bossy and strategyproof mechanism π. Given two preference profiles $\succ, \succ' \in \mathcal{P}^n$ where $\succ = (\succ_i, \succ_{-i})$ and $\succ' = (\succ_i, \succ'_{-i})$, if for all $j < i$ we have $\pi_{f_j}(\succ) = \pi_{f_j}(\succ')$, then $\pi_{f_i}(\succ) = \pi_{f_i}(\succ')$.*

Proof. For all $j < i$ we have $\pi_{f_j}(\succ) = \pi_{f_j}(\succ')$. By non-bossiness and strategyproofness, for all $\succ'_j$ such that $\pi_j(\succ) = \pi_j(\succ'_j, \succ_{-j})$ we have $\pi(\succ) = \pi(\succ'_j, \succ_{-j})$. In words, non-bossiness and strategyproofness prevent any agent to change the allocation of other agents with lower priority (those who are ordered after him), without changing its own allocation.

Let M' be the set of remaining objects such that $M' = M \setminus \bigcup_{k=1}^{j} \pi_{f_k}(\succ)$. Since $\pi_{f_j}(\succ) = \pi_{f_j}(\succ')$, the set of remaining objects M' under $\succ'$ is equivalent to those under $\succ$, implying that $\pi_{f_i}(\succ) = \pi_{f_i}(\succ')$ which concludes the proof. □

The following Lemma guarantees that the outcome of a strategyproof and non-bossy mechanism only changes when an agent states that some set of objects that are less preferred to $\pi_i(\succ)$ under $\succ_i$ is now preferred under $\succ'_i$. Intuitively, any preference ordering $\succ'_i$ which reorders only the sets of objects that are preferred to $\pi_i(\succ)$ or the sets of objects that are less preferred to the set of objects allocated via $\pi_i(\succ)$ keeps the outcome unchanged.

Lemma 3.3. *Let π be a strategyproof and non-bossy mechanism, and let $\succ, \succ' \in \mathcal{P}^n$. For all allocations $A \in \mathcal{A}$, if for all $i \in N, \pi_i(\succ) \succeq_i A_i$ and $\pi_i(\succ) \succeq'_i A_i$, then $\pi(\succ) = \pi(\succ')$.*

Proof. The proof follows similar to Lemma 1 in [134]. First, we show that $\pi(\succ'_i, \succ_{-i}) = \pi(\succ)$, that is changing i's preference only does not affect the outcome. From strategyproofness we know that $\pi_i(\succ_i) \succeq_i \pi_i(\succ'_i, \succ_{-i})$. By assumption of the lemma we can also write $\pi_i(\succ_i) \succeq'_i \pi_i(\succ'_i, \succ_{-i})$. However, strategyproofness implies that $\pi_i(\succ'_i, \succ_{-i}) \succeq'_i \pi_i(\succ_i)$. Since the preferences are strict, the only way for the above inequalities to hold is when $\pi_i(\succ'_i, \succ_{-i}) = \pi_i(\succ)$. The non-bossiness of π implies that $\pi(\succ'_i, \succ_{-i}) = \pi(\succ)$.

We need to show that the following argument holds for all agents. We do this by partitioning the preference profile into arbitrary partitions constructed partly from $\succ$ and partly from $\succ'$. Let $\succ^p = (\succ'_1, \ldots, \succ'_{p-1}, \succ_p, \ldots, \succ_n) \in \mathcal{P}^n$. Thus, a sequence of preference profiles can be recursively written as $\succ^{p+1} = (\succ'_p, \succ^p_{-p})$.

Using the first part of the proof and by the recursive representation, we can write $\pi(\succ^p) = \pi(\succ'_p, \succ^p_{-p}) = \pi(\succ^{p+1})$. Now using this representation, we shall write $\pi(\succ') = \pi(\succ^{n+1})$ and $\pi(\succ) = \pi(\succ^1)$, which implies that $\pi(\succ) = \pi(\succ')$. □

The next lemma states that when all agents' preferences are identical, any strategyproof, non-bossy, and Pareto C-efficient mechanism simulates the outcome of a serial dictatorship quota mechanism.

Lemma 3.4. *Let π be a strategyproof, non-bossy, and Pareto C-efficient mechanism with quota system q, and $\succ$ be a preference profile where all individual preferences coincide, that is $\succ_i = \succ_j$ for all $i, j \in N$. Then, there exists an ordering of agents, f, such that for each $k = 1, \ldots, |q|$, agent f_k receives exactly q_k items according to quota q induced by a serial dictatorship.*

Proof. Suppose the contrary and let $\succ$ be an identical preference profile $\succ_1 = \succ_2 = a \succ b \succ c$ such that agent 1 receives a and c while agent 2 receives b. For agents 1 and 2, assume that they both have received no other objects except the ones stated above (Alternatively, we can assume that the other objects received by these two agents so far are their highest ranked objects, and because these objects were assigned in some previous steps, they won't affect the assignment of the remaining objects). For all other agents $N \setminus \{1, 2\}$ assume that the allocation remains unchanged, *i.e.*, these agents will receive exactly the same objects after we change the preferences of agent 1. By Lemma 3.3, since the mechanism is non-bossy and strategyproof, agent 1's allocation remains unchanged under the following changes in its preference ordering:

$$\succ_1 = a \succ b \succ c \Rightarrow a \succ c \succ b \Rightarrow c \succ a \succ b$$

Thus, the new preference profile $\succ'$ would be

$$\begin{aligned} \succ'_1: &\quad \boxed{c} \succ \boxed{a} \succ b \\ \succ_2: &\quad a \succ \boxed{b} \succ c \end{aligned}$$

where $\pi(\succ') = \pi(\succ)$. The squares show the current allocation. Since agent 1 is receiving two objects and agent 2 receives one, for any ordering that is not prescribed by a serial dictatorship, agent 2 should be ordered second (otherwise, the ordering is a serial dictatorship).

More specifically, orderings (1,2) and (2,1) are serial dictatorships. Since agent 2 must be ordered second, it must be the case that agent 1 goes first and third (otherwise we are back at (1,2), which results in a serial dictatorship). Agent 1 first chooses object c according to $\succ'_1$, then agent 2 chooses object a according to $\succ_2$, and lastly agent 1 chooses the remaining object b. Therefore, agent 2 can benefit from manipulating the mechanism by choosing a instead of b, contradicting the assumption that π is strategyproof and non-bossy. This implies that such agents cannot exist, and concludes our proof. □

Theorem 3.2. *If preferences are lexicographic, an allocation mechanism is strategyproof, non-bossy, neutral, and Pareto C-efficient if and only if it is a serial dictatorship quota mechanism.*

Proof. It is clear that in the multiple-assignment problem any serial dictatorship mechanism is strategyproof, neutral, and non-bossy [72, 110]. For Pareto efficiency, in Proposition 3.2, we showed that the serial dictatorship mechanism is Pareto C-efficient for any quota, and in fact it becomes Pareto efficient in a stronger sense when all objects are allocated $C = m$.

Now, we must show that any strategyproof, Pareto C-efficient, neutral, and non-bossy mechanism, π, can be simulated via a serial dictatorship quota mechanism. Let π be a strategyproof, Pareto C-efficient, neutral, and non-bossy mechanism. Consider $\succ \in \mathcal{P}^n$ to be an arbitrary lexicographic preference profile. Given q, we want to show that π is a serial dictatorship mechanism. Thus, we need to find an ordering f that induces the same outcome as π when allocating objects serially according to quota q.

Take an identical preference profile and apply the mechanism π with a quota q. By Lemma 3.4, there exists a serial dictatorial allocation with an ordering f where agent f_1 receives q_1 of her favorite objects from M, agent f_2 receives q_2 of her best objects from $M \setminus \pi_{f_1}$, and so on. Therefore, given a strategyproof, non-bossy, neutral, and Pareto C-efficient mechanism with quota q, we can identify an ordering of agents $f = (f_1, \ldots, f_n)$ that receive objects according to $q = (q_1, \ldots, q_n)$. Note that since the ordering is fixed a priori, the same f applies to any non-identical preference profile.

From any arbitrary preference profile $\succ$, we construct an equivalent profile as follows: Given the ordering f, the first best q_1 objects (the set of size q_1) according to $\succ_{f_1}$ are denoted by A_{f_1} and are listed as the first objects (or set of objects of size q_1 since preferences are lexicographic) in $\succ'_i$. The next q_2 objects in $\succ'_i$ are the first best q_2 objects according to $\succ_{f_2}$ from $M \setminus A_{f_1}$, and so on. In general, for each $i = 2, \ldots, |q|$, the next best q_i objects

Algorithm 1: Constructing an identical preference profile

```
Data: A preference profile ≻, an ordering f, and quota q
Result: A profile with identical preferences ≻′ with π(≻′) = π(≻)
Initialize ≻_1 ← ∅
Initialize set Z = ∅
for (i ← 1 to |q|) do
    Z ← top(q_i, ≻_{f_i}) // Most preferred set of size q_i from the remaining objects.
    ≻′_1 ← append(≻′_1, Z) // Append this set to the preference ordering.
    Z ← ∅
for (i ← 1 to |f|) do
    ≻′_i ← ≻′_1
return ≻′.
```

are the best q_i objects according to $\succ_{f_i}$ from $M \setminus \bigcup_{j=1}^{j=i-1} A_j$. Algorithm 1 illustrates these steps.

Now we need to show that applying π to the constructed identical preference profile ($\succ'$) induces the same outcome as applying it to $\succ$. By Lemma 3.2 for each agent f_i, $\pi_{f_i}(\succ) = \pi_{f_i}(\succ')$ if for all $j < i$ we have $\pi_{f_j}(\succ) = \pi_{f_j}(\succ')$. That is, the allocation of an agent remains the same if the allocations of all previous agents remain unchanged. Now by Lemma 3.3, for any allocation $A \in \mathcal{A}$, if for each agent $i \in N$, $\pi_i(\succ') \succeq'_i A_i$ then we also have $\pi_i(\succ') \succeq_i A_i$. For each f_i where $i = 1, \ldots, |q|$, by Lemma 3.3 since π is strategyproof and non-bossy, for any allocation A_{f_i} given the quota q we have $\pi_{f_i} \succeq'_{f_i} A_{f_i}$ and $\pi_{f_i} \succeq_{f_i} A_{f_i}$, which implies that $\pi_{f_i}(\succ') = \pi_{f_i}(\succ)$. Therefore, we have $\pi(\succ') = \pi(\succ)$. Since $\succ'$ is an identical profile, $\pi(\succ') = \pi(\succ)$ assigns q_i objects to each agent according to the serial ordering f. Thus, π is a serial dictatorship quota mechanism. □

The following example illustrates how an equivalent preference profile with identical outcome is constructed given any arbitrary preference profile, ordering, and quota system.

Example 3.2. Consider allocating 4 objects to 3 agents with preferences illustrated in Table 3.1 (left), based on the following quota $q = (1, 2, 1)$. Assume the following ordering of agents $f = (1, 2, 3)$. To construct a profile with identical orderings, agent 1's first best object according to $\succ_1$, a, is considered the highest ranking object in $\succ'_i$. Agent 2's best two objects ($q_2 = 2$) among the remaining objects c and b are ranked next, and finally agent 3's remaining object d is ranked last. Given f and q, the two preference profiles depicted in Table 3.1 have exactly similar outcome (shown with squares).

$\succ_1$: $\boxed{a} \succ b \succ c \succ d$		$\succ'_1$: $\boxed{a} \succ c \succ b \succ d$
$\succ_2$: $\boxed{c} \succ a \succ \boxed{b} \succ d$		$\succ'_2$: $a \succ \boxed{c} \succ \boxed{b} \succ d$
$\succ_3$: $a \succ c \succ \boxed{d} \succ b$		$\succ'_3$: $a \succ c \succ b \succ \boxed{d}$

Table 3.1: Converting a preference profile to identical orderings, with exact same outcome.

Finally, we show that strategyproofness and non-bossiness are necessary and sufficient conditions for group-strategyproofness.

Proposition 3.3. *A mechanism is group-strategyproof if and only if it is strategyproof and non-bossy.*

Proof. It is easy to see that group-strategyproofness implies strategyproofness and non-bossiness. We need to show the converse, that is, if π is strategyproof and non-bossy then it is group-strategyproof.[4] Let $N' \subseteq N$ be a subset of agents, $N' = \{1, \ldots, n'\}$, with $\succ'_{N'}$ such that for all $i \in N'$ $\pi_i(\succ'_{N'}, \succ_{-N'}) \succeq_i \pi_i(\succ)$. Construct an alternative preference profile $\hat{\succ}$ such that for all $i \in N'$ the preference ordering $\hat{\succ}_i$ preserves the ordering but moves the set $\pi_i(\succ'_{N'}, \succ_{-N'})$ to the first rank in the ordering.

For agent 1, if $\pi_1(\succ'_{N'}, \succ_{-N'}) \succ_1 \pi_1(\succ)$ then by Lemma 3.2, $\pi_1(\succ'_{N'}, \succ_{-N'})$ is not in the list of available sets. Otherwise, $\pi_1(\succ'_{N'}, \succ_{-N'}) = \pi_1(\succ)$. Thus, strategyproofness implies that $\pi_1(\hat{\succ}_1, \succ_{-1}) = \pi_1(\succ)$, and by non-bossiness we have $\pi(\hat{\succ}_1, \succ_{-1}) = \pi(\succ)$. Repeating the same argument for all other agents in $\{2, \ldots, n'\}$, we get $\pi(\hat{\succ}_{N'}, \succ_{-N'}) = \pi(\succ)$. Now since π is strategyproof and non-bossy, using Lemma 3.3 we have that $\pi(\hat{\succ}_{N'}, \succ_{-N'}) = \pi(\succ'_{N'}, \succ_{-N'})$. This implies that $\pi(\succ_{N'}, \succ_{-N'}) = \pi(\succ)$, meaning that π is group-strategyproof. □

Note that group-strategyproofness prevents any subset of agents from misreporting their preferences and gaining as a group. It is critical to note that under group-strategyproofness agents in a group cannot gain by misreporting, without being able to swap their objects after the final allocation. Thus, a group-strategyproof mechanism does not rule out the possibility of manipulation by a subset of agents that misreport their preferences and then exchange their allocations ex post. This type of coalitional manipulation requires that two agents can swap their objects after being assigned by the mechanism. A property that rules out manipulation by exchange of objects ex post is called *reallocation-proofness* [108].

The following example illustrates a mechanism that is group-strategyproof but does not guarantee reallocation-proofness.

[4]The proof is inspired by Lemma 1 in [108] for single-object allocation.

$\succ_1$:	$\boxed{a} \succ c \succ b$
$\succ_2$:	$\boxed{c} \succ b \succ a$
$\succ_3$:	$c \succ a \succ \boxed{b}$

Table 3.2: An example showing a mechanism that is group-strategyproof but not reallocation-proof.

Example 3.3. Consider three agents with preferences as shown in Table 3.2. A serial dictatorship mechanism with ordering $f = (1, 2, 3)$ assigns objects to agents as shown with squares. Given the serial dictatorship, none of the subset of agents benefit from misreporting their preferences since the serial dictatorship mechanism is non-bossy and strategyproof. However, if agents are able to exchange objects ex post, agent 1 and 3 can form a coalition and strategically report preferences as $\succ_1'$: $c \succ a \succ b$ and $\succ_3'$: $a \succ c \succ b$. Agent 1 receives object c, agent 2 receives b, and agent 3 receives object a. Now, after the allocation is complete, if agents 1 and 3 swap their assignments, they both receive their top choices, and thus, benefit from this type of reallocation manipulation.

Consequently, it is easy to see that serial dictatorship quota mechanisms are guaranteed against group manipulation but do not prevent coalitional manipulation through reallocation. We rewrite Theorem 3.2 as the following:

Theorem 3.3. *Serial dictatorship quota mechanisms are the only neutral, Pareto C-efficient, and group-strategyproof mechanisms.*

3.5 Sequential Dictatorship Quota Mechanisms

In this section, we study a broader class of quota mechanisms by relaxing the neutrality requirement and allowing for the dictators to be identified in each sequence, as opposed to fixing the dictatorship orderings apriori.

Proposition 3.4. *A sequential dictatorship quota mechanism is Pareto C-efficient under lexicographic preferences.*

Proof. The proof exactly follows as of the proof of Proposition 3.2. □

Characterizing the set of strategyproof, non-bossy, and Pareto C-efficient quota mechanisms is similar to our characterization for serial dictatorship mechanisms, but requires a subtle change in Lemma 3.4.

Lemma 3.5. *Let π be a strategyproof, non-bossy, and Pareto C-efficient mechanism with quota q, and $\succ$ be a preference profile where all individual preferences coincide, that is $\succ_i = \succ_j$ for all $i, j \in N$. Then, there exists an ordering $f_1, f_2(\pi_{f_1}(\succ)), \ldots, f_k(\pi_{f_1}(\succ), \ldots, \pi_{f_{k-1}}(\succ))$ such that for each $i \in N$ agent i receives exactly q_i items according to quota q.*

Proof. Let π be a strategyproof, non-bossy, and Pareto C-efficient mechanism with quota q. By Lemma 3.4, we know that for each identical preference profile, there exists a fixed ordering $f' : (f'_1, \ldots, f'_k)$ such that agent f'_1 receives q_1 objects, agent f'_2 receives q_2, and so on. Let f be a dictatorship ordering such that $f_1, f_2(\pi_{f_1}(\succ)), \ldots, f_k(\pi_{f_1}(\succ), \ldots, \pi_{f_{k-1}}(\succ))$. We show that for each ordering of agents, there is an exact mapping from f' to f. For all preference profiles, map each agent ordering as follows: $f_1 = f'_1$, $f_2(\pi_{f_1}(\succ)) = f'_2$, …, $f_k(\pi_{f_1}(\succ), \ldots, \pi_{f_{k-1}}(\succ) = f'_k$. This implies that f is a dictatorial ordering, which concludes our existence proof. □

Theorem 3.4. *An allocation mechanism is strategyproof, non-bossy, and Pareto C-efficient if and only if it is a sequential dictatorship quota mechanism.*

Proof. Sequential dictatorship quota mechanisms are strategyproof and non-bossy. Proposition 3.4 states that when preferences are lexicographic sequential dictatorships are Pareto C-efficient. Sequential dictatorships are also Pareto efficient when $C = \sum_{i=1}^{|q|} q_i$.

We must show the converse. Let π be a strategyproof, Pareto C-efficient, and non-bossy mechanism with quota q. By Lemma 3.5, given an identical preference profile and a quota q, there exists a sequential ordering f where agent f_1 receives q_1 of her favorite objects from M, agent $f_2(\pi_{f_1}(\succ))$ receives q_2 of her best objects from $M \setminus \pi_{f_1}$, and so on. Therefore, since the choice of the first dictator is independent of preference profile, we can identify a sequential ordering $f_1, f_2(\pi_{f_1}(\succ)), \ldots, f_k(\pi_{f_1}(\succ), \ldots, \pi_{f_{k-1}}(\succ))$ that receive objects according to $q = (q_1, \ldots, q_k)$.

Similar to the proof of Theorem 3.2, we construct an alternate preference profile $\succ'$, based on the given preference profile, at which all agents have identical preferences, where $\succ' = (\succ'_i, \ldots, \succ'_i)$.

According to function f, the first best q_1 objects according to $\succ_{f_1}$ are denoted by $\pi_{f_1}(\succ)$ and are listed as the first objects in $\succ'_i$. The next q_2 objects in $\succ'_2$ are the first best q_2 objects according to $\succ_{f_2(\pi_{f_1}(\succ))}$ from $M \setminus \pi_{f_1}(\succ)$, and so on. In general, for each $i \in N \setminus f_1$, the next best q_i objects are the best q_i objects according to $\succ_{f_i(\pi_{f_1}(\succ), \ldots, \pi_{f_{i-1}}(\succ))}$ from $M \setminus \bigcup_{j=1}^{j=i-1} \pi_{f_j}(\succ)$. These steps are depicted in Algorithm 2.

Algorithm 2: Constructing an identical preference profile

Data: A preference profile $\succ$, first dictator f_1, and quota q
Result: A profile with identical preferences $\succ'$ with $\pi(\succ') = \pi(\succ)$
Initialize $\succ_1 \leftarrow \emptyset$
Initialize set $Z = \emptyset$
for $(i \leftarrow 1$ **to** $|q|)$ **do**
    if $(i = 1)$ **then**
        $k \leftarrow f_1$ *// The first dictator is known.*
    else
        $k \leftarrow f_i(\pi_{f_1}(\succ), \ldots, \pi_{f_{i-1}}(\succ))$ *// Identify the next dictator*
    $Z \leftarrow \text{top}(q_i, \succ_k)$ *// Most preferred set of size q_i from the remaining objects.*
    $\succ'_1 \leftarrow \text{append}(\succ'_1, Z)$ *// Append this set to the preference ordering.*
    $Z \leftarrow \emptyset$
for $(i \leftarrow 1$ **to** $|f|)$ **do**
    $\succ'_i \leftarrow \succ'_1$
return $\succ'$.

By Lemma 3.2, for any agent in f the outcome of $\pi(\succ')$ must remain unchanged if the outcome of all predecessor agents remains unchanged. Thus, by Lemma 3.3, for any allocation $A \in \mathcal{A}$, if for each agent $i \in N$, $\pi_i(\succ') \succeq'_i A_i$ then we also have $\pi_i(\succ') \succeq_i A_i$. For each $f_i(\cdot)$ where $i = 1, \ldots, |f|$, by Lemma 3.3 since π is strategyproof and non-bossy, for any allocation A_{f_i} given the quota q we have

$$\pi_{f_i(\pi_{f_1}(\succ),\ldots,\pi_{f_{i-1}}(\succ))} \succeq'_{f_i(\pi_{f_1}(\succ),\ldots,\pi_{f_{i-1}}(\succ))} A_{f_i(\pi_{f_1}(\succ),\ldots,\pi_{f_{i-1}}(\succ))}$$
$$\pi_{f_i(\pi_{f_1}(\succ),\ldots,\pi_{f_{i-1}}(\succ))} \succeq_{f_i(\pi_{f_1}(\succ),\ldots,\pi_{f_{i-1}}(\succ))} A_{f_i(\pi_{f_1}(\succ),\ldots,\pi_{f_{i-1}}(\succ))}$$

which implies that $\pi(\succ') = \pi(\succ)$. Therefore, we identified an sequential ordering of agents that induces the same outcome as the original mechanism. Thus, π is a sequential dictatorship quota mechanism. □

3.6 Randomized Quota Mechanisms

So far we identified the class of deterministic strategyproof, non-bossy, and Pareto C-efficient quota mechanisms. However, deterministic quota mechanisms generally have poor fairness properties: the first dictator always has strong advantage over the next dictator and so on. This unfairness could escalate when an agent gets to pick more objects than the successor agent, that is, $q_i > q_j$ for $i < j$. Thus, while any profile-independent randomization over a set of serially dictatorial mechanisms still maintains the incentive property, randomization over priority orderings seem to be a proper way of restoring some measure of randomized fairness.

We first need to define a few additional properties in the randomized settings. A random allocation is a stochastic matrix A with $\sum_{i \in N} A_{i,j} = 1$ for each $j \in M$. This feasibility condition guarantees that the probability of assigning each object is a proper probability distribution. Moreover, every random allocation is a convex combination of deterministic allocations and is induced by a lottery over deterministic allocations [141]. Hence, we can focus on mechanisms that guarantee Pareto C-efficient solutions ex post.

Definition 3.7 (Ex Post C-Efficiency)**.** *A random allocation is ex post C-efficient if it can be represented as a probability distribution over deterministic Pareto C-efficient allocations.*

The support of any lottery representation of a strategyproof allocation mechanism must consist entirely of strategyproof deterministic mechanisms. Moreover, if the distribution over orderings does not depend on the submitted preferences of the agents, then such randomized mechanisms are strategyproof [118].

We focus our attention on the *downward* lexicographic dominance relation to compare the quality of two random allocations when preferences are lexicographic.[5] Given two allocations, an agent prefers the one in which there is a higher probability for getting the most-preferred object. Formally, given a preference ordering $\succ_i$, agent i prefers any allocation A_i that assigns a higher probability to her top ranked object A_{i,o_1} over any assignment B_i with $B_{i,o_1} < A_{i,o_1}$, regardless of the assigned probabilities to all other objects. Only when two assignments allocate the same probability to the top object will the agent considers the next-ranked object. Throughout this thesis we focus on the downward lexicographic relation, as opposed to upward lexicographic relation due to [46]. The downward lexicographic notion compares random allocations by comparing the probabilities assigned to objects in order of preference. Thus, it is a more natural way of comparing allocations and has shown to be often used in consumer markets and other settings involving human decision makers [84, 137, 144].

Definition 3.8. *Agent i with preference $\succ_i$ downward lexicographically prefers random allocation A_i to B_i if*

$$\exists \, \ell \in M : A_{i,\ell} > B_{i,\ell} \ \wedge \ \forall k \succ_i \ell : A_{i,k} = B_{i,k}.$$

[5] In the general domain, this measure corresponds to a stronger notion based on first-order stochastic dominance [30, 69]

We say that allocation A **downward lexicographically dominates** another allocation B if there exists no agent $i \in N$ that lexicographically prefers B_i to A_i. Thus, an allocation mechanism is downward lexicographically efficient (*ld-efficient*) if for all preference profiles its induced allocation is not downward lexicographically dominated by any other random allocation. We can see that efficiency under general preferences immediately implies ld-efficiency under lexicographic preferences. However, some allocations may only guarantee efficiency when preferences are lexicographic.

Example 3.4. Consider four agents $N = \{1, 2, 3, 4\}$ and four objects $M = \{a, b, c, d\}$ with quota $q = (1, 1, 1, 1)$ at the following preference profile $\succ = ((cabd), (acdb), (cbda), (acbd))$. Note that preferences are only defined over single objects, and we write $(cabd)$ as a shorthand form of $\succ_1 = c \succ a \succ b \succ d$. Table 3.3 shows the stochastic efficient allocation in

	a	b	c	d
A_1	0	1/3	1/2	1/6
A_2	1/2	0	0	1/2
A_3	0	1/3	1/2	1/6
A_4	1/2	1/3	0	1/6

(a) sd-efficient allocation

	a	b	c	d
A_1	1/12	1/3	5/12	1/6
A_2	11/24	0	1/12	11/24
A_3	0	5/12	5/12	1/6
A_4	11/24	1/4	1/12	5/24

(b) ld-dominated but not sd-dominated

Table 3.3: An example showing an allocation that is ld-efficient but not sd-efficient.

comparison with ld-efficient allocation. Here, even though the allocation in Table 3.3b is ld-dominated by the sd-efficient allocation, it is not stochastically dominated under the first-order stochastic dominance. This is because agent 2 (similarly agent 4) weakly prefers the allocation in Table 3.3b if only considering her first two top objects. Thus, the two random allocations are in fact incomparable with respect to stochastic dominance.

Given an allocation A, we say that agent i is envious of agent j's allocation if agent i prefers A_j to her own allocation A_i. Thus, an allocation is envyfree when no agent is envious of another agent's assignment. Formally we write,

Definition 3.9. *Allocation A is envyfree if for all agents $i \in N$, there exists no agent-object pair $j \in N$, $\ell \in M$ such that,*

$$A_{j,\ell} > A_{i,\ell} \;\wedge\; \forall k \succ_i \ell : A_{i,k} = A_{j,k}$$

A mechanism is envyfree if at all preference profiles $\succ \in \mathcal{P}^n$ it induces an envyfree allocation.

3.6.1 Random Serial Dictatorship Quota Mechanisms

Recall that $|q|$ denotes the number of agents that are assigned at least one object. Given a quota of size $|q|$, there are $\binom{n}{|q|} \times |q|!$ permutations (sequences without repetition) of $|q|$ agents from N. Thus, a *Random Serial Dictatorship mechanism with quota* q is a uniform randomization over all permutations of size $|q|$. Formally,

Definition 3.10 (Random Serial Dictatorship Quota Mechanism (RSDQ))**.** *Let $\mathbb{P}(N)$ be the power set of N, and $f \in \mathbb{P}(N)$ be any subset of N. Given a preference profile $\succ \in \mathcal{P}^n$, a random serial dictatorship with quota q is a convex combination of serial dictatorship quota mechanisms and is defined as*

$$\frac{\sum_{f \in \mathbb{P}(N):|f|=|q|} \pi_f(\succ)}{\binom{n}{|q|} \times |q|!} \tag{3.1}$$

In this randomized mechanism agents are allowed to pick more than one object according to q and not all the agents will be allocated ex post. We can think of such mechanisms as extending the well-known Random Serial Dictatorship (RSD) for the house assignment problem wherein each agent is entitled to receive exactly one object. Thus, an RSD mechanism is a special case of our quota mechanism with $q_i = 1, \forall i \in N$ and $|q| = n$.

Example 3.5. Consider three agents and four objects. Agents' preferences and the probabilistic allocation induced by RSDQ with quota $q = (2, 1, 1)$ are presented in Table 3.4. Note that the size of q can potentially be smaller than the number of agents, meaning that some agents may receive no objects ex post.

$\succ_1$	$c \succ a \succ b \succ d$
$\succ_2$	$a \succ c \succ d \succ b$
$\succ_3$	$c \succ b \succ d \succ a$

	a	b	c	d
A_1	3/6	1/6	2/6	2/6
A_2	3/6	0	2/6	3/6
A_3	0	5/6	2/6	1/6

Table 3.4: RSDQ allocation with $q = (2, 1, 1)$.

The weakest notion of fairness in randomized settings is the equal treatment of equals. We say an allocation is fair (in terms of equal treatment of equals) if it assigns an identical random allocation (lottery) to agents with identical preferences.

Theorem 3.5. *Take any serial dictatorship mechanism π with a quota q. A uniform randomization over all permutations of orderings with size $|q|$ is strategyproof, ex post C-efficient, and fair (equal treatment of equals).*

Proof. Showing ex post C-efficiency is simple: any serial dictatorship mechanism satisfies Pareto C-efficiency, and thus, any randomization also guarantees a Pareto C-efficient solution ex post. The support of the random allocation consists of only strategyproof deterministic allocations, implying that the randomization is also strategyproof. The equal treatment of equal is the direct consequence of the uniform randomization over the set of possible priority orderings. □

Now, we present our main result for envyfreeness of RSDQ regardless of the selected quota system.

Theorem 3.6. *Random Serial Dictatorship Quota mechanism is envyfree with any quota q, under downward lexicographic preferences.*

Proof. Let A denote a random allocation induced by RSDQ with quota q at an arbitrary preference profile $\succ \in \mathcal{P}^n$. Suppose for contradiction that there exists an agent $i \in N$ with random allocation A_i that prefers another agent's random allocation A_j to her own assignment, that is, $A_j \succ_i A_i$. Assuming that preferences are downward lexicographic, there exists an object ℓ such that $A_{j,\ell} > A_{i,\ell}$ and for all objects that are ranked higher than ℓ (if any) they both receive the same probability $\forall k \succ_i \ell : A_{i,k} = A_{j,k}$. Thus, we can write: $\sum_{x \in A_i : x \succ_i \ell} A_{j,x} = \sum_{x \in A_i : x \succ_i \ell} A_{i,x}$. Since preferences are lexicographic, the assignments of objects less preferred to ℓ become irrelevant because for two allocations A_i and B_i such that $A_{i,\ell} > B_{i,\ell}$, we have $A_i \succ_i B_i$ for all $x \prec_i \ell$ where $B_{i,x} \geq A_{i,x}$. Thus, we need only focus on object ℓ.

Let $\mathcal{F}$ denote the set of all orderings of agents where i is ordered before j or i appears but not j. Note that since we allow for $|q| = |f| \leq n$, some agents could be left unassigned, and permuting i and j could imply that one is not chosen under $\binom{n}{|q|}$. For any ordering $f \in \mathcal{F}$ of agents where i precedes j, let $\bar{f} \in \bar{\mathcal{F}}$ be the ordering obtained from f by swapping i and j. Clearly, $|\mathcal{F}| = |\bar{\mathcal{F}}|$ and the union of the two sets constitute the set of orderings that at least one of i or j (or both) is present. Fixing the preferences, we can only focus on f and $\bar{f}$.

Let $\pi_f(\succ)$ be the serial dictatorship with quota q and ordering f at $\succ$. RSDQ is a convex combination of such deterministic allocations with equal probability of choosing an ordering from any of $\mathcal{F}$ or $\bar{\mathcal{F}}$.

Given any object $y \in M$, either i receives y in π_f and j gets y in $\pi_{\bar{f}}$, or none of the two gets y in any of π_f and $\pi_{\bar{f}}$. Thus, object ℓ is either assigned to i in π_f and to j in $\pi_{\bar{f}}$, or is assigned to another agent. If i gets ℓ in π_f for all $f \in \mathcal{F}$, then j receives ℓ in $\pi_{\bar{f}}$. The contradiction assumption $A_{j,\ell} > A_{i,\ell}$ implies that there exists an ordering f where i

receives a set of size q_i that does not include object ℓ while j's allocation set includes ℓ. Let X_i denote this set for agent i and X_j for agent j. Then, $X_i \succ_i X_j$. Thus, by definition there exists an object $\ell' \in X_i$ such that $\ell' \succ_i \ell$, where $\ell' \notin X_j$. Thus, the probability of assigning object $\ell' \succ_i \ell$ to i is strictly greater than assigning it to j, that is, $A_{i,\ell'} > A_{j,\ell'}$. However, by lexicographic assumption we must have $\forall k \succ_i \ell : A_{i,k} = A_{j,k}$, which is a contradiction. □

Theorem 3.7. *Under downward lexicographic preferences, a Random Serial Dictatorship Quota mechanism is ex post C-efficient, strategyproof, and envyfree for any number of agents and objects and any quota system.*

The well-known random serial dictatorship mechanism (RSD), also known as Random Priority, is defined when $n = m$ and assigns a single object to agents [6]. It is apparent that RSD is a special instance from the class of RSDQ mechanisms.

Corollary 3.1. *RSD is ex post efficient, strategyproof, and envyfree when preferences are downward lexicographic.*

Proof. The conventional RSD mechanism is equivalent to an RSDQ mechanism where agents receive exactly one object, that is, $\sum_i q_i = m$ and for each agent i, $q_i = 1$. Therefore, RSD satisfies ex post efficiency, strategyproofness, and envyfreeness. □

3.7 Discussion

We investigated the strategyproof allocation mechanisms when agents with lexicographic preferences may receive more than one object according to a quota. The class of sequential quota mechanisms enables the social planner to choose any quota without any limitations. For the general domain of preferences, however, the class of strategyproof, non-bossy, and Pareto efficient mechanisms is restricted to sequential dictatorships with equal quota sizes. Demanding neutrality, the set of such mechanisms gets restricted to quasi-dictatorial mechanisms, which are far more unfair [109, 110]. Thus, such mechanisms limit a social planner to specific quota systems while demanding the complete allocation of all available objects.

The class of strategyproof allocation mechanisms that satisfy neutrality, Pareto C-efficiency, and non-bossiness expands significantly when preferences are lexicographic. Our characterization shows that serial dictatorship quota mechanisms are the only mechanisms satisfying these properties in the multiple-assignment problem. Removing the neutrality

	a	b	c
A_1	1/2	0	1/2
A_2	1/2	1/6	1/3
A_3	0	5/6	1/6

Table 3.5: An allocation prescribed by RSD that is envyfree under downward lexicographic preferences but is not envyfree under the upward lexicographic notion.

requirement, this class of mechanisms further expands to sequential dictatorship quota mechanisms.

To recover some level of fairness, we extended the serial dictatorship quota mechanisms to randomized settings and showed that randomization can help achieve some level of stochastic symmetry amongst the agents. More importantly, we showed that RSDQ mechanisms satisfy strategyproofness, ex post C-efficiency, and envyfreeness for any number of agents, objects, and quota systems when preferences are downward lexicographic. The envyfreeness result is noteworthy: it shows that in contrast to the Probabilistic Serial rule (PS) [30] which satisfies strategyproofness when preferences are lexicographic only when $n \geq m$ [124], the well-known RSD mechanism in the standard assignment problem is envyfree for any combination of n and m. These results address the two open questions about the existence of a mechanism with more favorable fairness and strategyproofness properties [109, 124].

Note that our envyfreeness result for RSDQ mechanisms holds under the assumption of downward lexicographic notion, which is a more natural way of comparing probabilistic outcomes. The upward lexicographic extension (due to Cho [46]) is a dual to the downward lexicographic notion, which is based on lexicographically minimizing the probabilities of less preferred objects. Our envyfreeness results in Theorem 3.6 and Theorem 3.7 do not hold under the upward extension. For instance, consider three objects and three agents with the following preferences: $\succ = ((a \succ c \succ b), (a \succ b \succ c), (b \succ a \succ c))$. A simple RSD mechanism will assign probabilities as shown in Table 3.5. This allocation is envyfree under downward lexicographic preferences; however, under the upward lexicographic relation agent 2 is envious of agent 3's allocation because agent 3 is assigned less probability for object c, its least preferred object.

Serial dictatorship mechanisms are widely used in practice since they are easy to implement while providing stability and strategyproofness guarantees [116]. Serial dictatorship quota mechanisms and their randomized counterparts provide a richer framework for multiple allocation problems while creating the possibility of fair and envyfree assignments. Our characterization for deterministic quota mechanisms justifies the use of quotas in sequential settings. In randomized settings, however, an open question is whether RSDQ mechanisms are the only allocation rules that satisfy the above properties in the multiple assignment domain. Of course, answering this question, first, requires addressing the open question by Bade [22] in the standard assignment problem (where every agent gets at most one object): *is random serial dictatorship a unique mechanism that satisfies strategyproofness, ex post efficiency, and equal treatment of equals?*

3.8 Design Recommendations

Characterizing the set of desirable allocation mechanisms is key in choosing which mechanisms to adopt in practice and in designing effective mechanisms in multiagent systems. Consider an online task allocation system for assigning various tasks to workers through crowdsourcing. Workers may have different preferences over the set of tasks and a mechanism designer may require that a subset of workers get assigned to multiple tasks while the rest only get to complete one task each. In this case, the choice of serial dictatorship quota mechanisms and its randomized counterpart (RSDQ) would be appropriate to ensure Pareto efficiency of the final allocation while preventing agent manipulation by guaranteeing strategyproofness and non-bossiness.

Fairness is an essential property in many allocation and scheduling settings. Under downward lexicographic preferences, our proposed RSDQ mechanism satisfies a strong notion of envyfreeness. Consider an online system for scheduling shifts to nurses or caregivers based on their preferences. Not only does the RSDQ mechanism prevent manipulation, but also it satisfies envyfreeness for those with lexicographic preferences. Moreover, a serial dictatorship quota mechanism enables the social planner to set any desired quota based on seniority so that senior nurses get to have higher priorities but are required to pick more than one shift while other junior nurses are ordered last and get to pick shifts later.

Chapter 4

Empirically Investigating the Characteristics of One-Sided Matching Mechanisms

One-sided matching mechanisms have been extensively adopted in many resource allocation settings such as assigning dormitory rooms or offices to students, students to public schools, college courses to students, organs and medical resources to patients, and members to subcommittees [13, 38, 100, 117]. Two important (randomized) matching mechanisms that only elicit ordinal preferences from agents are *Random Serial Dictatorship* (RSD) [6] and *Probabilistic Serial Rule* (PS) [30]. Both mechanisms have important economic properties and are practical to implement. The RSD mechanism has strong truthful incentives but guarantees neither efficiency nor envyfreeness. PS satisfies efficiency and envyfreeness; however, it is susceptible to manipulation. Therefore, there are subtle points to be considered when deciding which mechanism to use. For example, given a particular preference profile, the mechanisms often produce random assignments which are simply incomparable and thus, without additional knowledge of the underlying utility models of the agents, it is difficult to determine which is the "better" outcome. Furthermore, properties like efficiency, truthfulness, and envyfreeness can depend on whether there is underlying structure in the preferences, and even in general preference models it is valuable to understand under what conditions a mechanism is likely to be efficient, truthful, or envyfree as this can guide designers choices.

In this chapter, we study the comparability of PS and RSD when there is only one copy of each object, and analyze the space of all preference profiles for different numbers of agents and objects. Working in the space of general ordinal preferences, we provide empirical results on the (in)comparability of RSD and PS and analyze their respective economic properties.

We show that despite the inefficiency of RSD, the fraction of random assignments at which PS is strictly outperforms RSD vanishes, especially when the number of agents is less than or equal to the available objects. We also investigate the manipulability of PS and show that PS is almost 99% manipulable for any combination of agents and objects, and the fraction of strongly manipulable profiles goes to 1 as the ratio of objects to agents increases. We show similar trends on these properties under lexicographic preferences, and further present results on envy of RSD. Our results show that although the fraction of envious agents grows with the number of agents, there is a sudden drop in the fraction of envious agents when there are equal numbers of agents and objects.

In Section 4.4, we instantiate utility functions for agents to gain deeper insights on the manipulability, social welfare, and envyfreeness of the two mechanisms under different risk attitudes. Our main result is that under risk aversion, the social welfare of RSD is as good as PS but RSD does create envy among the agents (though the fraction of envious profiles and the total envy are small). Moreover, when the number of agents and objects are equal, RSD assignments are less likely to be dominated by PS and overall RSD assignments create negligible envy among agents. We also show that PS is highly susceptible to manipulation in almost all combinations of agents and objects. The fraction of manipulable profiles and the gain from manipulation rapidly increases, particularly when agents become more risk averse.

4.1 Model

In this section we describe the basic one-sided matching problem and introduce the two mechanisms we study in detail, Random Serial Dictatorship (RSD) [6] and Probabilistic Serial Rule (PS) [30]. We then introduce a number of properties and criteria used to evaluate these mechanisms.

A one-sided matching problem consists of a set of n agents, N, and a set of m indivisible objects, M.[1] Each agent $i \in N$ has a private strict preference ordering, $\succ_i$, over M where $a \succ_i b$ indicates that agent i prefers to receive object a over object b. We represent the preference ordering of agent i by the ordered list of objects $\succ_i = a \succ_i b \succ_i c$ or $\succ_i = (abc)$, for short. We let $\mathcal{P}$ denote the set of all complete and strict preference orderings over M.

[1] This problem is sometimes called the assignment problem or house allocation problem in the literature.

A *preference profile* $\succ \in \mathcal{P}^n$ specifies a preference ordering for each agent, and we use the standard notation $\succ_{-i} = (\succ_1, \ldots, \succ_{i-1}, \succ_{i+1}, \ldots, \succ_n)$ to denote preferences orderings of all agents except i and thus $\succ = (\succ_i, \succ_{-i})$.

The goal in a one-sided matching problem is to assign the objects in M to the agents in N according to preference profiles, under the constraint that no object can be assigned to more than one agent. If $m = n$ then this means that each agent will receive exactly one object, however if $m < n$ then some agents will receive no object and if $m > n$ then some agents may receive multiple objects. An assignment is represented as a matrix

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,m} \\ A_{2,1} & A_{2,2} & \ldots & A_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \ldots & A_{n,m} \end{pmatrix}$$

where $A_{i,j} \in [0,1]$ is the probability that agent i is assigned object j. We let $\mathcal{A}$ denote the set of all *feasible* assignments where an assignment $A \in \mathcal{A}$ is *feasible* if and only if $\forall j \in M$, $\sum_{i \in N} A_{i,j} = 1$. If $A \in \mathcal{A}$ is such that $A_{i,j} \in \{0,1\}$ then we say that A is a *deterministic* assignment; otherwise, A is a *random* assignment. Every random assignment can be represented as a convex combination of deterministic assignments [141], and thus we view random assignments as a probability distribution over a set of deterministic assignments.

4.1.1 Matching Mechanisms

In general, a *matching mechanism*, $\mathcal{M}$, is a mapping from the set of preference profiles, $\mathcal{P}^n$ to the set of feasible assignments, $\mathcal{A}$. That is, $\mathcal{M} : \mathcal{P}^n \mapsto \mathcal{A}$. We focus our attention on two widely studied mechanisms for one-side matching: Random Serial Dictatorship (RSD) [6] and Probabilistic Serial Rule (PS) [29].

RSD relies on the concept of priority orderings over agents. Such an ordering is an ordered list of agents where the first agent gets to select its most preferred object from the set of objects, the second agent then selects its most preferred object from the set of remaining objects and so on until no objects remain.[2] Given a preference profile $\succ \in \mathcal{P}^n$, RSD returns an assignment $RSD(\succ) \in \mathcal{A}$ which is a uniform distribution over all

[2]When $n < m$ and agents can receive more than one object, RSD requires a careful method for picking a sequence at each priority ordering to ensure strategyproofness. As discussed in Chapter 3, this picking sequence should be based on an arbitrary serial dictatorship quota mechanism, which directly affects the

deterministic assignments induced from all possible priority orderings over the set of agents. RSD has been widely adopted for fair and strategyproof assignments for the school choice problem, course assignment, house allocation, and room assignment [3, 6, 7, 131]

PS treats objects as a set of divisible goods of equal size and simulates a simultaneous eating algorithm. Each agent starts "eating" its most preferred object, all at the same rate. Once an object is gone (eaten away) then the agent starts eating its next preferred object among the remaining objects. This process terminates when all objects have been "eaten". Given a preference profile $\succ \in \mathcal{P}^n$, $PS(\succ) \in \mathcal{A}$ is a random assignment where $A_{i,j}$ is the probability (fraction) that object j is assigned to (or "eaten by") agent i.

4.1.2 General Properties

In this section we define key properties for matching mechanisms. To evaluate the quality of a random assignment, we use first-order stochastic dominance [30, 69]. Given a random assignment A_i, the probability that agent i is assigned an object that is at least as good as object ℓ is defined as follows

$$w(\succ_i, \ell, A_i) = \sum_{j \in M: j \succeq_i \ell} A_{i,j} \tag{4.1}$$

We say an agent always prefers assignment A_i to B_i, if for each object ℓ the probability of assigning an object at least as good as ℓ under A_i is greater or equal that of B_i, and strictly greater for some object.

Definition 4.1 (Stochastic Dominance). *Given a preference ordering $\succ_i$, random assignment A_i stochastically dominates (sd) assignment $B_i (\neq A_i)$ if*

$$\forall \ell \in M, \; w(\succ_i, \ell, A_i) \geq w(\succ_i, \ell, B_i) \tag{4.2}$$

A matching mechanism is *sd*-efficient if at all preference profiles $\succ \in \mathcal{P}^n$, for all agents $i \in N$, the prescribed assignment is not stochastically dominated by any other assignment.

Definition 4.2 (*sd*-Efficiency). *A random assignment is sd-efficient if for all agents, it is not stochastically dominated by any other random assignment.*

efficiency and envy of the assignments [33, 75]. For simplicity, we use the variant of RSD based on a *quasi-dictatorial* mechanism [109] where the first agent selects its most preferred $(m - n + 1)$ objects, and the rest of the agents choose one object each.

An important desirable property in matching mechanisms is strategyproofness, that is the mechanism is designed so that no agent has incentive to misreport its preferences.

Definition 4.3 (*sd*-Strategyproofness)**.** *Mechanism $\mathcal{M}$ is sd-strategyproof if at all preference profiles $\succ \in \mathcal{P}^n$, for all agents $i \in N$, and for any misreport $\succ'_i \in \mathcal{P}^n$, such that $A = \mathcal{M}(\succ)$ and $A' = \mathcal{M}(\succ'_i, \succ_{-i})$, we have:*

$$\forall \ell \in M,\ w(\succ_i, \ell, A_i) \geq w(\succ_i, \ell, A'_i) \tag{4.3}$$

Sd-strategyproofness is a strict requirement. It implies that under any utility model consistent with the preference orderings, no agent can improve her expected utility by misreporting. We say that a mechanism is *weakly sd-strategyproof* if at all preference profiles there is no misreport such that for all $\ell \in M$, $w(\succ_i, \ell, A'_i) \geq w(\succ_i, \ell, A_i)$ with at least one $\ell' \in M$ such that $w(\succ_i, \ell', A'_i) > w(\succ_i, \ell', A_i)$. Clearly, *sd*-strategyproofness implies weak *sd*-strategyproofness but the converse does not hold.

An assignment is *manipulable* if it is not *sd*-strategyproof. If there exists some agent who strictly benefits from the manipulation, (*i.e.* the mechanism is not even weakly *sd*-strategyproof) then we say the assignment is *sd*-manipulable (or strictly manipulable).

We are also interested in whether mechanisms are fair and use the notion of envyfreeness to this end. An assignment is *sd*-envyfree if each agent strictly prefers her random allocation to any other agent's assignment.

Definition 4.4 (*sd*-Envyfreeness)**.** *Given agent i's preference $\succ_i$, assignment A_i is sd-envyfree if for all agents $\forall k \neq i \in N$,*

$$\forall \ell \in M,\ w(\succ_i, \ell, A_i) \geq w(\succ_i, \ell, A_k) \tag{4.4}$$

We say an assignment is weakly *sd*-envyfree if the inequality in Equation 4.4 is strict for some $\ell \in M$, but there exists at least one ℓ' for which the inequality in Equation 4.4 does not hold. A matching mechanism satisfies *sd*-envyfreeness if at all preference profiles $\succ \in \mathcal{P}^n$, it induces *sd*-envyfree assignments for all agents.

Lastly, we are interested in investigating efficiency, manipulation, and envy of the random mechanisms when preferences are lexicographic. Under lexicographic preferences, given two allocations, an agent prefers the one in which there is a higher probability for getting the most-preferred object. Recall our definition of downward lexicographic dominance in Chapter 3.

	$n \geq m$		$n < m$	
	PS	RSD	PS	RSD
sd-strategyproof	weak	✓	✗	✓
sd-efficiency	✓	✗	✓	✗
sd-envyfree	✓	weak	✓	weak

Table 4.1: Properties of PS and RSD.

Definition 4.5 (Lexicographic Dominance). *Given a preference ordering $\succ_i$, random assignment A_i lexicographically dominates (ld) assignment B_i if*

$$\begin{aligned} &\exists\ \ell \in M : w(\succ_i, \ell, A_i) > w(\succ_i, \ell, B_i)\ \wedge \\ &\forall\ k \succ_i \ell : w(\succ_i, \ell, A_i) = w(\succ_i, \ell, B_i). \end{aligned} \tag{4.5}$$

We say that allocation A *downward lexicographically dominates* another allocation B if there exists no agent $i \in N$ that lexicographically prefers B_i to A_i. Thus, an allocation mechanism is lexicographically efficient (*ld*-efficient) if for all preference profiles its induced allocation is not lexicographically dominated by any other random allocation.

4.1.3 Properties of RSD and PS

The theoretical properties of PS and RSD have been well studied in the economics literature [30], and we summarize the results in Table 4.1. Both mechanisms are ex post efficient, that is, their *realized outcomes* cannot be improved without making at least one agent worse off. PS has been shown to be both *sd*-envyfree and *sd*-efficient. However, it is not even weakly *sd*-strategyproof when $n < m$ [93] and is only weakly *sd*-strategyproof when $n \geq m$. On the other hand, RSD is always *sd*-strategyproof, but it is only weakly *sd*-envyfree and is not *sd-efficient*. Example 4.1 illustrates the *sd*-inefficiency of RSD.

Example 4.1. Suppose there are four agents $N = \{1, 2, 3, 4\}$ and four objects $M = \{a, b, c, d\}$. Consider the following preference profile $\succ= ((abcd), (abcd), (badc), (badc))$. Table 4.2 shows the outcomes for $PS(\succ)$ and $RSD(\succ)$. In this example, all agents strictly prefer the assignment induced by PS over the RSD assignment. Thus, RSD is inefficient at this preference profile.

	a	b	c	d
A_1	1/2	0	1/2	0
A_2	1/2	0	1/2	0
A_3	0	1/2	0	1/2
A_4	0	1/2	0	1/2

(a) Assignment under $PS(\succ)$

	a	b	c	d
A_1	5/12	1/12	5/12	1/12
A_2	5/12	1/12	5/12	1/12
A_3	1/12	5/12	1/12	5/12
A_4	1/12	5/12	1/12	5/12

(b) Assignment under $RSD(\succ)$

Table 4.2: Example showing the inefficiency of RSD

	a	b	c
A_1	1/2	0	1/2
A_2	1/2	1/4	1/4
A_3	0	3/4	1/4

(a) Assignment under $PS(\succ)$

	a	b	c
A_1	1/2	0	1/2
A_2	1/2	1/6	1/3
A_3	0	5/6	1/6

(b) Assignment under $RSD(\succ)$

Table 4.3: Incomparability of RSD and PS

4.2 Incomparability of RSD and PS

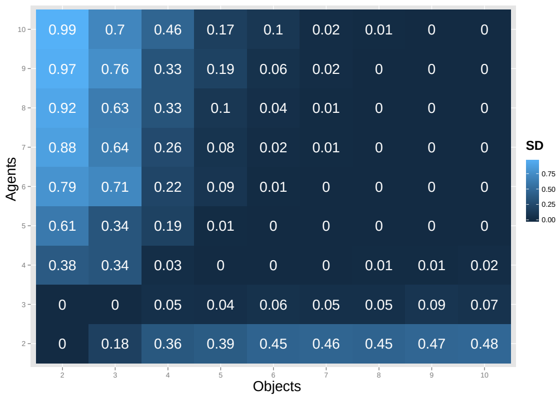
We argue that the theoretical findings on RSD and PS do not necessarily provide enough guidance to a market designer trying to select the correct mechanism for a specific setting. For example, while we know that PS is *sd*-efficient and RSD is not, this does not mean that PS assignments always stochastically dominate the assignments prescribed by RSD.

Example 4.2. Suppose there are three agents $N = \{1, 2, 3\}$ and three objects $M = \{a, b, c\}$. Consider the following preference profile $\succ = ((acb), (abc), (bac))$. Table 4.3 shows $PS(\succ)$ and $RSD(\succ)$. Neither assignment dominates the other since agent 1 is indifferent between the two assignments while agent 2 prefers $PS(\succ)$ and agent 3 prefers $RSD(\succ)$.

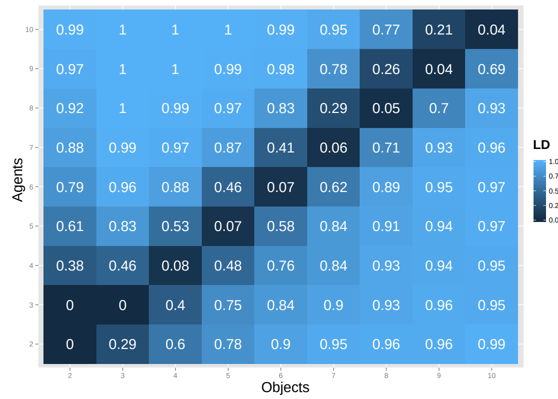
If we knew the utility functions of the agents, consistent with their ordinal preferences, then we might be able to use the notion of (utilitarian) social welfare to help determine the better assignment.[3] However, it is easy to construct different utility functions for the agents in Example 4.2 where both RSD and PS maximize social welfare.

Similarly, the envy of RSD and the manipulability of PS both depend on the structure of preference profiles, and thus, a compelling question, that justifies studying the practical

[3] Given utility functions for the agents, where $u_i(j)$ is the utility agent i derives from being assigned object j, the (utilitarian) social welfare of an assignment A is $\sum_i \sum_j A_{i,j} u_i(j)$.



(a) The fraction that PS stochastically dominates RSD.



(b) The fraction that PS lexicographically dominates RSD.

Figure 4.1: The fraction of preference profiles under which PS dominates RSD.

implications of deploying a matching mechanism, is to analyze the space of preference profiles to find the likelihood of inefficient, manipulable, or envious assignments under these mechanisms.

4.3 General and Lexicographic Preferences

The theoretical properties of PS and RSD only provide limited insight into their practical applications. In particular, when deciding which mechanism to use in different settings, the incomparability of PS and RSD leaves us with an ambiguous choice in terms of efficiency, manipulability, and envyfreeness. Thus, we examine the properties of RSD and PS in the space of all possible preference profiles as well as under lexicographic preferences. Lexicographic preferences are present in various applications and have been extensively studied in artificial intelligence and multiagent systems as a means of assessing allocations based on ordinal preferences [51, 58, 119]. Under lexicographic preferences, an allocation that assigns a higher probability to the top ranked object is always preferred to any other allocation, regardless of the probabilities assigned to objects in the next positions. Only when two allocations assign equal probabilities to the top ranked object, the probability of the next preferred object is considered. In the rest of this chapter, we denote the efficiency, strategyproofness, manipulability, and envyfreeness with *ld*- (lexicographically dominate) prefix.

The number of all possible preference profiles is super exponential $(m!)^n$. For each combination of n agents and m objects we performed a brute force coverage of all possible preference profiles. Thus, for all subsequent figures each data point shows the fraction of all possible preference profiles. For the cases of $n = 10$ and $m \in \{9, 10\}$, we randomly generated 1,000 instances by sampling from a uniform preference profile distribution. For each preference profile, we ran both PS and RSD mechanisms and compared their outcomes in terms of the stochastic dominance relation. Note that not only is computing RSD probabilities #P-complete (and thus intractable) [16, 121], but checking the desired properties such as envyfreeness, efficiency, and manipulablity of random allocations is shown to be NP-hard for general settings [20, 21]. Thus, for larger settings even if we randomly sample preference profiles it is not easy to verify the aforementioned properties.

4.3.1 Preliminary Results

Our experimentation discloses several intriguing observations, confirming theoretical results and providing additional insights into matching markets. A preliminary look at our empirical results illustrates the following: when $m \leq 2, n \leq 3$, PS coincides exactly with RSD, which results in the best of the two mechanisms, *i.e.*, both mechanisms are *sd*-efficient, *sd*-strategyproof, and *sd*-envyfree. Another interesting observation is that when $m = 2$, for all n, PS is *sd*-strategyproof (although the PS assignments are not necessarily equivalent to assignments induced by RSD), RSD is *sd*-envyfree, and for most instances when $m = 2$, PS stochastically dominates RSD, particularly when $n \geq 4$.

4.3.2 Efficiency

Our first finding is that the fraction of preference profiles at which RSD and PS prescribe identical random assignments goes to 0 when n grows. There are two conclusions that one can draw. First, this result confirms the theoretical results of Manea on asymptotic inefficiency of RSD [99], in that, in most instances, the assignments induced by RSD are not identical to the PS assignments. Second, this result suggests that the incomparability of outcomes is significant, that is, the social welfare of the random outcomes is highly dependent on the underlying utility models.

The fraction of preference profiles $\succ \in \mathcal{P}^n$ for which RSD is stochastically dominated by PS at $\succ$ converges to zero as $\frac{n}{m} \to 1$. Figure 4.1a shows that when m grows beyond $m > 5$, due to incomparability of RSD and PS with regard to the stochastic dominance relation, the

RSD assignments are rarely stochastically dominated by sd-efficient assignments prescribed by PS.

We also see similar results when we restrict ourselves to lexicographic preferences (Figure 4.1b). The fraction of preference profiles $\succ \in \mathcal{P}^n$ for which RSD is lexicographically dominated by PS at $\succ$ converges to zero as $\frac{n}{m} \to 1$.

For lexicographic preferences, we also observe that the fraction of preference profiles for which PS assignments strictly dominate RSD-induced allocations goes to 1 when the number of agents and objects diverge. The fraction of preference profiles $\succ \in \mathcal{P}^n$ for which RSD is lexicographically dominated by PS at $\succ$ converges to 1 as $|n - m|$ grows. Intuitively, when some agents can receive more than one object ($n < m$) or when there are not sufficient objects ($n > m$) for all agents, in each realized ordering of agents by RSD, those with higher priority are treated very differently than those in lower priority. Thus, the RSD outcome tend to be unfair and undesirable for most agents.

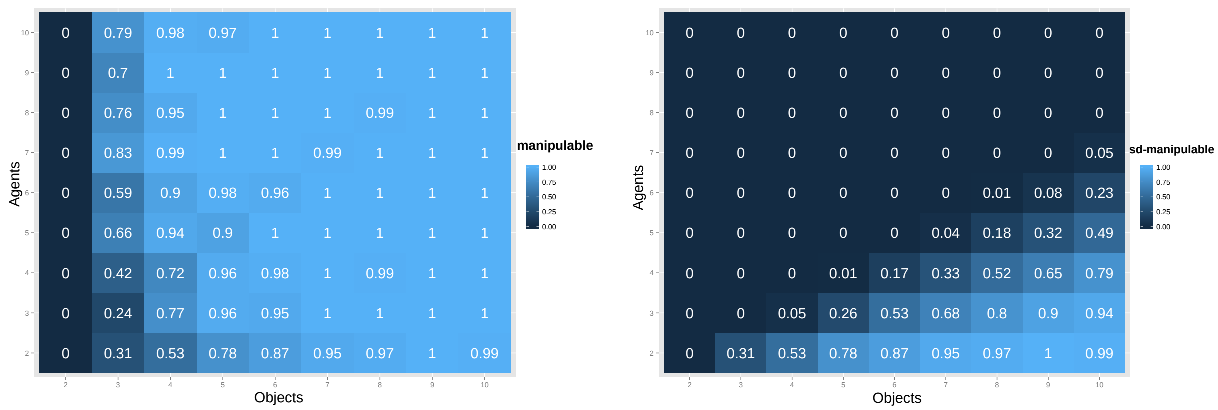
One immediate conclusion is that although RSD does not guarantee either sd-efficiency or ld-efficiency, in most settings when $\frac{n}{m} \to 1$ (and also $n \leq m$ for sd-efficiency according to Figure 4.1a), neither of the two mechanisms is preferred in terms of efficiency. Hence, one cannot simply rule out the RSD mechanism.

4.3.3 Manipulability of PS

One critical issue with deploying PS is that it does not provide incentives for honest reporting of preferences. Although for $n \geq m$ PS is weakly sd-strategyproof [30] and ld-strategyproof [124], when $n < m$ PS no longer satisfies these two properties.[4] The real concern is that, in the absence of strategyproofness, PS allocations are only efficient (or envyfree) with respect to the reported preferences. Thus, if an agent decides to manipulate the outcome by misreporting its preferences, PS will no longer guarantee efficiency, nor envyfreeness with respect to the true underlying preferences. Thus, we are interested in understanding the degree to which PS allocations are manipulable.

Figure 4.2 shows that the fraction of manipulable profiles goes to 1 as n or m grow. PS is almost 99% manipulable for $n > 5, m > 5$. Another interesting observation is that, for all $n < m$, the fraction of sd-manipulable preference profiles goes to 1 as $m - n$ grows (Figure 4.2b). These results imply that when agents are permitted to receive more than a single object, agents can strictly benefit from misreporting their preferences.

[4] A recent experimental study on the incentive properties of PS shows that human subjects are less likely to manipulate the mechanism when misreporting is a Nash equilibrium. However, subjects' tendency for misreporting is still significant even when it does not improve their allocations [81].



(a) The fraction of manipulable preference profiles under PS.

(b) The fraction of sd-manipulable profiles under PS.

Figure 4.2: Heatmaps illustrating the manipulability of PS.

Moreover, at those instances of problem where PS is *sd*-strategyproof, the assignment prescribed by PS most often coincides with the RSD induced assignment. For example, when $n = m = 5$, PS is only *sd*-strategyproof at 11% of preference profiles, 7% of which are identical to the assignments induced by RSD. This insight further confirms the vulnerability of PS to misreporting (See Table A.1 in Appendix A for detailed numerical results).

As illustrated in Figure 4.3, the manipulability of PS under lexicographic preferences has a similar trend when there are more objects than agents ($n < m$) and the fraction of *ld*-manipulable preference profiles converges to 1 even more rapidly when $\frac{m}{n}$ grows.

4.3.4 Envy in RSD

The PS mechanism has a desirable fairness property and is guaranteed to satisfy *sd*-envyfreeness, whereas RSD is not *sd*-envyfree. To further investigate the envy among agents under RSD, we measured the fraction of agents that are weakly *sd*-envious of at least one other agent.

Figure 4.4 shows that for RSD, the percentage of agents that are weakly envious increases with the number of agents. Figure 4.4a reveals an interesting observation: fixing any $n > 3$, the percentage of agents that are (weakly) envious grows with the number of objects, however, there is a sudden drop in the percentage of envious agents when there are equal number of agents and objects.

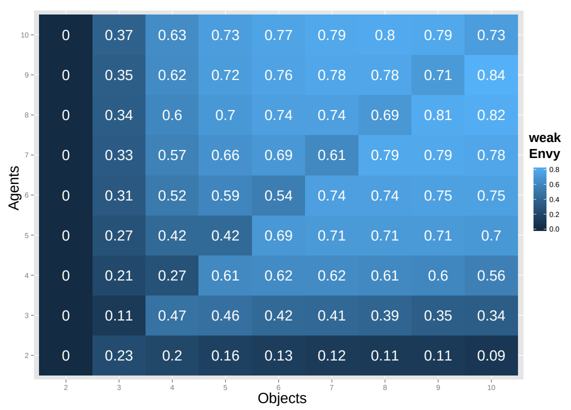
Figure 4.3: The fraction of ld-manipulable profiles under PS.

For better understanding of the population of agents who feel (weakly) envious under RSD, we illustrate the various envy profiles based on the percentage of envious agents in all instances of the problem when $n = m$ (Figure 4.4b). One observation is that there are few distinct envy profiles at each n, each representing a particular class of preference profiles, and by increasing n, the fraction of agents that are envious of at least one other agent increases.

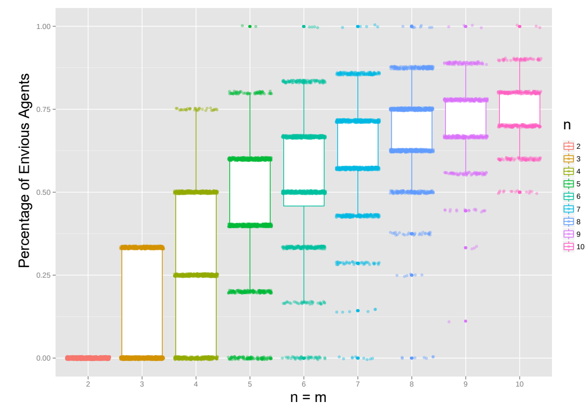
4.4 Utility Models

Given a utility model consistent with an agent's preference ordering, we can find the agent's expected utility for a random assignment. Let u_i denote agent i's Von Neumann-Morgenstern (VNM) utility model consistent with its preference ordering $\succ_i$. That is, $u_i(a) > u_i(b)$ if and only if $a \succ_i b$. Then, agent i's expected utility for random assignment A_i is $\mathbb{E}(u_i|A_i) = \sum_{j \in M} A_{i,j} u_i(j)$.

We say that agent i (strictly) prefers assignment A_i to B_i if and only if $\mathbb{E}(u_i|A_i) > \mathbb{E}(u_i|B_i)$. A mechanism is strategyproof if there exists no agent that can improve its expected utility by misreporting its preference ordering.



(a) A heatmap showing the percentage of envious agents.



(b) Boxplots showing the various envy profiles for $n = m$. The Y axis represents the percentage of envious agents.

Figure 4.4: Plots representing the percentage of (weakly) envious agents under RSD.

Definition 4.6 (Strategyproofness)**.** *Mechanism $\mathcal{M}$ is strategyproof if for all agents $i \in N$, and for any misreport $\succ_i' \in \mathcal{P}^n$, such that $A = \mathcal{M}(\succ)$ and $A' = \mathcal{M}(\succ_i', \succ_{-i})$, given a utility model u_i consistent with $\succ_i$, we have $\mathbb{E}(u_i|A_i) \geq \mathbb{E}(u_i|A_i')$.*

A matching mechanism is envyfree if for all preference profiles it prescribes an envyfree assignment.

Definition 4.7 (Envyfreeness)**.** *Assignment A is envyfree if for all $i, k \in N$, given utility model u_i consistent with $\succ_i$, we have $\mathbb{E}(u_i|A_i) \geq \mathbb{E}(u_i|A_k)$.*

Given utility functions for the agents, the (utilitarian) *social welfare* of an assignment A is $\sum_i \mathbb{E}(u_i|A_i)$. A random assignment A is *sd*-efficient if and only if there exists a profile of utility values consistent with $\succ$ such that A maximizes the social welfare ex ante [30, 101]. This existence result does not shed light on the social welfare when comparing two random assignments, since an assignment can be *sd*-efficient but may not have desirable ex ante social welfare. Consider the following random assignments: assignment A which is *sd*-efficient and assignment $B \neq A$ which is not stochastically dominated by A. Given a preference profile, A is guaranteed to maximize the social welfare for at least one profile of consistent utilities. However, there may be other profiles of utilities consistent with preferences at which B maximizes the sum of utilities (social welfare).

Example 4.3. Consider the problem introduced in Example 4.2 with assignments illustrated in Table 4.3. Let's assume that all agents have the same utility model $u_1 = u_2 = u_3$ where the utilities are $10, 9, 0$ for the first, second, and third objects respectively. The sum of expected utilities under the PS assignment is $(\frac{1}{2} \cdot 10 + \frac{1}{2} \cdot 9 + 0) + (\frac{1}{2} \cdot 10 + \frac{1}{4} \cdot 9 + \frac{1}{4} \cdot 0) + (\frac{3}{4} \cdot 10 + 0 \cdot 9 + \frac{1}{4} \cdot 0)$, while the sum of expected utilities under the RSD allocation is $(\frac{1}{2} \cdot 10 + \frac{1}{2} \cdot 9 + 0) + (\frac{1}{2} \cdot 10 + \frac{1}{6} \cdot 9 + \frac{1}{3} \cdot 0) + (\frac{5}{6} \cdot 10 + 0 \cdot 9 + \frac{1}{6} \cdot 0)$. It is easy to see that for this profile, the ex ante social welfare under RSD is larger than that of PS.

Thus, given a profile of utilities we investigate the (ex ante) social welfare of the assignments under PS and RSD.

4.4.1 Instantiating Utility Functions

To deepen our understanding as to the performance of the two mechanisms, we investigate different utility models. In particular we look at the performance of the mechanisms when the agents are all risk neutral (*i.e.* have linear utility functions), when agents are risk seeking and when agents are risk averse.

Our first utility model is the well-studied linear utility model. Given an agent i's preference ordering $\succ_i$, we let $r(\succ_i, j)$ denote the rank of object j. For example, given preference ordering $a \succ_i b \succ_i c$ then $r(\succ_i, a) = 1$, $r(\succ_i, b) = 2$ and $r(\succ_i, c) = 3$. The utility function for agent i, given object j is $u_i(j) = m - r(\succ_i, j)$.

We use an *exponential* utility model to capture risk attitudes beyond risk-neutrality. An exponential utility has been shown to provide an appropriate translation for individuals' utility models [11]. In particular, we define the exponential utility as follows:

$$u_i(j) = \begin{cases} (1 - e^{-\alpha(m - r(\succ_i, j))})/\alpha, & \alpha \neq 0 \\ m - r(\succ_i, j), & \alpha = 0 \end{cases} \tag{4.6}$$

The parameter α represents the agent's risk attitude. If $\alpha > 0$ then the agent is risk averse, while if $\alpha < 0$ then the agent is risk seeking. When $\alpha = 0$ then the agent is risk neutral and we have a linear utility model. The value $|\alpha|$ represents the intensity of the attitude. That is, given two agents with $\alpha_1 > \alpha_2 > 0$, we say that agent 1 is more risk averse than agent 2. Similarly if $\alpha_1 < \alpha_2 < 0$ then agent 1 is more risk seeking than agent 2. Figure 4.5 illustrates the risk curvature for various risk taking and risk averse α parameters.

Table 4.4 shows sample utility values for various risk taking, neutral, and risk averse utility profiles. These values show how a utility for objects in various ranking positions will change according to risk attitude models. Note that we do normalize the utilities such that all utilities add up to 1.

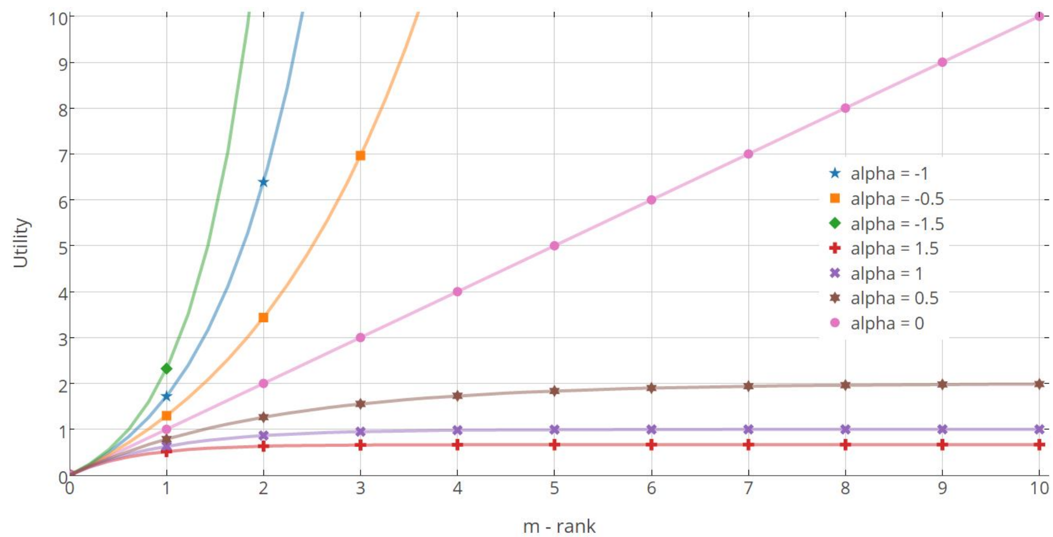


Figure 4.5: Utility values for various α under risk taking, risk neutral, and risk averse models. There are eleven objects ranked from 1 to 11, with linear utilities from 0 (the last object) to 10 (the top choice). The trendlines fit exponential trends to the discrete alpha parameters.

rank / α	$\alpha = -2$	$\alpha = -1$	$\alpha = 0$	$\alpha = 1$	$\alpha = 2$
1	26.799	6.389	2	0.865	0.491
2	3.195	1.718	1	0.632	0.432
3	0	0	0	0	0

Table 4.4: Sample utility values when there are 3 objects under different risk attitudes and risk intensities.

4.5 Results

For our experiments, we vary three parameters: the number of agents n, the number of objects m, and the risk attitude factor α. Each data point in the graphs shows the average over all possible preference profiles. We study the same settings as in Section 4.3 when $n \geq m$ and $n < m$. For each utility function, we look at homogeneous populations of agents where agents have the same risk attitudes but may have difference ordinal preferences.

To compare the social welfare, we investigate the percentage change (or improvement) in social welfare of PS compared to RSD under various utility models. That is,

$$\frac{\sum_i \mathbb{E}(u_i|PS(\succ)) - \sum_i \mathbb{E}(u_i|RSD(\succ))}{\sum_i \mathbb{E}(u_i|RSD(\succ))}.$$

To measure the manipulability of PS, we are interested in answering two key questions: *i*) In what fraction of profiles PS is manipulable by at least one agent? and *ii*) If manipulation is possible, what is the average percentage of maximum gain? That is,

$$\max_i \{\frac{\mathbb{E}(u_i|PS(\succ'_i, \succ_{-i})) - \mathbb{E}(u_i|PS(\succ))}{\mathbb{E}(u_i|PS(\succ))}\}.$$

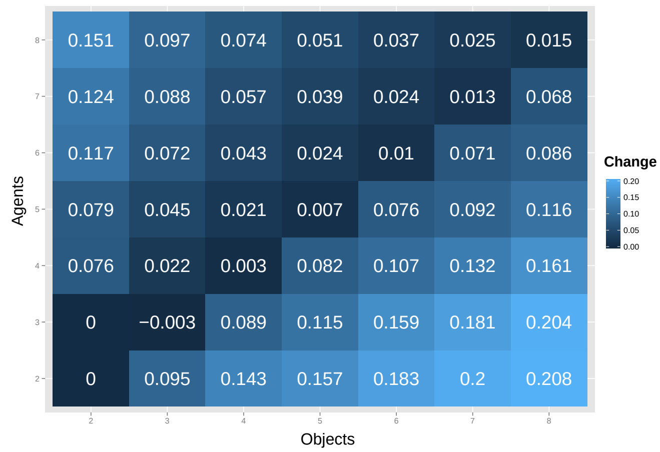
To study the envy under the RSD mechanism, we consider two measures: *i*) the fraction of envious agents, and *ii*) the total envy felt by all agents.

4.5.1 Linear Utility Model

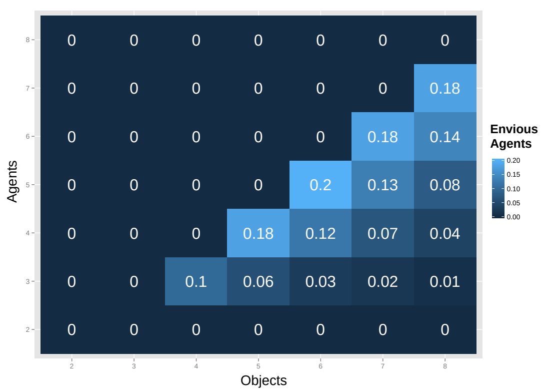
We first looked at how RSD and PS perform under the assumption that the utility models are linear (Figure 4.6). In most cases, the social welfare under PS increases compared to RSD; however, the social welfare of PS is very close to that of RSD when $n = m$ (less than 0.015 overall improvement in all cases). Interestingly, under RSD the fraction of envious agents gets close to 0 when $n \geq m$. With regards to strategyproofness, PS is manipulable in most combinations of n and m and the fraction of manipulable profiles and the utility gain from manipulation increases as the number of objects compared to agents increases.

4.5.2 Risk Seeking

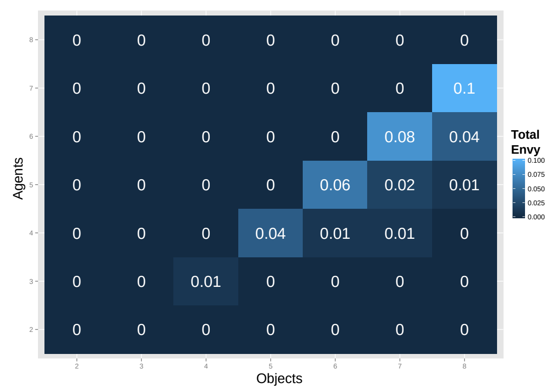
Figure 4.7 presents our results in terms of percentage change in social welfare between PS and RSD. Positive numbers show the percentage of improvement in social welfare from



(a) Welfare change, Linear

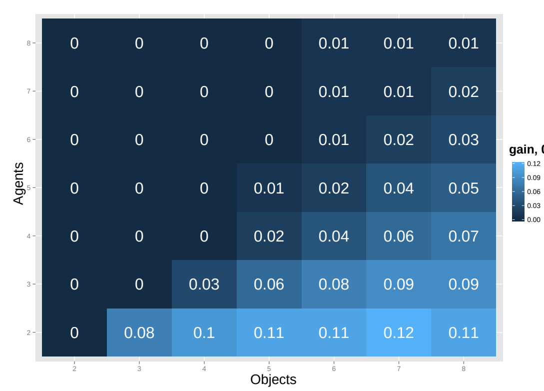


(b) Envious Agents, Linear



(c) Total Envy, Linear

(d) Manipulation, Linear



(e) Manipulation gain, Linear

Figure 4.6: Linear Utility: Welfare Change, Envy, and Manipulation.

PS to RSD. Negative values represent those cases where RSD has increased social welfare compared to PS.

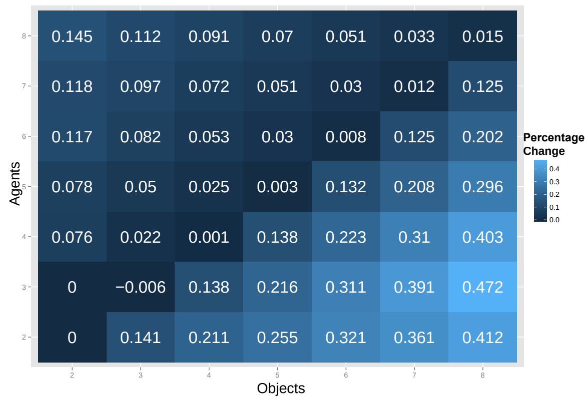
Social welfare: Fixing $\alpha < 0$, for $n \geq m$ when $\frac{n}{m}$ grows PS improves the social welfare compared to RSD in all instances of the problem and the percentage of improvement also increases. A similar trend holds when varying risk intensity α for fixed n and m where $n \neq m$. For $n < m$, when $\frac{m}{n}$ grows the fraction of profiles at which PS has higher social welfare compared to RSD rapidly increases and the percentage change is also noticeably larger, quickly getting close to 90% improvement (Figures 4.7a, 4.7c, and 4.7e). This social welfare gap between PS and RSD grows as the risk intensity $|\alpha|$ increases. Surprisingly, this trend changes for equal number of agents and objects $n = m$: the more risk-seeking agents are (larger $|\alpha|$), RSD becomes more desirable than PS, and in fact, RSD improves the social welfare in more instances.

Envy: Figure 4.8 shows that for $n \geq m$, the fraction of envious agents under all profiles vanishes and RSD becomes envyfree. This is more evident when agents are more risk-seeking. Intuitively, these observations confirm our theoretical findings in Chapter 3 about the envyfreeness of RSD under lexicographic preferences. This is because one can consider lexicographic preferences as risk-seeking preferences where an object in a higher ranking is infinitely preferred to all objects that are ranked less preferably [75]. When $n < m$, our quasi-dictatorial extension of RSD creates some envy among the agents, because the agent with the highest priority receives $m - n + 1$ objects, while all other agents receive at most one object. An interesting result is the envy created by RSD starts to fade out when the risk intensity $|\alpha|$ increases.

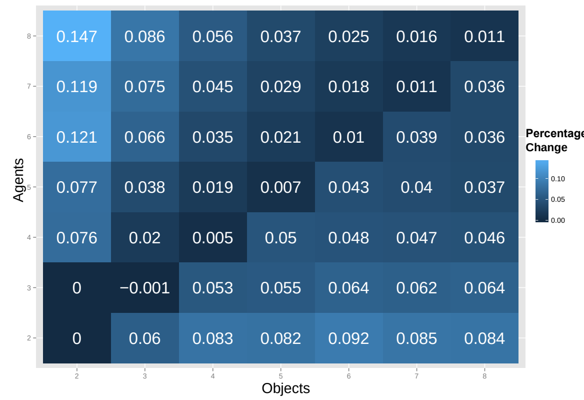
Manipulability: Figure 4.9 shows the manipulability of the PS assignments when agents are risk seeking. We see that the possibility of manipulation (and any gain) decreases as the risk intensity increases. When $n \geq m$ the fraction of manipulable profiles goes to 0 the more risk seeking agents become. However, when $n < m$ even though the fraction of manipulable profiles (and manipulation gain) decreases, the fraction of manipulable profiles goes to 1 as $\frac{m}{n}$ grows.

4.5.3 Risk Aversion

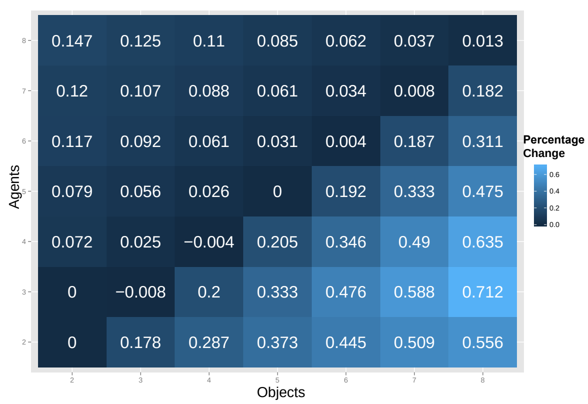
Social welfare: Figures 4.7b, 4.7d, and 4.7f show that fixing risk factor $\alpha > 0$, when $\frac{n}{m}$ grows, PS assignments are superior to that of RSD in terms of social welfare in more instances, and the percentage change in social welfare increases. Fixing risk factor $\alpha > 0$ and when $\frac{m}{n}$ grows, RSD is more likely to have the same social welfare as PS, and in fact in some instances the social welfare under RSD is better than the social welfare under PS.



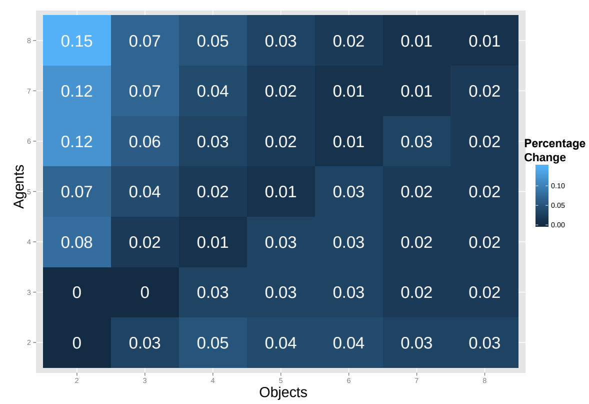
(a) *Risk seeking,* $\alpha = -0.5$.



(b) *Risk averse,* $\alpha = 0.5$.



(c) *Risk seeking,* $\alpha = -1$.

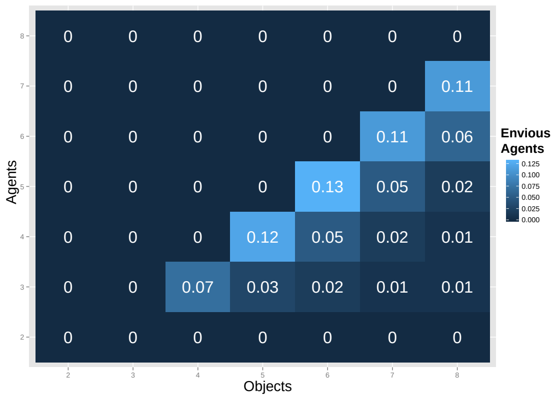


(d) *Risk averse,* $\alpha = 1$.

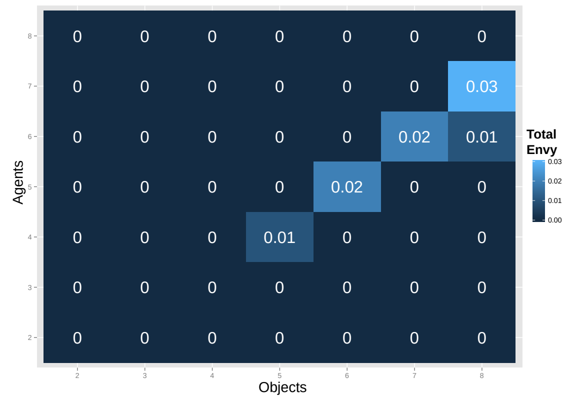
(e) *Risk seeking,* $\alpha = -2$.

(f) *Risk averse,* $\alpha = 2$.

Figure 4.7: The percentage change in social welfare. The negative values show that RSD outperforms PS.

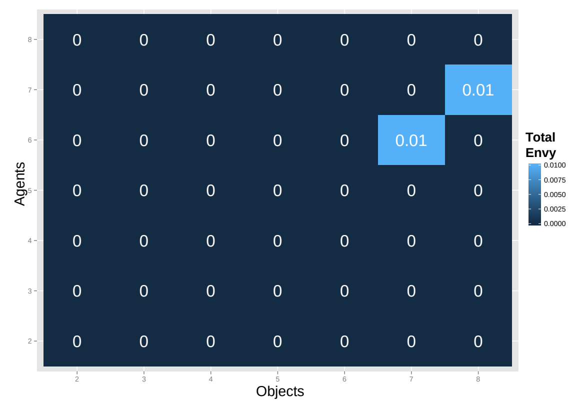


(a) *Fraction of envious agents,* $\alpha = -0.5$.



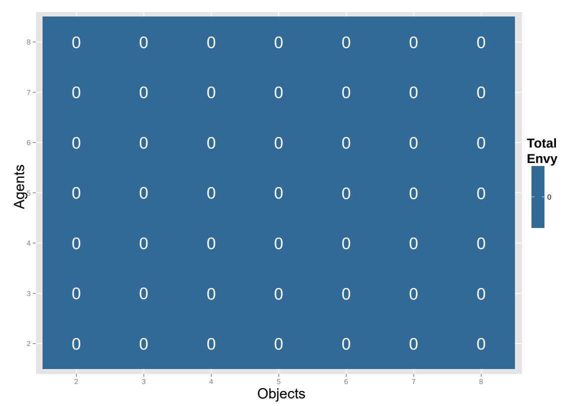
(b) *Total envy,* $\alpha = -0.5$.

(c) *Fraction of envious agents,* $\alpha = -1$.



(d) *Total envy,* $\alpha = -1$.

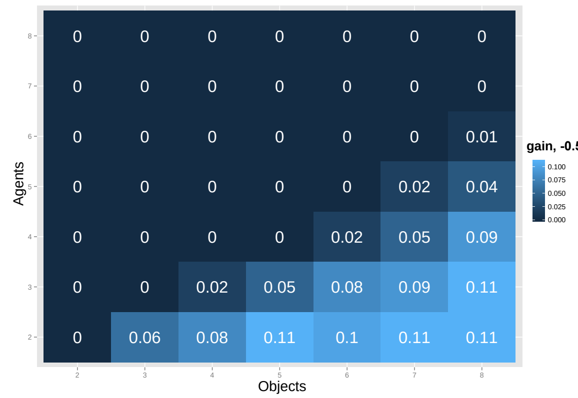
(e) *Fraction of envious agents,* $\alpha = -2$.



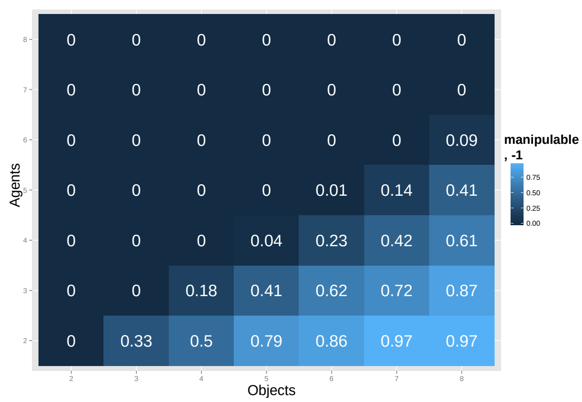
(f) *Total envy,* $\alpha = -2$.

Figure 4.8: The fraction of envious agents and total envy perceived by agents under risk-seeking utilities.

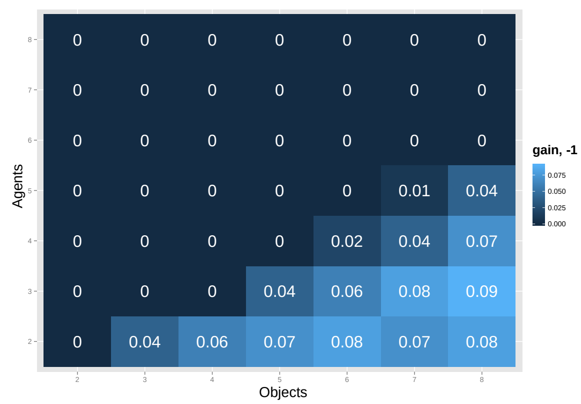
(a) Manipulation, $\alpha = -0.5$.



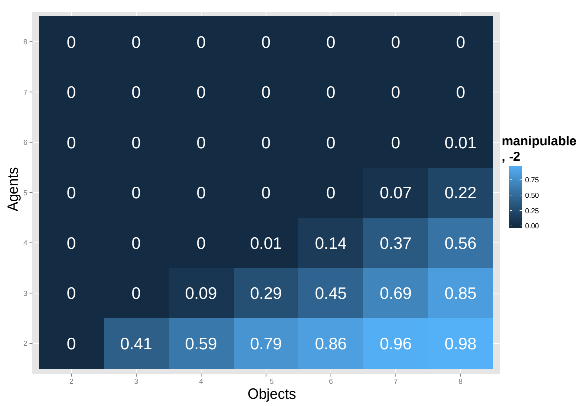
(b) Gain, $\alpha = -0.5$.



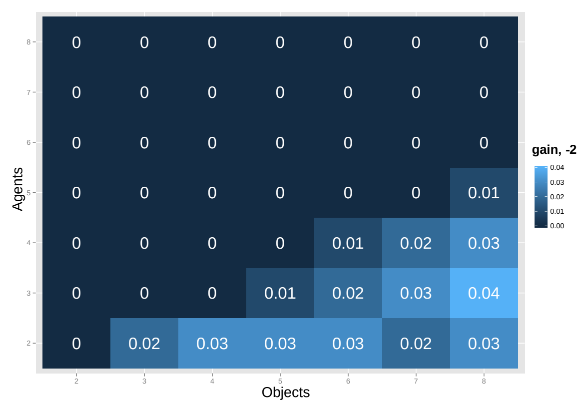
(c) Manipulation, $\alpha = -1$.



(d) Gain, $\alpha = -1$.



(e) Manipulation, $\alpha = -2$.



(f) Gain, $\alpha = -2$.

Figure 4.9: The fraction of manipulable instances and manipulation gain of PS under risk-seeking preferences.

Fixing m and n, when the risk intensity α increases RSD is more likely to have the same social welfare as PS, that is, the welfare gap between PS and RSD closes when agents are more risk averse (α increases). This result is insightful and states that under risk aversion the random allocations prescribed by RSD are either as good as PS or in some cases even are superior to the allocations prescribed by PS due to the underlying shape of the utility models. Figure 4.12 illustrates the percentage change in social welfare based on the difference between available objects and agents ($m - n$) for risk seeking, linear, and risk averse utilities with different risk intensities.

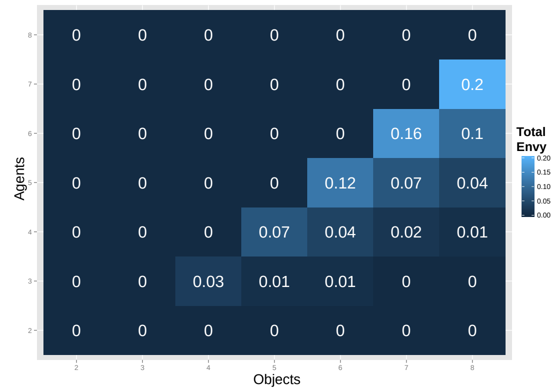
Envy: In Figure 4.10 we observe that when $n \geq m$, the fraction of envious agents and total envy grows as $\frac{n}{m} \to 1$. Increasing the risk intensity ($|\alpha|$), the fraction of envious agents increases; however, the total envy among the agents remains considerably low. For $n < m$, the fraction of envious agents and total envy grows as risk intensity increases. An interesting observation is that envy is maximized when $m = n + 1$, and it decreases as $\frac{m}{n}$ grows. This is mostly due to the choice of using randomized quasi-dictatorial mechanism for implementing RSD where the first dictator receives $m + n - 1$ objects and all other agents only receive a single object. Lastly, we noticed that in all instance where RSD creates envy among the agents, around 25% of agents bear more than 50% of envy. That is, few agents feel extremely envious while all other agents are either envyfree or only feel a minimal amount of envy.

Manipulability: Figures 4.11 illustrates the manipulability of the PS assignments when agents have risk averse preferences. The fraction of manipulable profiles rapidly goes to 1 as $\frac{m}{n}$ grows. Similarly, as agents become more risk averse (α increases) the fraction of manipulable profiles goes to 1 and the manipulation gain increases.

4.6 Related Literature

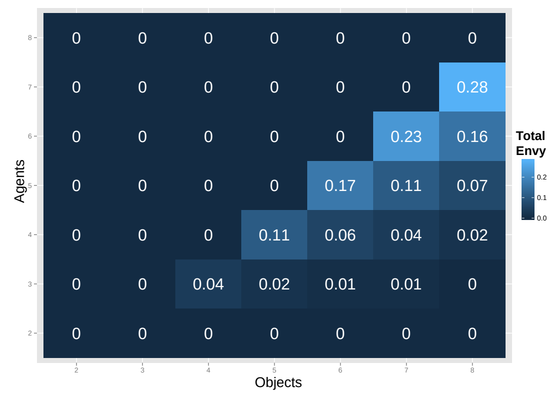
Assignment problems with ordinal preferences have attracted interest from many researchers. Svensson showed that serial dictatorship is the only deterministic mechanism that is strategyproof, nonbossy, and neutral [134]. Random Serial Dictatorship (RSD) (uniform randomization over all serial dictatorship assignments) satisfies strategyproofness, proportionality, and ex post efficiency [6]. Bogomolnaia and Moulin noted the inefficiency of RSD from the ex ante perspective, and characterized the matching mechanisms based on first-order stochastic dominance [30]. They proposed the probabilistic serial mechanism as an efficient and envyfree mechanism with regards to ordinal preferences. While PS is not strategyproof, it satisfies weak strategyproofness for problems with equal number of agents and objects. However, PS is strictly manipulable (not weakly strategyproof) when there

(a) Fraction of envious agents, $\alpha = 0.5$.

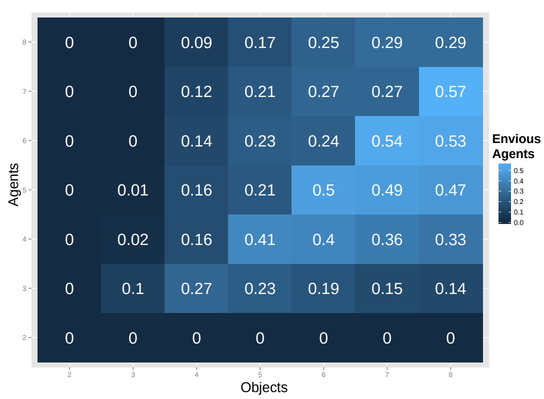


(b) Total envy, $\alpha = 0.5$.

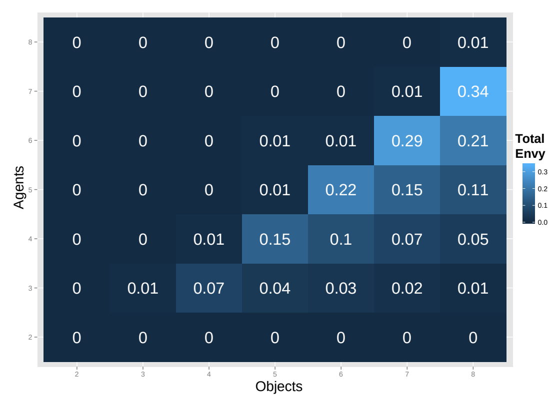
(c) Fraction of envious agents, $\alpha = 1$.



(d) Total envy, $\alpha = 1$.



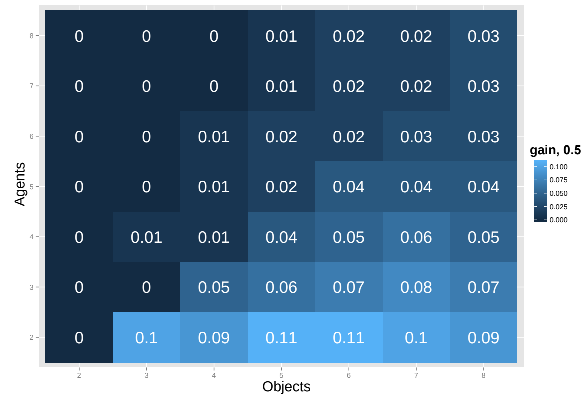
(e) Fraction of envious agents, $\alpha = 2$.



(f) Total envy, $\alpha = 2$.

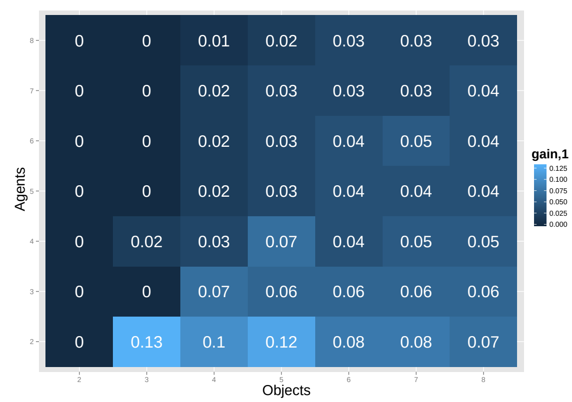
Figure 4.10: The fraction of envious agents and total envy perceived by agents under risk aversion. The total envy is shown up to two decimal points.

(a) Manipulation, $\alpha = 0.5$.



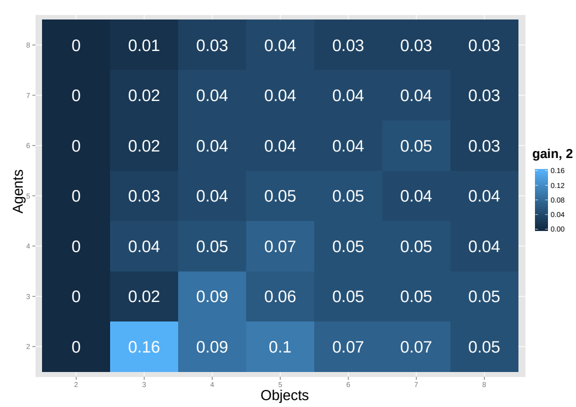
(b) Gain, $\alpha = 0.5$.

(c) Manipulation, $\alpha = 1$.



(d) Gain, $\alpha = 1$.

(e) Manipulation, $\alpha = 2$.



(f) Gain, $\alpha = 2$.

Figure 4.11: The fraction of manipulable instances and manipulation gain of PS under risk aversion.

are more objects than agents [92]. Kojima and Manea, showed that in large assignment problems with sufficiently many copies of each object, truth-telling is a weakly dominant strategy in PS [93]. In fact PS and RSD mechanisms become equivalent [44], that is, the inefficiency of RSD and manipulability of PS vanishes when the number of copies of each object approaches infinity.

The practical implications of deploying RSD and PS have been the center of attention in many one-sided matching problems [3,103]. In the school choice setting with multi-capacity alternatives, Pathak observed that many students obtained a more desirable random assignment through PS in public schools of New York City [113]; however, the efficiency difference was quite small. These equivalence results and their extensions to all random mechanisms [97], do not hold when the quantities of each object is limited to one.

Other interesting aspects of PS and RSD such as computational complexity and best-response strategies have also been explored [19, 20, 56]. In this vein, Aziz et al. proved the existence of pure Nash equilibria, but showed that computing an equilibrium is NP-hard [19]. Nevertheless, Mennle et al. [106] showed that agents can easily find near-optimal strategies by simple local and greedy search. In the absence of truthful incentives, the outcome of PS is no longer guaranteed to be efficient or envyfree with respect to agents' true underlying preferences, and this inefficiency may result in outcomes that are worse than RSD, especially in 'small' markets [56]. The utilitarian and egalitarian welfare guarantees of RSD have been studied under ordinal and linear utility assumptions [18, 26]. For arbitrary utilities, RSD provides the best approximation ratio for utilitarian social welfare when $m = n$ among all mechanisms that rely only on ordinal preferences [57].

4.7 Discussion

We studied the space of general preferences and provided empirical results on the incomparability of RSD and PS. It is worth mentioning that at preference profiles where PS and RSD induce identical assignments, RSD is sd-efficient, sd-envyfree, and sd-strategyproof. However, PS is still highly manipulable. We further strengthen this argument by providing an observation in Example 4.4:

Example 4.4. Consider the following preference profile $\succ = ((bca), (cab), (bca))$. Table 4.5 shows the prescribed random assignment. In this example, with PS as the matching mechanism, agent 1 can misreport her preference as $\succ'_1 = (cba)$, and manipulate her assignment to $1/4(b), 1/2(c), 1/4(a)$. It is easy to see that agent 1's misreport improves her expected outcome for all utility models where $\frac{2}{6}u_1(c) > \frac{1}{4}u_1(b) + \frac{1}{12}u_1(a)$ (for example utilities $10, 9, 0$ for b, c, a respectively.).

	a	b	c
A_1	1/3	1/2	1/6
A_2	1/3	0	2/3
A_3	1/3	1/2	1/6

Table 4.5: A random assignment for a preference profile wherein PS and RSD both prescribe an identical matching, i.e. $PS(\succ) = RSD(\succ)$.

We investigated various utility models according to different risk attitudes. Our main results are:

- In terms of efficiency, the fraction of preference profiles $\succ \in \mathcal{P}^n$ for which PS stochastically (or lexicographically) dominates RSD converges to zero as $\frac{n}{m} \to 1$. When instantiating the preferences with actual utility functions, PS allocations are only slightly better than RSD allocations in terms of social welfare when varying n and m, particularly under risk averse utilities. In fact, in some cases RSD allocations are superior in terms of social welfare (see Figure 4.12).
- PS is almost 99% manipulable when $n \leq m$ and the fraction of *sd*- and *ld*- manipulable profiles rapidly goes to 1 as $\frac{m}{n}$ grows. When instantiating the preferences with utility functions, the manipulability of PS increases as agents become more risk averse. Moreover, an agent's utility gain from manipulation also grows when the risk intensity increases.
- For risk seeking utilities, when $n \geq m$ the fraction of envious agents under all profiles vanishes and RSD becomes envyfree. For risk averse utilities, the fraction of envious agents increases as agents become more risk averse. However, the total amount of envy just slightly grows, and surprisingly, only few agents feel extremely envious while all other agents are either envyfree or only feel a minimal amount of envy.

An interesting future direction is to study egalitarian social welfare of the matching mechanisms in single and multi unit assignment problems as well as in the full preference domain. Another open direction is to provide a parametric analysis of the matching mechanisms according to the risk aversion factor.

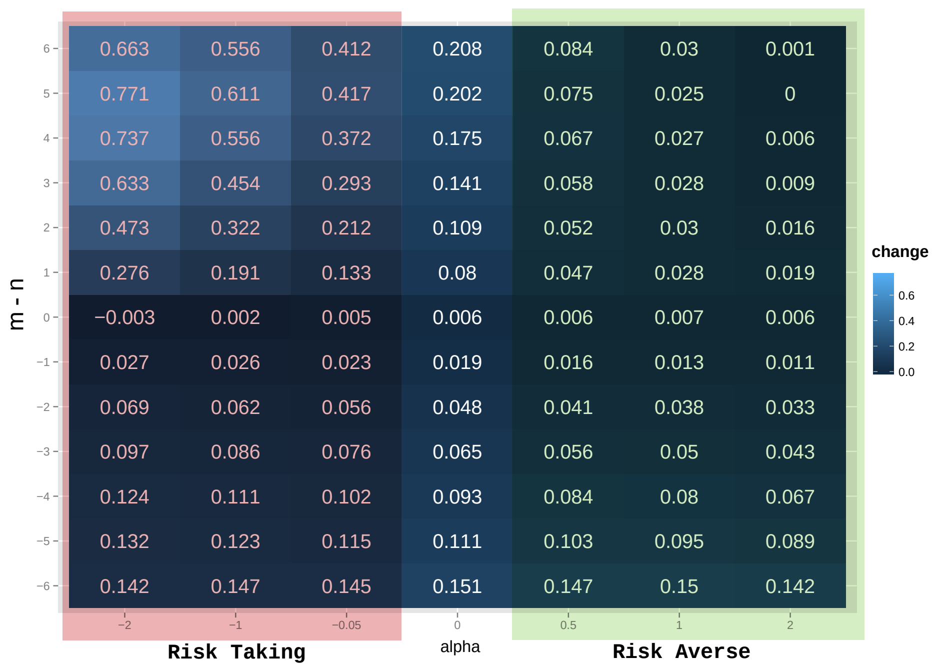


Figure 4.12: The percentage change in social welfare between RSD and PS for $\alpha \in (-2, -1, -0.5, 0, 0.5, 1, 2)$ *and different combinations of* $m - n$*. Positive* α *indicates risk averse and negative* α *risk taking profiles. Linear utility is indicated by* $\alpha = 0$*. As agents become more risk averse the social welfare gap between RSD and PS closes.*

4.8 Design Recommendations for Multiagent Systems

Our work in this chapter can be used to help guide designers of multiagent systems who need to solve allocation problems. If a designer strongly requires *sd*-efficiency then the theoretical results of PS indicate that it is better than RSD. However, our results show that PS is highly prone to manipulation for various combinations of agents and objects. This manipulation and the possible gain from manipulation become more severe particularly when agents are risk averse, and designers need to take this into consideration. On the other hand, while RSD does not theoretically guarantee *sd*-efficiency, our results show that it tends to do quite well – sometimes even outperforming PS in terms of social welfare. RSD also has the added advantage of being *sd*-strategyproof and thus is not prone to the manipulation problems of PS.

Although computing RSD probabilities (fractional assignments) is #P-hard [16, 121], just like PS, RSD is easy to implement in practice. However, the welfare cost of adopting manipulable mechanisms such as PS raises concern and has real consequences [38, 114]. Even though computing optimal manipulation strategies is computationally hard for both PS and RSD, individuals can easily figure out how to manipulate such mechanisms using simple greedy heuristics [38, 106]. Our investigations show that in many instances RSD performs as desirably as PS in terms of social welfare. Conversely, PS assignments are highly susceptible to manipulation especially when agents are risk averse.

These findings suggest that in multiagent settings where mechanism designers are unsure of sincere reporting of their preferences or when agents are mostly risk averse, the use of RSD is more desirable to ensure truthful reporting while providing reasonable social welfare. However, PS is still a desirable allocation mechanism for its fairness and efficiency properties, particularly in settings where agents are sincere.

Part II

Dynamic Matching Mechanisms

Chapter 5

Sequential Matching With Dynamic Preferences

One-sided matching problems have been extensively studied in the context of microeconomics, artificial intelligence, and mechanism design. Despite the interest in matching problems, little work has focused on dynamic settings where agents' preferences evolve over time. In the real world, decisions do not exist in isolation, but rather are situated in a temporal context with other decisions and possibly stochastic events. Dynamic mechanism design [111] is a compelling research area that has attracted attention in recent years. In these settings, agents act to improve their outcomes over time, and decisions both in the present and in the past influence how the preferences look in the future. The dynamic pivot mechanism for dynamic auctions [25], dynamic Groves mechanisms [41], and many others [15, 140] are a few of myriad examples of mechanisms in dynamic settings that consider agents with private dynamic preferences. However, almost all of these works (excluding a recent study on dynamic social choice [112]) assume an underlying utility function with possible utility transfers.

In this chapter, we study dynamic matching problems in which a sequence of decisions must be made for agents whose private underlying preferences may change over time.[1] In each period, the mechanism elicits ordinal preferences from agents, and each agent declares its preferences (truthfully or strategically) so as to improve its overall assignment, now or in the future. We propose a generic model to study the various properties of such environments including the strategic behavior of agents. Our model captures a diverse

[1] We use the terms matching, allocation, and assignment interchangeably, and further describe our analogous terminologies in dynamic settings.

set of real-life scenarios: assigning members to subcommittees each year, tasks among team members on various projects, nurses to various hospital shifts, teaching loads among faculty, and assigning students to college housing each year.

Consider the problem of scheduling nurses to shifts (nurse rostering) in multiple planning periods.[2] Self-rostering is one of the most advocated methods in nurse scheduling that caters to individual preferences [40, 129]. Each nurse has some internal preference over hospital shifts at various times. The process is as follows: every 4–6 weeks, nurses log in to an online system which enables them to rank their preferences for various shifts (day, night, weekends, etc.) during that period. At each planning period (typically 4–6 weeks), an assignment decision is made based on the self-reported preferences. This is in contrast with traditional nurse scheduling systems which do not take nurse preferences into account. Although self rostering reduces administrative burden and improves nurse satisfaction, there is evidence that it encourages strategic behavior among nurses [10, 49]. The preferences of nurses may change dynamically according to their internal desires and past assignments and "those who are savvy enough to game the system will always have an advantage over the procrastinators" [24]. For example, knowing how more challenging (or less challenging) shifts would affect the preferences of less experienced nurses, a more senior nurse, given the information about demands and severity of various shifts, may strategically misreport its preference to influence the preferences of other agents in hopes of benefiting in the future periods. This example and many other real-life applications raise several intriguing questions when designing matching mechanisms in dynamic and uncertain settings.

5.1 Our Model and Contributions

We consider a setting where a sequence of assignments should be made for a fixed number of agents and alternatives. Agents' preferences are represented as strict orderings over a set of alternatives, where these preferences may change over time. Our contributions in this chapter are:

- We initiate the study of repeated matching with dynamic preferences and provide a framework for modeling and analysis of ordinal matching decisions in uncertain domains. We formulate a general dynamic matching problem using a *history-dependent matching process*. The state of the matching process corresponds to a history of preference profiles and matching decisions.

[2] Assume skill category and hospital constraints are incorporated into shift schedules.

- We show that one cannot simply deploy conventional matching mechanisms to dynamic settings, because strategyproof mechanisms for static and one-shot settings are prone to manipulation when preferences change dynamically over time. More specifically, We show that simply running a sequence of independent assignments induced by a strategyproof mechanism in one-shot settings, namely Random Serial Dictatorship (RSD), does not satisfy global strategyproofness. Furthermore, we investigate weak and strong manipulation in sequential matching with dynamic preferences.

- Subsequently, we design a stochastic matching policy based on the RSD mechanism which takes agents' histories into account. Our key idea is to extend a notion that was first introduced for multi-period contracts [136], where future matching decisions are used to incentivize desirable behavior in the current time period. Nonetheless, we argue that in the absence of monetary transfers, designing generic mechanisms that provide truthful incentives in dynamic settings raises several complications.

- Our proposed mechanism satisfies some notion of fairness in repeated matching problems while providing incentives for truthfulness when agents have similar preference dynamics. Moreover, given any history of preferences and decisions and particular preference dynamics, our mechanism prevents any harmful manipulation while satisfying desirable local properties of ex post efficiency, strategyproofness, and equitability. We further show that under some mild assumptions an agent's successful manipulation (if possible) results in a Pareto improving sequence of matchings in expectation. Finally, we formulate the notion of envy in the context of sequential matching by providing a systematic way of measuring the degree of envy towards individual agents, arguing that our mechanism provides a constant degree of individual envy by balancing priority orderings.

5.2 A Matching Model for Dynamic Ordinal Preferences

In this section, we introduce our model for matching in sequential settings when there are dynamic preferences. We start by introducing key preference and matching terminologies used in static settings. We then generalize them to the dynamic setting studied in this chapter.

5.2.1 Basic Terminology

There is a set $N = \{1, \ldots, n\}$ of agents who have preferences over a finite set of alternatives $M = \{1, \ldots, m\}$, where $n \geq m$.[3] Agents have private preferences over alternatives. We use the notation $a \succ_i b$ to mean that agent i strictly prefers alternative a to alternative b. We let $\mathcal{P}(M)$ or $\mathcal{P}$ denote the class of all strict linear preferences over M where $|\mathcal{P}| = m!$. Agent i's preference is denoted by $\succ_i \in \mathcal{P}$, thus, $\succ = (\succ_1, \ldots, \succ_n) \in \mathcal{P}^n$ denotes the *preference profile* of agents. We write $\succ_{-i}$ to denote $(\succ_1, \ldots, \succ_{i-1}, \succ_{i+1}, \ldots, \succ_n)$, and thus $\succ = (\succ_i, \succ_{-i})$.

A *matching*, $\mu : N \rightarrow M$, is a mapping from agents to alternatives. We let $\mu(i)$ denote the alternative allocated to agent i under matching μ. A matching is *feasible* if and only if for all $i, j \in N$, $\mu(i) \neq \mu(j)$ when $i \neq j$. We let $\mathcal{M}$ denote the set of all feasible matchings over the set of alternatives M. We also allow for randomization where $\bar{\mu}$ denotes a probability distribution over the set of (deterministic) feasible matchings. That is, $\bar{\mu} \in \Delta(\mathcal{M})$.

5.2.2 Dynamic Preferences

The decision problem in many multiagent settings is dynamic rather than static. Agents participate in a sequence of allocation decisions and their preferences may change (or evolve) over time. For example, in a nurse scheduling scenario, one might prefer a day shift to a night shift on weekdays but prefer a night shift over a day shift on weekends. The preferences could be more complicated when considering the subtle relations between allocations: a day shift immediately after a night shift is often considered extremely undesirable. Another interesting example is the problem of assigning computational resources to scientific groups over time. A researcher may require access to a powerful processor for running high-performance computations while in the next occasion she prefers a less powerful core but for a longer period of time. The preference evolution may also be dependent on previous allocations or some idiosyncratic dynamics. In our example, a researcher's preference may change given the allocation of a high-performance core (ergo the completion of a task).

Let $o_1, o_2, \ldots, o_m$ denote the set of objects in ranks $1, 2, \ldots, m$ respectively. Thus, if $a \succ_i b$ then $o_1 = a$ and $o_2 = b$. We use this notation because objects may be ranked differently in each period. The preference ordering of agent i at time t is denoted by $\succ_i^t$.

[3] We accommodate the possibility of $n > m$ by adding dummy alternatives corresponding to a null assignment.

For the sake of clarity, at each time t we rank alternatives under preference ordering $\succ_i^t$ such that $o_1 \succ_i^t o_2 \succ_i^t \ldots \succ_i^t o_m$ denotes the ranking positions, where the index of each alternative indicates its ranking for agent i at time t. Formally, for each agent $i \in N$ it may be the case that $a \succ_i^t b$ while $b \succ_i^{t+1} a$. Thus, it is possible that $\succ_i^t \neq \succ_i^{t'}$ for two periods $t \neq t'$. A preference profile at time t is denoted by $\succ^t = (\succ_1^t, \ldots, \succ_n^t) \in \mathcal{P}^n$. Following the convention, we write $\succ_{-i}^t$ to denote the preferences of all agents except i, thus $\succ^t = (\succ_i^t, \succ_{-i}^t)$.

Example 5.1. Consider three agents 1, 2, and 3 and three objects a, b, and c. Table 5.1 illustrates the preference of each agent in different time steps. The preferences of agents 1 and 2 changes from $t = 1$ to $t = 2$ while agent 3's preference ordering remains unchanged.

	$t = 1$	$t = 2$	$t = 3$
1	$a \succ_1^1 c \succ_1^1 b$	$a \succ_1^2 b \succ_1^2 c$	$\ldots$
2	$b \succ_2^1 c \succ_2^1 a$	$b \succ_2^2 a \succ_2^2 c$	$\ldots$
3	$a \succ_3^1 c \succ_3^1 b$	$a \succ_3^2 c \succ_3^2 b$	$\ldots$

Table 5.1: Three agents with dynamic preferences over time.

There are various plausible reasons as to why an agent's preference may change over time. We will elaborate on these changes in preferences, and further discuss a model where preferences change depending on previous assignments or uncertain idiosyncratic reasons.

5.2.3 Sequential Matching Mechanisms

In many real-life applications, a set of alternatives are to be allocated repeatedly over time to a set of agents. In these settings, a series of matching decisions has to be made sequentially over time, with new private information potentially obtained after each matching decision. Figure 5.1 illustrates a discrete time-line where a (random) matching decision is made in each period.[4] We are interested in settings where agents' preferences *evolve* over time. In particular, we assume that the preference held by an agent at time t depends on the preferences it held earlier along with allocations (*i.e.* matchings) made previously.

We model the sequential matching mechanism with a discrete-time stochastic decision process. A matching process consists of a set of joint *states* denoted by preference profiles, a set of joint *actions* corresponding to matching decisions, and a *transition* model

[4]Note that this is different from multiple-assignment problems [55, 110] that was discussed in Chapter 3, where multiple objects can be assigned to each agent. Here, the set of objects is fixed in all sequences, which are then re-assigned to the agents in each step of the repeated matching.

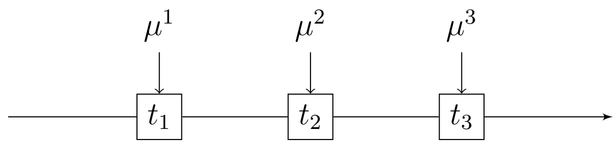


Figure 5.1: Timeline for a sequence of matching decisions.

which specifies a possibly stochastic transition function given an allocation determined by a matching decision and previous states.

We denote a matching at time t by μ^t. We let h^t denote the joint history of the joint states defined by agents' preferences and realized matchings up to time $t-1$ and including the joint preferences (preference profile) at time t, that is, $h^t = (\succ^1, \mu^1, \ldots, \succ^{t-1}, \mu^{t-1}, \succ^t)$. The set $\mathcal{H}^t$ contains all possible joint histories at time t. In other words, $h^t \in \mathcal{H}^t$ indicates the *trajectory* of states and actions from the first time step until time t.

5.2.4 Transition Models

A matching mechanism that bases the decisions only on the current state, and not any additional information in the history of the process, is Markovian and requires that the transition probabilities to also be Markovian. Following the convention, we distinguish between Markovian (or memoryless) transitions and history-dependent transitions [115].

Assuming a memoryless transition function, we define a *Markovian model* (first order Markov process) where the transitions are independent of the past history of actions and decisions, and transitioning to the next state $\succ^{t+1}$ is only dependent on the matching decision μ^t in a state with joint preference $\succ^t$ (see Figure 5.2).

Definition 5.1 (Markovian transition)**.** *A transition function is Markovian if the transitions in each state only depend on the current state and action, that is,* $T(\succ^{t+1} \mid \succ^t, \mu^t)$.

For a single agent i, the probability that agent i's preference ordering changes to $\succ_i^{t+1}$ in the next time step, given its current preference $\succ_i^t$ and its allocation $\mu^t(i)$ is denoted by $T_i(\succ_i^{t+1} \mid \succ_i^t, \mu^t(i))$. For all agents, the probability of transitioning to preference profile $\succ^{t+1}$ given the current matching and preference profile is written as $T(\succ^{t+1} \mid \succ^t, \mu^t)$, and we have

$$\sum_{\succ^{t+1} \in \mathcal{P}^n} T(\succ^{t+1} \mid \succ^t, \mu^t) = 1$$

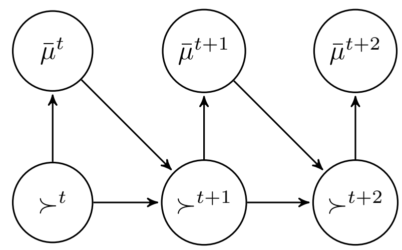


Figure 5.2: Influence diagram representing the conditional dependencies in a Markovian matching process.

Throughout this work, we assume independence of priors and transitions and write

$$T(\succ^{t+1} \mid \succ^t, \mu^t) = \prod_{i \in N} T_i(\succ_i^{t+1} \mid \succ_i^t, \mu^t(i))$$

We also consider a different dynamic model where the transitions depend on the history of previous preferences and matching decisions. Thus, we sometimes assume there is an underlying *history-dependent stochastic model* $T(h^{t+1}|h^t, \mu^t)$ which denotes the probability that agents will transition to joint state h^{t+1} after trajectory h^t and matching decision μ^t. Similar to the Markovian model described above, throughout this work, we assume independence of priors and transitions. Figure 5.3 illustrates the conditional dependencies of a history-dependent process. The transition function assumes that

$$\sum_{h^{t+1} \in \mathcal{H}^{t+1}} T(h^{t+1}|h^t, \mu^t) = 1$$

Definition 5.2 (History-dependent transition)**.** *A transition function is history-dependent if the probability of transitioning to the next state depends on the history up to time t plus the decision made at time t, that is, $T(h^{t+1}|h^t, \mu^t)$.*

Clearly, a history-dependent transition can be thought of as a Markovian transition of order λ, where λ is a finite planning horizon. Recall that a trajectory (or execution history) is the entire sequence of state-action pairs up to time t plus the state visited at time t. A history-dependent function is a more generic model where transitions could potentially be influenced by the complete history of preferences and decisions. We study both of these transition models, and following the convention in dynamic mechanism design literature [25, 42, 112], assume that the stochastic model, regardless of being Markovian or history-dependent, is *common knowledge*.

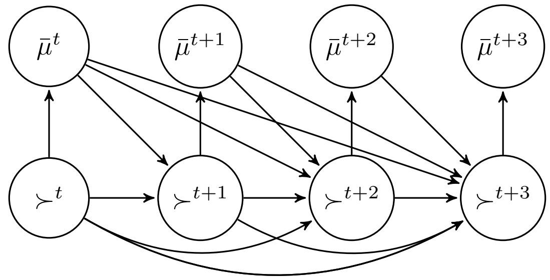


Figure 5.3: Influence diagram representing the conditional dependencies in a history-dependent matching process.

5.2.5 Sequential Matching Policy

A policy prescribes a matching decision to be made at each period, given the information available to the decision maker at that point in time, which may consist of the entire sequence of states and actions up to the present time step. We define a discrete-time *sequential matching decision process* as a sequence of matchings prescribed by a matching policy. A matching policy is a sequence of (possibly randomized) matchings, that is, $\pi = (\bar{\mu}^1, \bar{\mu}^2, \ldots)$. We assume *stationary* policies where the matching mechanism does not change over time. Given a history $h^t \in \mathcal{H}^t$, a *matching policy* $\pi(\mu|h^t)$ returns the probability of applying matching μ. Given a matching policy, π, and a history h^t the probability of agent i being allocated alternative x at time t is

$$p_i^t(x \mid h^t) = \sum_{\mu \in \mathcal{M}:\mu(i)=x} \pi(\mu \mid h^t). \tag{5.1}$$

where $\sum_{i\in N} p_i^t(x \mid h^t) = 1$. The definition of a matching policy, π, incorporates randomized or deterministic matching policies. We denote the (random) allocation of agent i by $\pi_i(h^t)$. When it is clear from the context, we will abuse notation and use $\pi(h^t)$ to also refer to the (random) matching prescribed by policy π given the history h^t.

Example 5.2. Consider two decision periods and preferences as defined in Example 5.1. In Table 5.2, policy π prescribes feasible matchings $\mu^1 = \pi(\succ^1)$ and $\mu^2 = \pi(\succ^2)$ in the first and the second periods.

	$t = 1$	$t = 2$
1	$\mu^1(1) = a$	$\mu^2(1) = b$
2	$\mu^1(2) = b$	$\mu^2(2) = c$
3	$\mu^1(3) = c$	$\mu^2(3) = a$

Table 5.2: A deterministic policy for two time periods.

5.2.6 Agents' Strategies

We are interested in *dynamic mechanisms* in which each agent interacts with the mechanism simply by declaring its (perhaps untruthful) preference ordering at each decision period. These mechanisms are desirable in practice because they remove the computational and cognitive burden on agents by eliciting agents' preferences at each time step. In contrast, under one-shot allocation mechanisms, agents are required to report preferences for all future (possibly uncertain) periods at the beginning, which could be computationally challenging.

In the dynamic mechanism, at each period t, the mechanism elicits reports from agents regarding private preferences. Each agent observes its private preference $\succ_i^t$, and the history of past preferences and realized matchings h_i^{t-1}, and based on the underlying transition model, takes an action by declaring a preference ordering for time t, $\succ_i^t$. Then, given the local elicited preferences, in each period the mechanism draws a matching $\mu^t \in \mathcal{M}$ according to matching policy $\bar{\mu}^t = \pi(\succ^t)$. Figure 5.4 shows these steps for two periods of matching decisions.

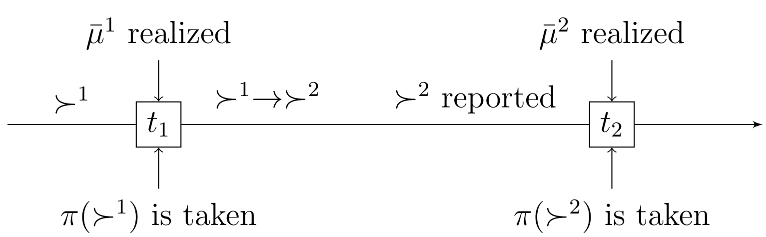


Figure 5.4: Timeline of two decisions and evolution of preferences.

In each period, an agent's action is to report an ordering of the objects according to a strategy, which does not need to be equivalent to the agent's true underlying preference. Agent i's reporting *strategy* $\sigma_i(\succ_i^t)$ specifies its declared preference $\hat{\succ}_i^t = \sigma_i(\succ_i^t)$ when its true private preference is $\succ_i^t$. An agent may potentially be aware of its entire history of

(a) Preference profile at $t = 1$		(b) Preference profile at $t = 2$	
$\succ_1^1$:	$a \succ c \succ b$	$\succ_1^2$:	$a \succ c \succ b$
$\succ_2^1$:	$b \succ c \succ a$	$\succ_2^2$:	$b \succ c \succ a$
$\succ_3^1$:	$a \succ c \succ b$	$\succ_3^2$:	$a \succ b \succ c$

Table 5.3: Preferences revealed by three agents at $t = 1$ and $t = 2$.

preferences and allocations, h_i^t. For simplicity, we use $\sigma_i(\succ_i^t)$ instead of $\sigma_i(h_i^t)$ throughout this chapter but assume that in each period an agent may have access to its entire history.

Example 5.3. There are three agents and three objects and two-time periods. Table 5.3a illustrates the preferences reported by agents at $t = 1$. After a matching decision $\bar{\mu}^1$ is realized, agent's preference may change to the preferences shown in Table 5.3b.

5.3 Evaluating Sequential Outcomes

A policy prescribes a matching given the current preference profile. To determine whether a particular matching policy, π, is a *good* policy, one must make comparisons between policies. In this chapter, we are interested in settings where agents have *ordinal* preferences and so do not rely on particular utility functions. Instead, we compare two sequences of matchings using a *scoring function* based on the expected sum of probabilities that a sequence assigns a more preferred object. For example in a two-period setting, assigning the top-ranked object to an agent in both periods is preferred to assigning its top-ranked object in the first period and the last-ranked object in the second period.

Example 5.4. Consider the preferences revealed by agent 3 in Table 5.3. Let (x, y) be a sequence of decisions in the first and the second periods, respectively. Agent 3 would prefer receiving (a, a) to (c, a) and (c, a) to (a, c), because it preferred c more in the first period than the second period.

Given a policy π, we can evaluate it by looking at the *score* of being allocated alternatives in a particular ranking position in the sequence of random matchings from time t onward. More concretely, let o_ℓ be any alternative ranked in position ℓ. Given h^t and transition function T, the score that agent i receives alternatives with rankings as good as ℓ under matching policy π is defined recursively as

$$W_i^\pi(h^t, o_\ell) = \sum_{x=o_1}^{o_\ell} p_i^t(x|h^t) \ + \gamma \sum_{\mu \in \mathcal{M}} \sum_{h^{t+1} \in \mathcal{H}^{t+1}} \pi(\mu|h^t) T(h^{t+1}|h^t, \mu) W_i^\pi(h^{t+1}, o_\ell) \qquad (5.2)$$

where $h^{t+1} = (h^t, \mu^t, \succ^{t+1})$ is the history at time $t+1$, and $0 < \gamma \leq 1$ denotes the *discounting factor*. Throughout this chapter, we always assume that the future outcomes are as important as the immediate outcomes, and thus $\gamma = 1$.[5] Intuitively, when $\gamma = 1$ then $W_i^\pi(h^t, o_\ell)$ is the sum of probabilities that agent i receives alternatives as good as ℓ in current period and all future sequences of matchings up to a desired *planning horizon* λ. For Markovian processes, we can simplify the above formulation by replacing h^t with $\succ^t$. Example 5.5 shows how two sequential outcomes can be compared using our scoring function.

Example 5.5. In this example, we show how to compare two sequential outcomes using our scoring function. Consider assigning 3 objects in two decision periods to 3 agents with preference orderings, as shown in Figure 5.5, at periods 1 and 2. For simplicity assume deterministic transitions that are independent of the realized matching decisions, and let $h^1 = \succ^1$ so that the matching starts at $t = 1$. A matching $\mu = xyz$ denotes that agents $1, 2, 3$ receive objects x, y, z respectively. Consider a policy π that prescribes random matchings $\bar{\mu}^1 = (\frac{1}{2}abc, \frac{1}{2}cba)$ and $\bar{\mu}^2 = (\frac{1}{2}abc, \frac{1}{3}acb, \frac{1}{6}cab)$ at periods 1 and 2 respectively. Using Equation 5.3, the score that agent 1 receives its first rank alternatives (*i.e.* o_1) in periods 1 and 2 given the above random decisions is calculated by (highlighted as the left subtree in Figure 5.5):

$$W_1^\pi(\succ^1, o_1) = (\tfrac{1}{2}) + (\tfrac{1}{2} + \tfrac{1}{3}) = \tfrac{8}{6}$$

Similarly, we can compute the score that agent 1 receives alternatives with rankings as good as rank 2 (*i.e.* o_1 and o_2) in periods 1 and 2:

$$W_1^\pi(\succ^1, o_2) = (\tfrac{1}{2} + \tfrac{1}{2}) + (\tfrac{1}{2} + \tfrac{1}{3}) = \tfrac{11}{6}$$

Intuitively, $W_1^\pi(\succ^1, o_2)$ is the score of the following possible sequences $[\mu^1(1) = o_1, \mu^2(1) = o_1]$, $[\mu^1(1) = o_1, \mu^2(1) = o_2]$, $[\mu^1(1) = o_2, \mu^2(1) = o_1]$, $[\mu^1(1) = o_2, \mu^2(1) = o_2]$ (highlighted areas in Figure 5.5).

The outcome of a matching policy for all rankings can be represented as the sum of weights for alternatives in all ranking position $\alpha_1 o_1 + \alpha_2 o_2 + \ldots + \alpha_m o_m$, where o_ℓ represents objects in ranking ℓ and coefficients α_j are real non-negative numbers equal to $W_i^\pi(h^t, o_\ell)$. Thus, given h^t and transition function T, an ordinal sequential outcome for agent i can be written as

$$W_i^\pi(h^t) = \sum_{x \in M} W_i^\pi(h^t, x)\ x \tag{5.3}$$

[5] Without loss of generality, our results can be extended to problems with diminishing discount factors.

$$\begin{pmatrix} \succ_1^1\colon & a \succ c \succ b \\ \succ_2^1\colon & b \succ c \succ a \\ \succ_3^1\colon & a \succ c \succ b \end{pmatrix} \begin{pmatrix} \succ_1^2\colon & a \succ b \succ c \\ \succ_2^2\colon & b \succ a \succ c \\ \succ_3^2\colon & b \succ c \succ a \end{pmatrix}$$

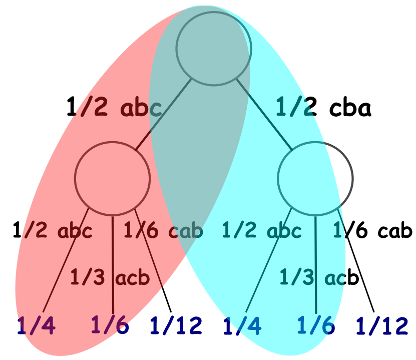


Figure 5.5: The preferences of agents and the probability tree for the two-time matching problem in Example 5.5.

When there is only one matching period, the above formulation can be seen as a convex combination where $\sum_j \alpha_j = 1$.

Our scoring function extends the concept of *stochastic dominance* [69] to dynamic and repeated settings, and enables us to compare a large set of matching decisions by looking at fractional ordinal allocations, without having any information about the underlying utility models. Even though matching under ordinal preferences provides a natural framework for a large class of allocation problems, it opens up new challenges with regards to expressiveness of ordinal allocations. Next, we discuss this issue of expressiveness when comparing allocations under ordinal preferences.

5.3.1 Incomparability of Sequences

In the absence of cardinal utilities, two matching sequences can potentially be incomparable under ordinal preferences. This incomparablity states that the quality of a sequence of matchings compared to another sequence depends on the underlying utility functions. The following example illustrates this incomparability under ordinal outcomes.

Example 5.6. Consider the problem of assigning three objects to agent i in two periods, where $a \succ_i b \succ_i c$ is a fixed preference in both periods. Let (x, y) be a sequence of deterministic decisions that assigns object x in period 1 and y in period 2. There are 9 possible sequences of assignments; $(a, a), (a, b), (b, a), (a, c), (c, a), (b, b), (b, c), (c, b), (c, c)$. Clearly, assigning the top choice in both steps, (a, a), is the best sequence and assigning the last choice (c, c) in both steps is the worst sequence. Since we assumed no discounting, sequences like (a, b) and (b, a) are treated as equal. Every pair of these sequences are easily comparable, meaning that for any utility model consistent with the agent's preference one sequence is always preferred to another sequence. However, two sequences (a, c) and (b, b) are not comparable according to agent i's preferences: (a, c) is preferred to (b, b) if $u_i(a) = 10, u_i(b) = 2$, and $u_i(c) = 1$, while (b, b) is preferred to (a, c) under $u_i(a) = 10, u_i(b) = 9$, and $u_i(c) = 1$.

Similarly, a sequence of randomized matchings prescribed by a policy may be incomparable with a sequence prescribed by another matching policy. Later in this chapter, we define how incomparability of outcomes come to play when investigating various properties such as strategyproofness and manipulation. In the ordinal domain, we define *weak* notions for our desirable properties, providing means to measure these in-comparabilities in sequential settings.

5.4 Properties for the Model

Our goal is to implement stochastic matching policies so that agents truthfully reveal their preferences, no matter what other agents do, now or in the future. We are interested in preserving the local properties in each matching period while satisfying global properties for a sequence of matchings. Local properties hold valid for each matching decision in each step, independent of future or past decisions, while global properties hold over sequences of decisions. More specifically, we are interested in matching policies that satisfy global strategyproofness while inducing a sequence of locally strategyproof and ex post efficient random matchings. In this section, we formally define these local and global properties.

5.4.1 Local Properties

In each period, a matching decision must satisfy some desired local properties, regardless of any of the past or future matching decisions. A random matching induced by matching mechanism π stochastically dominates another random matching induced by π' if for each item $x \in M$, the probability of selecting an outcome as good as x by π is greater than or equal to π'.

Definition 5.3. *Given a preference profile $\succ^t$ at time t, a random matching induced by π* ***stochastically dominates (sd)*** *another random matching prescribed by π', if for all* *agents $i \in N$,*

$$\forall y \in M, \sum_{\substack{x \in M: \\ x \succ_i^t y}} \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i)=x}} \pi(\mu| \succ^t) \geq \sum_{\substack{x \in M: \\ x \succ_i^t y}} \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i)=x}} \pi'(\mu| \succ^t)$$

In other words, in a single-shot setting, every agent with any utility model that is consistent with the private ordinal preferences will prefer π over π' if the random matching selected by π first-order stochastically dominates the random matching selected by π'. A single-shot matching process is a special case that coincides precisely with the random assignment problem [6, 30].

A matching policy is *locally sd-strategyproof* in period t, if an agent's truthful report always results in a random matching that stochastically dominates its random matching under an untruthful misreport. In other words, at each period t no agent can improve its current random assignment by a strategic misreport.

Definition 5.4. *A matching policy π is **locally sd-strategyproof (lsd-strategyproof)** if and only if truthfulness is a stochastic dominant strategy at all times t, that is, $\forall t$ for any $\succ^t \in \mathcal{P}^n$, for all agents $i \in N$, $\forall y \in M$*

$$\sum_{\substack{x \in M: \\ x \succ_i^t y}} \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i)=x}} \pi(\mu|(\succ_i^t, \succ_{-i}^t)) \geq \sum_{\substack{x \in M: \\ x \succ_i^t y}} \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i)=x}} \pi(\mu|(\sigma_i(\succ_i^t), \succ_{-i}^t))$$

Incentive compatibility notions generally guarantee that no agent can benefit from misreporting preferences. However, an agent may still behave in a "bossy" manner by affecting the prescribed random allocation for some other agents while keeping its allocation unchanged. Satterthwaite and Sonnenschein [122] first introduced the *non-bossiness* property as one of the desirable axioms in matching and assignment problems.

A mechanism is *non-bossy* if an agent cannot change the random allocation without changing the random allocation for itself. The non-bossiness property is particularly important in sequential settings as a non-truthful agent may benefit from changing the evolution of other agents' preferences and consequently improve its sequence of outcomes in expectation.

Definition 5.5. *A mechanism is **non-bossy** at time t if for all $\succ^t \in \mathcal{P}^n$ and agent $i \in N$, for all $\hat{\succ}_i^t$ such that $\pi_i(\succ^t) = \pi_i((\hat{\succ}_i^t, \succ_{-i}^t))$ we have $\pi(\succ^t) = \pi((\hat{\succ}_i^t, \succ_{-i}^t))$.*

A matching is *Pareto efficient* if there is no other matching that makes all agents weakly better off and at least one agent strictly better off.

Definition 5.6. *A random matching is **ex post efficient** if it can be represented as a probability distribution over Pareto efficient deterministic matchings.*

5.4.2 Global Properties

In sequential settings, the desirability of a policy is often characterized by their global properties. In contrast to local properties that only hold within each period, the global properties should hold over any sequence of decisions, and thus, guarantee the effectiveness of policies when operating in dynamic and repeated settings.

Given a transition function, policy π *stochastically dominates (sd)* policy π' when for each rank ℓ, the expected score that alternatives with rankings as good as ℓ get selected under π, is greater or equal to the expected score under π'.

Definition 5.7. *Given a transition function T, matching policy π* ***stochastically dominates (sd)*** *π', if at all states $h^t \in \mathcal{H}^t$, for all agents $i \in N$,*

$$\forall o_\ell \in M,\ W_i^{\pi}(h^t, o_\ell) \geq W_i^{\pi'}(h^t, o_\ell)$$

Strategyproofness in sequential settings states that not only an agent cannot improve its current immediate outcome by misreporting its preferences but also the agent's current misreport will not make it better off in the future.

Definition 5.8. *A matching policy is* ***globally sd-strategyproof (gsd-strategyproof)*** *if and only if for any transition function T, given any misreport $\hat{\succ}_i^t = \sigma_i(\succ_i^t)$ such that $\hat{h}_i^t = (\succ^1, \ldots, (\hat{\succ}_i^t, \succ_{-i}^t); \mu^1, \ldots, \mu^{t-1})$ at time t, for all agents $i \in N$,*

$$\forall o_\ell \in M,\ W_i^{\pi}(h^t, o_\ell) \geq W_i^{\pi}(\hat{h}^t, o_\ell)$$

Global sd-strategyproofness is an incentive requirement which states that under any possible transition of preference profiles, no agent can improve its sequence of random matchings (now or in future) by a strategic report. We define a weaker notion for strategyproofness in sequential settings. Weak gsd-strategyproofness guarantees that an agent's misreport can never result in an outcome that stochastically dominates its sequence under truthful reporting. In other words, a misreport either results in a stochastically dominated outcome or in a sequence that is incomparable to the outcome under truthfulness.

Definition 5.9. *A matching policy is* ***weakly gsd-strategyproof*** *if and only if truthfulness does not yield to a stochastically dominated outcome, i.e., for any transition function T, given any misreport $\hat{\succ}_i^t = \sigma_i(\succ_i^t)$ such that $\hat{h}_i^t = (\succ^1, \ldots, (\hat{\succ}_i^t, \succ_{-i}^t); \mu^1, \ldots, \mu^{t-1})$ at time t, for all agents $i \in N$,*

$$\exists\ o_\ell \in M,\ W_i^{\pi}(h^t, o_\ell) > W_i^{\pi}(\hat{h}^t, o_\ell)$$

Clearly, a matching mechanism that satisfies global sd-strategyproofness is also weakly gsd-strategyproof, but the reverse does not hold.

If a policy is not gsd-strategyproof, an agent can benefit by misreporting its preferences. A matching policy is weakly manipulable if it is not gsd-strategyproof.

Definition 5.10. *A matching policy is **weakly manipulable** if it is not gsd-strategyproof.*

Intuitively, weak manipulation states that there may exist some utility model consistent with an agent's preference ordering, under which misreporting would improve the agent's expected utilities. Similarly, a matching policy is strongly manipulable if it is not weakly gsd-strategyproof.

Definition 5.11. *A matching policy π is **strongly manipulable** if it is not weakly gsd-strategyproof.*

In other words, given any (possibly time-invariant) utility model consistent with the agent's dynamic ordinal preferences, agent i is better off deviating from truthful reporting.

5.4.3 Pareto Improvement in Expectation

Recall that we denote a sequence of preference profiles (states) and matchings (actions) up to time t by h^t, that is, $h^t = (\succ^1, \mu^1, \ldots, \succ^t)$. A sequence of matchings prescribed by policy π dominates another sequence prescribed by π' if there exists one agent i that strictly prefers π_i to π'_i and all other agents are weakly better off under π. A sequence of matchings prescribed by a policy π can be determined recursively according to the following: given a preference profile at time t, policy π determines a random matching $\bar{\mu}^t \sim \pi(h^t)$, and after realization of $\bar{\mu}^t$ the system transitions to the next state $h^{t+1} \sim T(\cdot|h^t, \mu^t)$.

Given a sequence of matching decisions, a change in the decision trajectory that makes at least one agent better off without making any other individual worse off is called a **Pareto improvement**. Clearly, a sequence of allocations is Pareto efficient when no further Pareto improvement can be made.

Definition 5.12. *Given a transition function T, a sequence prescribed by policy π is a **Pareto improvement** over a sequence prescribed by π' in expectation when the following two conditions hold at time t:*

- *There exists an agent $i \in N$ such that, $\exists \ell$ we have $W_i^{\pi}(h^t, o_\ell) > W_i^{\pi'}(h^t, o_\ell)$ and for all other ranking positions $\ell', W_i^{\pi}(h^t, o_{\ell'}) \geq W_i^{\pi'}(h^t, o_{\ell'})$, and;*

- *For all other agents $j \in N \setminus \{i\}$, for all ℓ, $W_j^{\pi}(h^t, o_\ell) \geq W_j^{\pi'}(h^t, o_\ell)$.*

With fixed preferences, any sequence of ex post efficient matchings (see Definition 5.6) belongs to the set of Pareto efficient sequential allocations since no improvement is possible in each local matching and any improvement over the sequence of matchings will require to make at least one agent worse off in some local matchings. Under dynamic preferences, given a transition model a Pareto improvement may be possible by changing the trajectory of decisions to less conflicting states, where some agents receive equally good sequences of outcomes and at least one agent's allocation is improved. Note that the converse does not hold since an optimal policy may sometimes yield a locally dominated allocation to maximize the quality of expected allocations.

5.5 Sequential RSD Policy

A typical goal of a mechanism designer is to allocate the objects to agents efficiently and fairly. However, ex ante strategyproofness and ex ante efficiency are incompatible in one-shot settings [30]. This incompatibility persists in dynamic domains, and several other efficiency notions under ordinal preferences (*e.g.* rank-maximal matching [83]) have also been shown to be highly manipulable. Thus, a social planner needs to carefully decide which one of the two properties he or she deems more crucial. If the goal is to satisfy the latter, randomizing over all possible serial dictatorships can guarantee strategyproofness while satisfying some notion of fairness (equal treatment of agents with equal preferences). Random Serial Dictatorship (RSD) [6] is a widely adopted mechanism that satisfies lsd-strategyproofness while being extremely easy to implement. In this section, we focus our attention on the RSD mechanism and ask the question of *whether RSD remains strategyproof in sequential matchings with possibly dynamic preferences.*

To formally define the RSD mechanism, we first introduce priority orderings and serial dictatorships. A *priority ordering* $f : \{1, \ldots, n\} \to N$ is a one-to-one mapping that specifies an ordering of agents: agent $f(1)$ is ordered first, agent $f(2)$ is ordered second, and so on.

Given a priority ordering $f \in \mathcal{F}$ and a preference profile $\succ$, a **Serial Dictatorship**, $SD(f, \succ)$, is as follows: agent $f(1)$ receives its favorite object $m_1 \in M$ according to its preference $\succ_{f(1)}$; agent $f(2)$ receives its favorite object $m_2 \in M \setminus \{m_1\}$; $f(n)$ receives its best object $m_n \in M \setminus \{m_1, \ldots, m_{n-1}\}$.

Random serial dictatorship is a *convex combination* of all feasible serial dictatorships, induced by a uniform distribution over all priority orderings, and it is formally defined as

$$RSD(\succ) = \frac{1}{n!} \sum_{f \in \mathcal{F}} SD(f, \succ). \tag{5.4}$$

In one-shot settings, RSD is sd-strategyproof, equitable (in terms of equal treatment of equals), and ex post efficient [6]. Although the RSD mechanism does not satisfy a stronger notion of sd-efficiency, it is weakly envyfree for any number of agents and objects [30].

5.6 One-Shot Direct Mechanisms

In dynamic mechanism design global strategyproofness is a strong requirement, which requires agents to be truthful under all possible stochastic events and dynamics. One plausible way to achieve global strategyproofness in dynamic settings is assigning probability distributions uniformly over all possible sequences of priority orderings, thus treating the problem as a single-shot matching problem. The idea is to set all possible priority orderings at the initial start state and for any consecutive states up to a planning horizon (see Figure 5.6 for an example with 3 time steps). In this case, given λ decision steps, a mechanism designer must sample from a uniform distribution over $(n!)^\lambda$ possible sequences of priority orderings.

The mechanism can be thought of as a generalization to the RSD mechanism when all future decisions are made in the start state; instead of randomizing over priority orderings, this mechanism randomizes over all sequences of priority orderings.

Although this mechanism satisfies global strategyproofness, it imposes a high cognitive burden on the agents by trying to elicit agents' preferences for all the future periods. Such preference elicitation requires that each agent reasons over its preferences for every future step while considering uncertainties over idiosyncratic preferences, which could be computationally expensive. Moreover, randomizing over all sequences of priority orderings may result in extremely unfair allocations ex post; for example, choosing a sequence of priority orderings in which agent i is ranked before agent j in every period.

5.6.1 Incentives in Sequential RSD

Due to the shortcomings discussed in the previous section, we focus attention on direct sequential policies where agents report their preferences in each step and a matching deci-

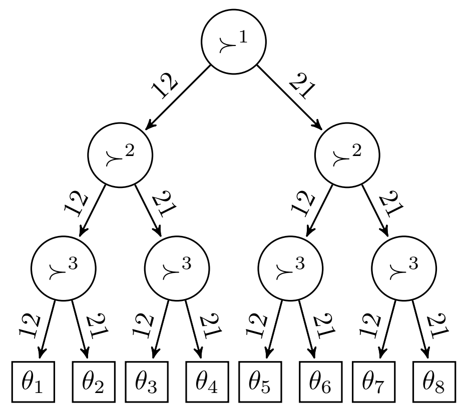


Figure 5.6: Possible priority ordering sequences for three-time periods. The labels on the edges represent a priority ordering, i.e., 21 denotes an ordering f with $f(1) = 2$, and $f(2) = 1$.

sion is made according to the reported preferences (as illustrated in Figure 5.4). A natural matching policy is to elicit agents' preferences and apply the RSD mechanism in each period to obtain a random allocation.

Definition 5.13. *Sequential RSD is a stochastic policy consisting of a sequence of RSD-induced matchings in each period.*

While RSD satisfies lsd-strategyproofness at each decision period, we show that *sequential RSD* is prone to manipulation when agents have dynamic preferences. Sequential RSD selects a random matching at each round independent of the past history of decisions and preferences.

Theorem 5.1. *Sequential RSD (a sequence of RSD induced matchings) is not gsd-strategyproof.*

(a) Truthful		(b) Misreport	
$\succ_1$:	$a \succ c \succ b$	$\succ_1$:	$a \succ c \succ b$
$\succ_2$:	$b \succ c \succ a$	$\succ_2$:	$b \succ c \succ a$
$\succ_3$:	$a \succ c \succ b$	$\hat{\succ}_3$:	$a \succ b \succ c$

Table 5.4: Preferences revealed by three agents.

Proof. Consider 2 decision periods with deterministic preference dynamics known to agents. Let $\bar{\mu}^2_{\mu_1}$ denote the random matching at $t = 2$ after assignment μ_1 at $t = 1$. With truthful preferences (Table 5.4a), RSD induces the following random matching: $(\frac{1}{2}\mu_1, \frac{1}{2}\mu_2) =$

	$W_3(\succ^t, o_\ell)$	$W_3((\hat{\succ}^t_i, \succ^t_{-i}), o_\ell)$
o_1	$\frac{1}{2} + [\frac{1}{2}(\bar{\mu}^2_{\mu_1}) + \frac{1}{2}(\bar{\mu}^2_{\mu_2})]$	$\frac{1}{2} + [\frac{3}{6}(\bar{\mu}^2_{\mu_1}) + \frac{2}{6}(\bar{\mu}^2_{\mu_2}) + \frac{1}{6}(\bar{\mu}^2_{\mu_3})]$
o_2	$1 + [\frac{1}{2}(\bar{\mu}^2_{\mu_1}) + \frac{1}{2}(\bar{\mu}^2_{\mu_2})]$	$\frac{5}{6} + [\frac{3}{6}(\bar{\mu}^2_{\mu_1}) + \frac{2}{6}(\bar{\mu}^2_{\mu_2}) + \frac{1}{6}(\bar{\mu}^2_{\mu_3})]$
o_3	$1 + [\frac{1}{2}(\bar{\mu}^2_{\mu_1}) + \frac{1}{2}(\bar{\mu}^2_{\mu_2})]$	$1 + [\frac{3}{6}(\bar{\mu}^2_{\mu_1}) + \frac{2}{6}(\bar{\mu}^2_{\mu_2}) + \frac{1}{6}(\bar{\mu}^2_{\mu_3})]$

Table 5.5: Evaluation of the matching policy for agent 3.

$(\frac{1}{2}abc, \frac{1}{2}cba)$. If agent 3 misreports (as shown in Table 5.4b), the probability distribution would be $(\frac{2}{6}\mu_1, \frac{3}{6}\mu_2, \frac{1}{6}\mu_3) = (\frac{2}{6}abc, \frac{3}{6}cba, \frac{1}{6}acb)$. Assuming truthfulness in the second period, given a preference profile, identical decisions always result in identical next states. For each ranking position ℓ, we compute $W_3(\cdot, o_\ell)$ as shown in Table 5.5.

To prove that sequential RSD is not gsd-strategyproof, we need to show that there exists a ranking position ℓ such that $W_3((\hat{\succ}^t_3, \succ^t_{-3}), o_\ell) > W_3(\succ^t, o_\ell)$. For o_1 we need to show that $\frac{1}{2} + [\frac{3}{6}(\bar{\mu}^2_{\mu_1}) + \frac{2}{6}(\bar{\mu}^2_{\mu_2}) + \frac{1}{6}(\bar{\mu}^2_{\mu_3})] > \frac{1}{2} + [\frac{1}{2}(\bar{\mu}^2_{\mu_1}) + \frac{1}{2}(\bar{\mu}^2_{\mu_2})]$, which is simplified to $\bar{\mu}^2_{\mu_3} > \bar{\mu}^2_{\mu_2}$. Therefore, for any matching such that $\bar{\mu}^2_{\mu_3}(3)$ is preferred to $\bar{\mu}^2_{\mu_2}(3)$, agent 3 would be better off under misreporting. For example, consider a matching that allocates to agent 3 its top choice with certainty under $\bar{\mu}^2_{\mu_3}(3) = (1o_1)$, but assigns all objects with equal probability under $\bar{\mu}^2_{\mu_2}(3) = (\frac{1}{3}o_1, \frac{1}{3}o_2, \frac{1}{3}o_3)$.

For the sake of completeness, we also consider objects that are ranked as good as second. For o_2, we must show that $\frac{5}{6} + [\frac{3}{6}(\bar{\mu}^2_{\mu_1}) + \frac{2}{6}(\bar{\mu}^2_{\mu_2}) + \frac{1}{6}(\bar{\mu}^2_{\mu_3})] > 1 + [\frac{1}{2}(\bar{\mu}^2_{\mu_1}) + \frac{1}{2}(\bar{\mu}^2_{\mu_2})]$, which further simplifies to $\bar{\mu}^2_{\mu_3} > \bar{\mu}^2_{\mu_2} + 1$. Since any allocation at best provides a matching with the sum of 1 for o_1 and o_2, we have that $\bar{\mu}^2_{\mu_3} \leq 1$. Thus, the assignment under truthfulness is always preferred for objects that are as good as o_2.

Even though for o_2 misreporting is not beneficial, it is still possible to benefit from changing the allocation for the top ranking position. For o_1 we have $W_3((\hat{\succ}^t_3, \succ^t_{-3}), o_1) > W_3(\succ^t, o_1)$, implying that sequential RSD is not gsd-strategyproof. This concludes our proof. □

The following example illustrates a simple scenario wherein sequential RSD can be manipulated.

Example 5.7. Consider two agents and two objects. Assume the preferences of agents are as follows: agent 1's preference is fixed as $\succ_1$: $a \succ b$ in all periods, while agent 2's preference is $\succ_2$: $a \succ b$ in the initial state and remains the same if it receives object b. However, agent 2's preference will change to $b \succ a$ if it receives object a and will remain fixed thereafter, no matter what it receives in the future.

In this case, agent 1 may misreport its preference in the first period as $b \succ a$ so that agent 2 obtains a with certainty in the first period and agent 2's preference evolve to $b \succ a$ in all future periods. This way, agent 1 can avoid competition in the future periods and benefit from its initial misreport.

In the next section, we will elaborate on possible types of manipulation and how strategyproofness also depends on whether the transitions are Markovian or history dependent.

5.7 Strong and Weak Manipulation of Sequential RSD

So far we showed that manipulation is possible under sequential RSD. However, we did not show the quality of manipulation (whether weak or strong) nor discussed the possibility of manipulation with regards to the planning horizon. Recall that a matching policy is weakly manipulable if it is not gsd-strategyproof, and it is strongly manipulable if it does not satisfy weak gsd-strategyproofness. This distinction between strong and weak manipulation is an important factor when measuring the susceptibility of matching policies. Strong manipulation states that an agent can strategically influence the matching decisions such that its allocation improves, for all possible utility models that are consistent with its preference orderings. On the other hand, under weak manipulation an agent may benefit from a strategic report only for some (but not *all*) particular utility models.

In this section, we are interested in characterizing the conditions wherein sequential RSD is susceptible to strong manipulation. We dig deeper on the possibility of weak or strong manipulation under various transition dynamics and provide results on the number of future steps required for a strategic agent to benefit from misreporting, looking forward into the future. We first introduce a few lemmas that we will use later to analyze manipulation strategies.

Lemma 5.1. *Let $\bar{\mu}$ be a random matching decision, and $n = |N|$. Then, the probability of assigning an object to an agent under RSD satisfies the following lower bound and upper bound:*

- *Worst-case matching outcome: for all ranking position ℓ, $\sum_{x=o_1}^{o_\ell} p_i^t(x|h^t) = \frac{\ell}{n}$. In other words, for all ℓ, $p_i^t(o_\ell|h^t) \geq \frac{1}{n}$.*
- *Best-case matching decision: for top object o_1, $p_i^t(o_1|h^t) = 1$.*

Proof. Worst-case matching outcome under RSD is when all n participating agents have identical preferences. In this case, at least in $\frac{(n-1)!}{n!}$ orderings, an agent will be ordered first and has exactly $\frac{1}{n}$ chance of receiving its top choice. Since agents have identical preferences, then there is exactly $\frac{1}{n}$ probability of being ordered first, second, third, and so on, implying that for any ranking position $p_i^t(o_\ell|h^t) \geq \frac{1}{n}$. Adding these probabilities for all ranking positions we have $\sum_{x=o_1}^{o_\ell} p_i^t(x|h^t) = \frac{\ell}{n}$.

Best-case matching outcome is when no two agents have identical top preferences, that is, each object is a top choice of only one agent. In this case, for all $n!$ priority orderings of agents, all agents receive their top object, and thus $p_i^t(o_1|h^t) = 1$. □

The following lemma states that if a misreport does not change the current matching decision, it cannot impact the evolution of preferences and the next state, and thus the decision trajectory, remains unchanged.

Lemma 5.2. *Fixing $\mu \in \mathcal{M}$, for any $\succ^{t+1} \in \mathcal{P}^n$ given agent i's misreport $\hat{\succ}_i^t$,*

$$T(\succ^{t+1} \mid \succ^t, \mu) = T(\succ^{t+1} \mid (\hat{\succ}_i^t, \succ_{-i}^t), \mu)$$

Proof. Since agent i's true transition is according to its true preference and not the misreport, we have $T_i(\succ_i^{t+1} \mid \hat{\succ}_i^t, \mu(i)) = T_i(\succ_i^{t+1} \mid \succ_i^t, \mu(i))$. By the independence of priors and transitions and the assumption of no externalities, given a matching decision μ, for all other agents $j \in N \backslash i$ we can write,

$$T_j(\succ_j^{t+1} \mid (\hat{\succ}_i^t, \succ_{-i}^t), \mu(j)) = T_j(\succ_j^{t+1} \mid \succ_j^t, \mu(j)).$$

The overall transition is the Cartesian product of all individual transitions, and thus we have,

$$\prod_{j \in N} T_j(\succ_j^{t+1} \mid (\hat{\succ}_i^t, \succ_{-i}^t), \mu(j)) = \prod_{j \in N} T_j(\succ_j^{t+1} \mid \succ_j^t, \mu(j)),$$

which implies that $T(\succ^{t+1} \mid \succ^t, \mu) = T(\succ^{t+1} \mid (\hat{\succ}_i^t, \succ_{-i}^t), \mu)$. □

In other words, given the independence of priors and transitions, agent i's misreport only affects the decision trajectory through changing the allocation. The immediate consequence of Lemma 5.2 for random matchings is the following: If a misreport does not change the random outcome at t, it does not impact the evolution of preferences, and hence, the decision trajectory. Thus, a misreport may change the decision trajectory if and only if it changes the random allocation for some (or all) other agents.

In each step, sequential RSD prescribes a matching decision based on current reported preferences and independent of the past history. However, agents' preferences may evolve based on the history of past allocations and preferences. Therefore, in the sections that follow we distinguish between history-dependent and Markovian transitions when analyzing agents' strategic behavior.

5.7.1 Two Agents

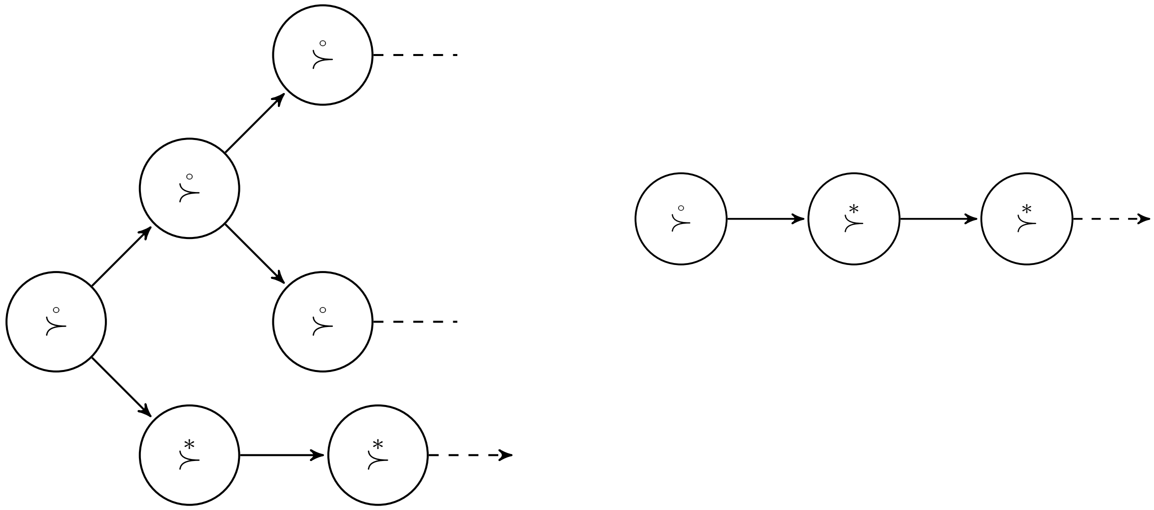
Several decision problems deal with periodically assigning tasks or responsibilities between two agents with dynamic preferences. For example, co-funders of a small startup may want to re-assign executive tasks among each other biannually. Similarly, co-advisors may decide to assign supervision activities periodically based on their (dynamic) preferences over the tasks.

We first study the strategyproofness of sequential RSD when allocating objects to two agents. We show that gsd-strategyproofness depends on the planning horizon and the transition of preferences. Our results show that the sequential RSD policy is not even weakly gsd-strategyproof when the transitions are history dependent.

Theorem 5.2. *Sequential RSD is strongly manipulable for two agents when $\lambda \geq 4$, under history-dependent transitions. For $\lambda < 4$ sequential RSD is gsd-strategyproof under history-dependent transitions.*

Proof. We provide a proof by construction. With two agents and two objects, there are four distinct preference profiles: two profiles where both agents have the same top choice and two where the top choices differ. For agent i, we construct the worst-case scenario under truthful reporting and the best-case scenario under a strategic misreport as follows:

- Let $\overset{*}{\succ}$ denote a preference profile where agent i receives its top choice. By Lemma 5.1, the best-case outcome at each period is $\pi_i(\overset{*}{\succ}) = (1o_1)$, which for simplicity we denote this outcome by $\overset{*}{\mu}$.
- Let $\overset{\circ}{\succ}$ be a profile with the worst possible outcome for agent i. By Lemma 5.1, the worst-case outcome for agent i at each period is $\pi_i(\overset{\circ}{\succ}) = (1/2o_1, 1/2o_2)$, which we denote by $\overset{\circ}{\mu}$.



(a) Worst-case sequence

(b) Best-case sequence

Figure 5.7: History-dependent transitions for two agents.

By Lemma 5.2, agent i can only affect the evolution of preferences by changing the random outcome prescribed by π. At current time t, given history-dependent transitions, agent i reports preference $\hat{\succ}_i^t$ to change the matching outcome to $\pi(\hat{\succ}_i^t, \succ_{-i}^t)$.

Considering the best trajectory, $\pi(\hat{\succ}_i^t, \succ_{-i}^t)$ will transition the system to a trajectory with a sequence of best outcomes as shown in Figure 5.7b. Agent i's misreport would cause an immediate loss at t; however, it may take its state to $\overset{*}{\succ}$ according to $T(h^{t'} = \overset{*}{\succ} | h^{t'-1}, \mu^{t'-1}) = 1$ for all $t' > t$ where $h^{t'} = (\mathring{\succ}, \overset{*}{\mu}, \overset{*}{\succ}, \overset{*}{\mu}, \ldots, \succ^{t'})$ and $\succ^{t'} = \overset{*}{\succ}$. However, under truthful reporting for the worst-case trajectory (as shown in Figure 5.7a) we have $T(h^{t'} = | h^{t'-1}, \mu^{t'-1}) = 1$ for all $t' > t$, where $h^{t'} = (\mathring{\succ}, \mathring{\mu}, \mathring{\succ}, \ldots, \succ^{t'})$ and $\succ^{t'} = \mathring{\succ}$.

We can now compute the score of receiving objects in each ranking position for both best-case and worst-case scenarios. First consider the top ranked object for $\lambda = 4$ steps. For the best-case scenario under misreport, at the current period agent i must report a preference $\hat{\succ}_i^t \neq \succ_i^t$, because there are only two possible preference orderings with two objects. By lsd-strategyproofness agent i receives a less preferred outcome at t but, given the best-outcome scenario, in all future steps receives its top choice. Thus, by backward induction we write $W_i^\pi(\hat{\succ}^t, o_1) = 3$, since for at least 3 steps it receives its top choice. Similarly, for the worst-case scenario under truthful reporting, by backward induction we can write $W_i^\pi(\succ^t, o_1) = 2 + \frac{3}{4}$. Therefore, $W_i^\pi(\hat{\succ}^t, o_1) > W_i^\pi(\succ^t, o_1)$. Since there are only two objects, the allocation of the second ranking position is the complement of the first ranking position, implying that for $\lambda \geq 4$ sequential RSD is strongly manipulable.

We now show that for $\lambda < 4$, sequential RSD is gsd-strategyproof for two agents. According to Equation 5.3 for the best-case scenario under the strategic report when $\lambda = 3$ we can write:

$$\begin{aligned} W_i^\pi(\succ^t) =& \frac{1}{4}(\frac{1}{2}(o_1 + o_2)) + \frac{1}{4}(\frac{1}{2}(o_1 + o_2)) + \\ & \frac{1}{2}(\frac{1}{2}(o_1 + o_2)) + \frac{1}{2}(2o_1) + \frac{1}{2}(o_1 + o_2) \end{aligned}$$

Thus, the score of receiving top choice under truthful reporting is $W_i^\pi(\succ^t, o_1) = 2$ which is equal to that of misreport $W_i^\pi(\hat{\succ}^t, o_1) = 2$. Thus, the sequence of matching under truthfulness stochastically dominates the best possible allocation sequence under a misreport. □

Theorem 5.3. *Given Markovian transitions, sequential RSD is gsd-strategyproof for two agents.*

Proof. Proof by construction. We consider two scenarios for agent i: the best trajectory after misreporting at time t, $\pi(\hat{\succ}_i^t, \succ_{-i}^t)$ which will transition the system to a sequence of best outcomes similar to the sequence shown in Figure 5.7b, and the worst-case scenario under a truthful report at time t. By Lemma 5.1, the worst-case outcome assigns 1/2 probability of getting top object o_1, and 1/2 of getting o_2. By lsd-strategyproofness at t, agent i strictly prefers the immediate outcome under truthfulness, that is $\pi_i(\succ^t) \succ_i^t \pi_i(\hat{\succ}_i^t, \succ_{-i}^t)$.

To construct the best-case scenario after misreporting, we assume the following: (i) given any non-truthful report and any matching the preference profile transitions to a profile that agent i receives its top choice, that is, $T(\overset{*}{\succ}|\overset{\circ}{\succ}, \mu) = 1$ for any $\overset{\circ}{\succ} \neq \succ$ and μ, and (ii) $T(\overset{*}{\succ}|\overset{*}{\succ}, \mu) = 1$ for any profile that assigns top choices to agents.

By the Markovian assumption, in any state $\succ^{t'} = \overset{\circ}{\succ}$ because $\overset{\circ}{\succ}$ is a profile with the worst-outcome, there is 1/2 probability of choosing decision μ and slipping into the best trajectory as shown in Figure 5.8, and 1/2 probability of making a decision that transitions agents to profile $\overset{\circ}{\succ}$. Thus, under truthfulness, for two agents the score (expected sum of probabilities) can be written in the following closed form

$$W_i^\pi(\succ^t) = \sum_{k=1}^{\lambda} [(\frac{1}{2})^k (\lambda - k) o_1 + (\frac{1}{2})^k (o_1 + o_2)]$$

To show that sequential RSD is gsd-strategyproof when $n = 2$, for any λ, we verify that $W_i(\succ^t, o_1) \geq W_i(\hat{\succ}^t, o_1)$ by induction. For $\lambda = 1$ the base case holds by lsd-strategyproofness of sequential RSD. For the induction step, assume that for λ we have,

$$W_i(\succ^t, o_1) \geq W_i(\hat{\succ}^t, o_1) \tag{5.5}$$

$$\sum_{k=1}^{\lambda}[(\frac{1}{2})^k(\lambda - k) + (\frac{1}{2})^k] \geq \lambda - 1$$

Then we need to show that the above holds for $\lambda + 1$. Since the hypothesis holds for any arbitrary λ, we rewrite it for $\lambda + 1$:

$$\sum_{k=1}^{\lambda+1}[(\frac{1}{2})^k((\lambda+1) - k) + (\frac{1}{2})^k] \geq (\lambda+1) - 1$$

$$\underbrace{\sum_{k=1}^{\lambda}[(\frac{1}{2})^k(\lambda - k) + (\frac{1}{2})^k]}_{\text{hypothesis}} + \sum_{k=1}^{\lambda}(\frac{1}{2})^k + \frac{1}{2} \geq (\lambda+1) - 1$$

The hypothesis statement is greater than or equal to $\lambda - 1$, we need to show that the rest of the left-hand side is greater than or equal 1:

$$\sum_{k=1}^{\lambda}(\frac{1}{2})^k + \frac{1}{2} \geq 1$$

which holds for any $\lambda \geq 1$. Thus, the inequality 5.5 holds for all λ, which concludes our proof by induction. The sum of probability assignments for second ranking position complements that of the first ranking position, which implies that sequential RSD is gsd-strategyproof under Markovian transitions. □

5.7.2 Three or More Agents

In this section, we investigate the weak and strong manipulability of sequential RSD for settings with three agents, and discuss how these results hold for larger problems with more

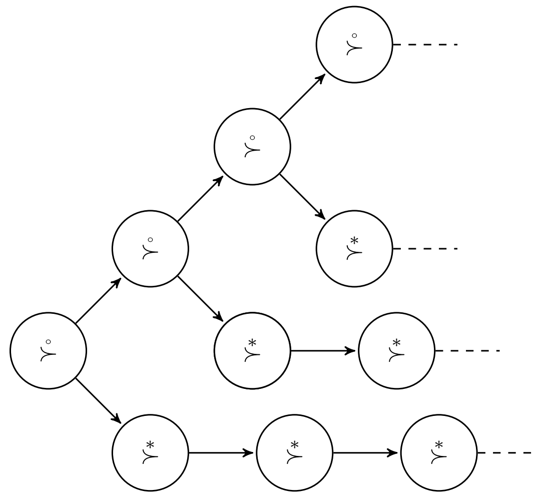


Figure 5.8: Worst-case scenario with Markovian transitions.

agents. We establish a connection between the number of decision steps (planning horizon) and the possibility of manipulation in such matching mechanisms. We show findings for three agents and then argue that our impossibility results persist for any larger number of agents.

History-Dependent Transitions

Theorem 5.4. *Let the transition functions be history-dependent. For $n = 3$, sequential RSD is strongly manipulable if $\lambda \geq 4$, and it is weakly manipulable if $1 < \lambda < 4$.*

The proof closely follows the proof of Theorem 5.2 by contrasting the worst-case and best-case scenarios after truthful and strategic reports respectively, and is relegated to Appendix B.

Markovian Transitions

When the matching policy is independent of history, Theorem 5.4 shows that strong manipulation is possible under history-dependent transitions. In this section, we focus attention on Markovian transitions as a common practice in the literature of sequential decision making under uncertainty [25, 31, 32, 42, 115].

The following lemma states that if an agent's strategic report changes the probability distribution over deterministic matchings, then for at least one deterministic matching decision, the matching remains in the support of the new random matching.

Lemma 5.3. *Let* $\succ_i^t$ *and* $\hat{\succ}_i^t$ *denote agent* i*'s true and strategic report respectively. There always exists a subset of Pareto efficient matchings* $\mathcal{M}' \subseteq \mathcal{M}$ *with* $\sum_{\mu \in \mathcal{M}'} \pi(\mu|\hat{\succ}^t) > 0$ *such that* $\sum_{\mu \in \mathcal{M}'} \pi(\mu| \succ^t) \geq \frac{1}{n}$.

Proof. The matching decision is a uniform randomization over serial dictatorships that are prescribed by the set of all possible priority orderings (permutations of agent orderings). For all serial dictatorships where agent i has the lowest priority, it chooses an object after all agents receive their objects. Thus, agent i will receive what is left over no matter what preference ordering it declares, implying that a subset of matchings $\mathcal{M}' \subseteq \mathcal{M}$ exists with $\sum_{\mu \in \mathcal{M}'} \pi(\mu|\hat{\succ}^t) > 0$.

Now we need to show that the probability of this subset of matching decisions is always greater than or equal to $1/n$. Sequential RSD is a uniform distribution over serial dictatorships prescribed by $n!$ possible priority orderings. Fixing the preferences of all agent $j \in N \setminus i$, there are $(n-1)!$ priority orderings where all $n-1$ agents are ordered before agent i. Thus, for $(n-1)!$ orderings $\pi_j(\mu|\hat{\succ}^t) = \pi_j(\mu| \succ^t)$. This implies that there exists a subset $\mathcal{M}' \subseteq \mathcal{M}$ in the support of both random matchings with the probability of at least $\frac{(n-1)!}{n!}$. □

Theorem 5.5. *Let transition functions be Markovian. Then sequential RSD is weakly gsd-strategyproof for* $1 < \lambda < 4$ *and is strongly manipulable for* $\lambda \geq 4$.

The proof is relegated to Appendix C. It uses Lemma 5.3 and follows closely the proof of Theorem 5.3.

The results on weak and strong gsd-strategyproofness of sequential RSD when there are three agents are essentially negative results that show the lack of truthful incentives when using sequential RSD as the matching policy. Therefore, these results can easily be generalized to larger problems with more agents $n \geq 3$. The following example shows how these findings can be extended to problems with larger number of agents. Consider a group of agents, three of which are competing over objects and the rest do not impact the sequential outcome for those three and among themselves. Theorem 5.5 and 5.4 show that under sequential RSD, there will still be one agent (in the group of three) that can strongly manipulate the outcome and gain from misreporting its preferences.

Table 5.6 summarizes our results on manipulability of sequential RSD according to Markovian or history-dependent transitions. Sequential RSD is gsd-strategyproof for two

	Markovian	**History-Dependent**
2 agents	gsd-strategyproof	gsd-strategyproof if $\lambda < 4$ not weakly gsd-strategyproof if $\lambda \geq 4$
3 agents (or more)	weakly gsd-strategyproof if $\lambda < 4$ not weakly gsd-strategyproof if $\lambda \geq 4$	weakly gsd-strategyproof if $\lambda < 4$ not weakly gsd-strategyproof if $\lambda \geq 4$

Table 5.6: Summary of manipulation results under Markovian and history-dependent transitions.

agents under Markovian transitions. For three or more agents, sequential RSD does not even satisfy weak gsd-strategyproofness given either Markovian or history-dependent transitions, and therefore, it is strongly manipulable.

5.8 Sequential RSD with Adjusted Priorities

In the previous sections, we showed that sequential RSD is prone to manipulation when agents have dynamic preferences. In this section, we introduce a modification of RSD, which uses information contained in histories of interactions between agents to overcome strategic behavior. We start this section with some observations about relationships between agents.

A key property of RSD is that it prioritizes agents in each round, and an agent with higher priority gets to choose its more preferred item from the set of *remaining objects*. This gives rise to the concept of dictatorial domination.

Definition 5.14. *Given a preference profile and the realization of a matching decision, we say that agent i* ***dictatorially dominates*** *agent j at time t if and only if $\mu^t(i) \succ_j^t \mu^t(j)$.*

We can represent each agent's dictatorial dominance as a binary relation between every pair of agents. Each agent's dictatorial dominance on other agents at time t is represented by $\boldsymbol{\omega}_i^t = (\omega_{i,1}^t, \ldots, \omega_{i,n}^t)$, where $\omega_{i,j}^t = 1$ denotes that agent i has dominated agent j at time t. A *dominance profile* is a matrix of agents' dictatorial dominances, $\boldsymbol{\omega}^t = (\boldsymbol{\omega}_i^t, \ldots, \boldsymbol{\omega}_n^t)$.

Given a random matching mechanism, the probability that agent i dominates agent j is equal to the sum of the probabilities of all deterministic matchings wherein agent j prefers the outcome of agent i to its own outcome.[6] Given a random matching policy π and h^t,

[6] Deterministic matching is a special case where given μ^t, $\omega_{i,j}^t = 1$ iff $\mu^t(i) \succ_j^t \mu^t(j)$.

the probability of i dominating j at period t is:

$$\omega_{ij}^t(\pi(h^t)) = \sum_{\mu \in \mathcal{M}:\mu(i) \succ_j^t \mu(j)} \pi(\mu \mid h^t) \tag{5.6}$$

Similarly the probability that agent j prefers its own outcome to agent i's outcome is

$$\bar{\omega}_{ij}^t(\pi(h^t)) = 1 - \omega_{ij}^t(\pi(h^t)) \tag{5.7}$$

RSD ensures equal chance of dictatorships to agents, thus, an agent's strategic misreport can only increase its random dictatorial dominance on another agent to $\frac{1}{2}$. This is because agent i, knowing the preferences of others, can change the allocation of another agent j only in those permutations where i is prioritized before j. For all other priority orderings, agent i is ordered after agent j and cannot change agent j's allocation and increase its dominance on agent j.

Proposition 5.1. *Given RSD, for any $\succ^t \in \mathcal{P}^n$ we have $\forall i, j \in N, \omega_{ij}^t(\pi(\succ^t)) \leq \frac{1}{2}$, that is, the probability of agent i dominating another agent j is always bounded.*

The proof is relegated to the Appendix D.

The intuition comes from the fact that RSD is a uniform distribution over all priority orderings. Thus, when two agents have conflicting preferences over some alternatives, RSD assigns equal probabilities to all orderings that prioritize one agent lower than the other one.

5.8.1 Adaptive RSD

In this section, we provide a modification to the RSD mechanism based on adjusting agents' priority ordering. We introduce a simple structure to preserve the history of dominations throughout the matching process. Formally, let $\boldsymbol{d}^{t-1}$ be a matrix representing the complete *dominance history* of agents up to time t, where $d_{i,j}^{t-1}$ is agent i's history of dominating agent j up to time t. Since we only require to know one-to-one dominance history of agents, and because we update the dominance history in each step using *exclusive disjunction*, we can summarize the dominance history in a single global variable. As the dominance profile of agents is attained after realization of decisions (ex post), the state of the mechanism at t can be summarized by $h^t = (\succ^t, \boldsymbol{d}^{t-1})$. As shown in Algorithm 3, Adaptive RSD runs as follows:

Algorithm 3: RSD with adjusted priorities (Adaptive RSD)

Input: preference profile $\succ^t$, dominance history $\boldsymbol{d}^{t-1}$

Output: A probability distribution prescribed by π

foreach *priority ordering* $f \in \mathcal{F}$ **do**
- $f' \leftarrow f$;
- **for** $(i = n$ **to** $1)$ **do**
  - **for** $(j = n$ **to** $1)$ **do**
    - **if** $(j < i$ **and** $d^{t-1}_{f'(j),f(i)} = 1)$ **then**
      - $f' \leftarrow$ `Shift`$(f'(j), i - j)$;
- **foreach** $(\mu \in \mathcal{M})$ **do**
  - **if** $SD(f', \succ^t) = \mu$ **then**
    - $\pi(\mu \mid (\succ^t, \boldsymbol{d}^{t-1})) \leftarrow \pi(\mu \mid (\succ^t, \boldsymbol{d}^{t-1})) + \frac{1}{n!}$;

`// Updating` $\boldsymbol{d}^t$ `after realization of` π

for $(i = 1$ **to** $n)$ **do**
- **for** $(j = 1$ **to** $n)$ **do**
  - $d^t_{ij} = \omega^t_{ij}(\mu^t) \oplus d^{t-1}_{ji}$;

- If for a pair of agents the dominance history is unbalanced, that is, $d^{t-1}_{ij} = 1$, then for each priority ordering $f \in \mathcal{F}$, if agent j is ranked before i, we change the priority ordering to f' by shifting the orderings such that j has higher priority than agent i, and add the new priority ordering to a new multiset $\mathcal{F}'$.

- The matching mechanism then draws a priority ordering from a uniform probability distribution over the priority orderings in the multiset $\mathcal{F}'$. Then, agents select alternatives according to the selected ordering. Note that the multiset $\mathcal{F}'$ may contain several copies of a single priority ordering, and the matching mechanism draws uniformly from all $n!$ members of this set, thus, the matching may no longer be uniform over the set of all possible orderings $\mathcal{F}$.

- After the realization of matching decision, and given the dominance profile $\boldsymbol{\omega}^t$ at time t, the mechanism updates the dominance history according to the following exclusive disjunction: $\boldsymbol{d}^t = \boldsymbol{\omega}^t \oplus (\boldsymbol{d}^{t-1})^{Tr}$, where $(\boldsymbol{d}^{t-1})^{Tr}$ is the transpose matrix of $\boldsymbol{d}^{t-1}$, *i.e.*, for each element d^{t-1}_{ij} the transpose element would be d^{t-1}_{ji}.

$$\begin{pmatrix} \succ_1: a \succ b \succ c \\ \succ_2: b \succ a \succ c \\ \succ_3: a \succ b \succ c \end{pmatrix}$$

Table 5.7: A sample preference profile.

RSD	Adaptive RSD
$1, 2, 3$	$1, 3, 2$
$1, 3, 2$	$1, 3, 2$
$2, 1, 3$	$1, 3, 2$
$2, 3, 1$	$3, 2, 1$
$3, 1, 2$	$3, 1, 2$
$3, 2, 1$	$3, 2, 1$

Table 5.8: Priority orderings under RSD and Adaptive RSD given the dominance history described in the example.

The `Shift` operation updates the agent ordering by shifting the dominating agent to a lower ranking immediately after the dominated agent. Consider a problem with 4 agents with priority ordering of $4, 1, 3, 2$; agent 1 is ordered second and agent 2 is ordered last (fourth). If $d_{12}^{t-1} = 1$, then at time t agent 1's priority is shifted $4 - 2 = 2$ positions right to be ordered last, and the priority ordering is updated to $4, 3, 2, 1$. Note that when shifting an agent's ordering to right, the agents with lower initial priority will be relatively shifted left. The following example illustrates how Adaptive RSD changes the assigned probabilities according to the dominance history.

Example 5.8. Consider 3 agents and 3 objects with preferences as shown in Table 5.7. Assume that the dominance history is $d_{2,3} = 1$ and 0 for all others. Algorithm 3 balances the priority orderings according to the dominance history. Table 5.8 illustrates the set of all priority orderings under RSD and Adaptive RSD. For instance, the ordering $1, 2, 3$ would change to $1, 3, 2$ because agent 2 has dominated agent 3 in some period in the past. Table 5.9 shows the random allocations under simple RSD and Adaptive RSD.

Proposition 5.2. *When the dominance history is balanced, i.e.* $d_{ij}^{t-1} = 0, \forall i, j \in N$, *Adaptive RSD is equivalent to RSD.*

Proof. According to Algorithm 3, for each priority ordering $f \in \mathcal{F}$ since for all $i, j \in N$ the dominance history is balanced, *i.e.*, $d_{ij}^{t-1} = 0$, it is easy to see that the updated

	a	b	c
1	3/6	1/6	2/6
2	0	4/6	2/6
3	3/6	1/6	2/6

(a) RSD

	a	b	c
1	3/6	1/6	2/6
2	0	2/6	4/6
3	3/6	3/6	0

(b) Adaptive RSD

Table 5.9: Allocations under RSD and Adaptive RSD.

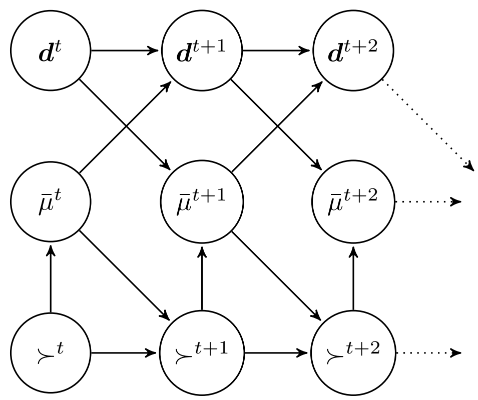


Figure 5.9: Influence diagram representing the conditional dependencies for the Adaptive RSD mechanism.

priority ordering $f' = f$. Thus at time t, the multiset $\mathcal{F}'$ is equivalent to the set of all priority orderings $\mathcal{F}$, and the a uniform distribution over $\mathcal{F}'$ is equivalent to the RSD mechanism. □

Figure 5.9 illustrates the influence diagram for the Adaptive RSD mechanism. Note that the history of matchings has been compressed into a single dominance variable in each state.

Avoiding Cycles

One notable property of Adaptive RSD in Algorithm 3 is that it prevents the possibility of pairwise *cycles* when updating priority orderings based on the dominance history. Given the pairwise dominance relation between each two agents, it may be the case that, using a simple swapping mechanism between two agents, a dominating agent would be re-prioritized higher than the agent that was dominated. Consider for example an ordering of $1, 2, 3$ where $d_{2,3} = 1$ and $d_{1,2} = 1$. If a mechanism first updates the priorities of

2 and 3 by swapping their orderings, the ordering will be $1, 3, 2$. Now, since agent 1 has also dominated agent 2, swapping their orderings leads into a priority of $2, 3, 1$, wherein a dominating agent 2 has higher priority than agent 3 which was dominated by agent 2. Likewise, if a mechanism first swaps the priorities of agents 1 and 2, and then agents 2 and 3, the priority ordering will change first to $2, 1, 3$ and finally to $3, 1, 2$, leaving agent 1 in higher priority than agent 2, which was dominated by agent 1.

The Adaptive RSD starts by comparing the *lowest priority* agents with the those in higher priority orderings, shifting the dominating agent to an order immediately after the dominated agent. In the above example, the priority ordering will be updated to $3, 2, 1$. Thus, under the Adaptive RSD mechanisms dominated agents always get priority over their dominating agent counterparts.

Local Properties of Adaptive RSD

Another implication of the Adaptive RSD mechanism is that it preserves the local properties of the RSD mechanism, that is, in each period the allocation is ex post efficient locally. Moreover, the non-bossiness of Adaptive RSD ensures that any attempt to changing the allocation of other agents also changes the allocation of a non-truthful agent.

Proposition 5.3. *Adaptive RSD is ex post efficient and non-bossy in each period.*

Proof. Adaptive RSD prescribes a random matching over serial dictatorship mechanisms induced by priority orderings. A serial dictatorship mechanism is Pareto efficient, and thus, randomization over a set of Pareto efficient mechanisms satisfies ex post efficiency.

The non-bossiness is the direct implication of serial dictatorships: the support of the Adaptive RSD mechanism consists solely of serial dictatorships. Any serial dictatorship mechanism is non-bossy [133, 134], implying that any distribution over this set that is independent of reported preferences is also non-bossy. Thus, Adaptive RSD is locally non-bossy. □

5.8.2 Strategic Behavior

At each period, an agent's strategic behavior is either through manipulating the immediate outcome (at the current period) or affecting the decision trajectory to gain advantage sometime in the future. We first focus on strategic behavior of agents at the current state and then investigate the types of strategies that an agent may choose to play to obtain a more preferred outcome (now or in future) by misrepresenting its preferences.

Lemma 5.4. *Given any dominance history $\boldsymbol{d}^{t-1}$, Adaptive RSD is lsd-strategyproof, i.e., for each agent $i \in N$,*

$$\forall j \in N, \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ_i^t \mu(j)}} \pi(\mu|(\succ^t, \boldsymbol{d}^{t-1})) \geq \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ_i^t \mu(j)}} \pi(\mu|((\hat{\succ}_i^t, \succ_{-i}^t), \boldsymbol{d}^{t-1}))$$

Proof. When the dominance history is balanced, the proof follows from Proposition 5.2 and *sd*-strategyproofness of RSD in single-shot settings.

Let $\mathcal{F}'$ be the multiset of priority orderings after updating the set $\mathcal{F}$ given the dominance history of agents up to now $\boldsymbol{d}^{t-1}$. Each priority ordering $f' \in \mathcal{F}'$ correspondents exactly to a single serial dictatorship. Agent i's strategic report $\hat{\succ}_i^t$ does not improve its assignment in any of the induced deterministic serial dictatorships, that is, for all orderings f', if $\mu^t = SD(\succ^t, f')$ and $\nu^t = SD((\hat{\succ}_i^t, \succ_{i-1}^t), f')$ then $\mu^t(i) \succeq_i^t \nu^t(i)$. Adaptive RSD draws from a uniform distribution over the multiset $\mathcal{F}'$, implying that for all agents $i \in N$, $\forall y \in M$

$$\sum_{\substack{x \in M: \\ x \succ_i^t y}} \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) = x}} \pi(\mu|(\succ^t, \boldsymbol{d}^{t-1}) \geq \sum_{\substack{x \in M: \\ x \succ_i^t y}} \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) = x}} \pi(\mu|((\hat{\succ}_i^t, \succ_{-i}^t), \boldsymbol{d}^{t-1}))$$

Since the equation is valid for every pair of agents, and each matching is a bijective from agents to alternatives, in the above inequality, for every $x, y \in M$ where $x \succ_i^t y$ we replace x with $\mu(i)$ and y with $\mu(j)$. This precisely implies the inequality in Lemma 5.4. □

According to Lemma 5.4, an agent cannot improve its allocation at time t by misreporting at time t. Lemma 5.2 implies that an agent can only change the decision trajectory by changing the matching (and hence the allocation of some agents) at time t. An agent's possible strategy to influence the sequential outcome is through reducing or increasing its dominance on a subset of agents to affect the priority orderings prescribed by the Adaptive RSD mechanism in the future.

Given the Adaptive RSD mechanism, the dominance relation between a pair of agents is not necessarily symmetric, that is, given a preference profile $\succ^t$ it is possible for two agents i and j to have $\omega_{ij}^t(\pi(\succ^t)) \neq \omega_{ji}^t(\pi(\succ^t))$. For example, consider a dominance history matrix of all zeros and three agents with preferences as follows: $\succ_1^t$: $a \succ b \succ c$, $\succ_2^t$: $b \succ a \succ c$, and $\succ_3^t$: $a \succ c \succ b$. Given Adaptive RSD (or sequential RSD), it is easy to see that $\omega_{12}^t(\pi(\succ^t)) = \frac{1}{6}$ while $\omega_{21}^t(\pi(\succ^t)) = \frac{2}{6}$ and $\omega_{23}^t(\pi(\succ^t)) = 0$ while $\omega_{32}^t(\pi(\succ^t)) = \frac{1}{6}$.

One implication of asymmetric dominance relations is that an agent may strategically change its dominance on other agents by misreporting its current preference to benefit in the future. We say that an agent's strategy is *dominance reducing* if it reports a preference that minimizes its current dominance on some other agents. This type of strategizing is similar to dynamic house allocation for 2 periods of decisions where agents can opt out in the first period to get priority in the second period [2].

Definition 5.15. *Agent i's reporting strategy $\sigma_i(\succ_i^t)$ at t is **dominance reducing** if given a matching policy π, for some $j \in N$,*

$$\omega_{ij}^t(\pi(\sigma_i(\succ_i^t), \succ_{-i}^t)) < \omega_{ij}^t(\pi(\succ_i^t, \succ_{-i}^t))$$

The next lemma states that given a random matching prescribed by Adaptive RSD, agent i's strategy to minimize its expected dictatorial dominance on agent j conversely results in reducing expected dominance caused by agent j on agent i.

Lemma 5.5. *Given the Adaptive RSD mechanism, and strategy profile $\sigma_i(\succ_i^t)$, if $\omega_{ij}^t(\pi(\sigma_i(\succ_i^t), \succ_{-i}^t)) < \omega_{ij}^t(\pi(\succ_i^t, \succ_{-i}^t))$ for some $j \in N$, then*

$$\omega_{ji}^t(\pi(\sigma_i(\succ_i^t), \succ_{-i}^t)) < \omega_{ji}^t(\pi(\succ_i^t, \succ_{-i}^t))$$

Proof. For the ease of notation, and since the history up to now would be fixed for any strategic report, we write $\omega_{ij}^t(\pi(h^t)) = \omega_{ij}^t(\pi(\succ^t))$. By Equation 5.6, the probability that an agent prefers its own allocation to another agent's allocation complements the probability that it prefers the other agent's allocation to its own, that is, $\forall \succ \in \mathcal{P}^n$, $\omega_{ij}^t(\pi(\succ^t)) + \bar{\omega}_{ij}^t(\pi(\succ^t)) = 1$. According to the dominance reducing strategy, for misreport $\hat{\succ}_i = \sigma_i(\succ_i^t)$ we should have $\omega_{ij}^t(\pi((\hat{\succ}_i^t, \succ_{-i}^t))) < \omega_{ij}^t(\pi((\succ_i^t, \succ_{-i}^t)))$. Using Equation 5.7, we can write this inequality as

$$1 - \bar{\omega}_{ij}^t(\pi((\hat{\succ}_i^t, \succ_{-i}^t))) < 1 - \bar{\omega}_{ij}^t(\pi((\succ_i^t, \succ_{-i}^t)))$$
$$\bar{\omega}_{ij}^t(\pi((\hat{\succ}_i^t, \succ_{-i}^t))) > \bar{\omega}_{ij}^t(\pi((\succ_i^t, \succ_{-i}^t)))$$

Thus, by reporting a dominance reducing preference, agent i's outcome gets improved, indicating that the probability of obtaining an object that is more preferred compared to $\mu(j)$ according to the report $\hat{\succ}_i^t$ will be higher than the same probability according to the truthful report $\succ_i^t$, *i.e.*, $\bar{\omega}_{ji}^t(\pi((\hat{\succ}_i^t, \succ_{-i}^t))) > \bar{\omega}_{ji}^t(\pi((\succ_i^t, \succ_{-i}^t)))$.

For contradiction assume that agent i increased the domination of agent j on himself by misreporting $\omega_{ji}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) > \omega_{ji}^t(\pi(\succ_i^t, \succ_{-i}^t))$. By feasibility of the matching mechanism we have $\bar{\omega}_{ji}^t(\pi((\hat{\succ}_i^t, \succ_{-i}^t))) < \bar{\omega}_{ji}^t(\pi((\succ_i^t, \succ_{-i}^t)))$, which contradicts the above finding. □

We say that an agent's strategy is *dominance increasing* if it reports a preference that increases its current dominance on some other agents.

Definition 5.16. *Agent i's reporting strategy $\sigma_i(\succ_i^t)$ at t is **dominance increasing** if given a matching policy π for some $j \in N$,*

$$\omega_{ij}^t(\pi(\sigma_i(\succ_i^t), \succ_{-i}^t)) > \omega_{ij}^t(\pi(\succ_i^t, \succ_{-i}^t))$$

Note that an agent may play a strategy that is dominance increasing on some agents while being dominance reducing on some another agents. An agent can manipulate the allocation of other agents by playing a dominance reducing or increasing strategy to change the evolution of preferences.

We show that dominance reducing or increasing strategies do not improve an agent's outcome, if not accompanied by changing the evolution of preferences. Thus, manipulation in the global sense may still be possible by changing the allocation, and thus, influencing the evolution of preferences. Recall that our desired solution concept states that given any transition model, no agent can obtain a more preferred sequence of outcomes, no matter how other agents play now or in the future. Therefore, an agent's misreport requires changing the current random outcome either to decrease its dominance in the current time step, in the hope for improved priority (and thus, matchings) in the future, or to increase its dominance and changing the evolution of preferences to a more preferred trajectory.

Algorithm 3 ensures that after realization of the random decision, a dominating agent loses priority against a dominated agent in future steps. The next lemma follows directly from the Adaptive RSD algorithm (see Appendix E for the detailed proof).

Lemma 5.6. *Given two states $h^{t+1} = (\succ^{t+1}, \boldsymbol{d}^t)$ and $\hat{h}^{t+1} = (\succ^{t+1}, \hat{\boldsymbol{d}}^t)$ such that for agents $i, j \in N$, $\hat{d}_{ij}^t > d_{ij}^t$, we have*

$$\sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ_i^{t+1} \mu(j)}} \pi(\mu \mid h^{t+1}) \geq \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ_i^{t+1} \mu(j)}} \pi(\mu \mid \hat{h}^{t+1}) \tag{5.8}$$

The implication of this lemma is that if an agent is dominated by another agent in some previous decisions, for all next steps where their dominance history is unbalanced (*i.e.*, the dominated agent has not had a chance to benefit from its priority over the dominating agent), the agent always prefers its own outcome to the outcome of the dominating agent. In other words, the dominated agent receives higher priority in all future steps until it dominates the other agent in a matching decision.

5.8.3 Two Agents

In Section 5.7.1 we showed that gsd-strategyproofness of the sequential RSD mechanism depends on the planning horizon and the transition functions, and in fact, under history dependent transitions, sequential RSD is not gsd-strategyproof. Here, we show an intriguing result about the Adaptive RSD mechanism, which states that given history-dependent transitions, for any planning horizon, a successful manipulation of Adaptive RSD results in a Pareto improvement.

Theorem 5.6. *Given Adaptive RSD and history-dependent transition, for two agents any strong manipulation results in a Pareto improvement.*

Proof. There are two agents i and j with private preferences over two objects. Assume agent i's strategic report $\hat{\succ}_i^t$ results in gaining from strong manipulation, *i.e.*, $W_i^\pi(((\hat{\succ}_i^t, \succ_j^t), \boldsymbol{d}^{t-1}), o_\ell) \geq W_i^\pi((\succ^t, \boldsymbol{d}^{t-1}), o_\ell), \forall o_\ell \in M$, with at least one strict inequality. Consider the case where agent i's strategy is dominance increasing at t, that is

$$\omega_{ij}^t(\pi((\hat{\succ}_i^t, \succ_j^t), \hat{\boldsymbol{d}}^{t-1})) > \omega_{ij}^t(\pi(\succ^t, \hat{\boldsymbol{d}}^{t-1}))$$

Since i has increased its dominance on j, then by Lemma 5.6 in the future steps, Adaptive RSD prioritized agent j over i. However, by assumption agent i has improved its outcome by misreporting. By lsd-strategyproofness of Adaptive RSD, agent i's immediate outcome at t after misreport would not improve, and by Proposition 5.3 it must be the case that $\pi_i((\succ^t, \boldsymbol{d}^{t-1})) \succ_i^t \pi_i(((\hat{\succ}_i^t, \succ_j^t), \boldsymbol{d}^{t-1}))$, which implies that agent j may only benefit in the future steps. With two agents and two objects, there are four possible profiles, $(2!)^2$, in each period. Thus, increasing dominance means that for agent j's immediate outcome, similar to agent i, we have $\pi_j((\succ^t, \boldsymbol{d}^{t-1})) \succ_j^t \pi_j(((\hat{\succ}_i^t, \succ_j^t), \boldsymbol{d}^{t-1}))$.

By assumption, since agent i's outcome has improved in the future, we have $W_i^\pi(((\hat{\succ}_i^t, \succ_j^t), \boldsymbol{d}^{t-1}), o_\ell) \geq W_i^\pi((\succ^t, \boldsymbol{d}^{t-1}), o_\ell), \forall o_\ell \in M$. However, agent i's outcome improves only on those profiles that its preference ordering is not equal to that of agent j. Thus, in all periods that i receives its top choice, by symmetry, agent j also receives its top choice, implying that $\forall o_\ell \in M,\ W_j^\pi(((\hat{\succ}_i^t, \succ_j^t), \boldsymbol{d}^{t-1}), o_\ell) \geq W_j^\pi((\succ^t, \boldsymbol{d}^{t-1}), o_\ell)$.

For dominance reducing strategy, when there are two agents, any dominance reducing strategy by agent i would strictly improve agent j's immediate outcome. By Lemma 5.5, agent i's reduction in dominance on agent j would also lead to symmetric decrease in agent j's dominance on i. Thus, agent i would not receive a higher priority in the future periods. The rest of the proof follows exactly as it was stated for the dominance increasing case. □

$$\begin{pmatrix} \succ_1\colon a \succ b \succ c \\ \succ_2\colon b \succ a \succ c \\ \succ_3\colon a \succ c \succ b \end{pmatrix}$$

(a) A sample preference profile.

	1	2	3
1	0	1/6	3/6
2	2/6	0	0
3	3/6	1/6	0

(b) The dominance profile $\boldsymbol{\omega}^t$, denoting the dominance relation for every pair of agents.

Table 5.10: An example showing that two agents' dominance relation depends on the mechanism and preferences of other agents.

For two agents, based on Theorem 5.3 it is easy to see that under Markovian transitions Adaptive RSD and sequential RSD are both gsd-strategyproof. Under history-dependent transitions, Adaptive RSD aligns the incentives of a strategic agent with improving the overall optimality of the problem. This alignment of incentives is due to the fact that under ordinal preferences, the strategic agent's gain is tightly coupled with the other competing agent: both agents' incentives are aligned only when their preferences are unequal (or unaligned).

In the next section, we show that adding one more agent to the setting requires other subtle assumptions about how preferences evolve to ensure desirable incentive properties.

5.8.4 Manipulation and Preference Dynamics

Measuring how two agents' preferences evolve to more conflicting or non-conflicting profiles depends heavily on the matching mechanism and the preferences of other agents. Not only does the dominance relation between two agents depend on how their preferences are similar to one another, but it also depends on other agents' preferences.

Example 5.9. Consider a preference profile as shown in Table 5.10. Looking only at the preferences of each pair, it is easy to see that without considering the matching mechanism or the other agents, since agent 2's top choice differs from agent 1 and 3's top choices, agent 2's preference does not conflict with any of the other two. Thus, neither would dominate one another. However, a more careful look at the preferences, given RSD as the matching mechanism, reveals that in a serial dictatorship with agent ordering $3, 1, 2$, both agents 1 and 3 dominate agent 2. Hence, given RSD as matching policy π, the dominance relation for each i and j, $\omega^t_{i,j}(\pi(\succ^t))$, can be shown as illustrated in Table 5.10b.

Using the notion of dictatorial dominance, for each pair of agents we compute the expected sum over matchings where agent i dominates agent j up to the planning horizon.

Formally, given a preference profile $\succ^t$, matching policy π, and a planning horizon λ, the expected number of matchings wherein agent i dominates agent j in the future (not including the current decision) is written as

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{ij}^{\tau}(\pi(\succ^{\tau}, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right] \tag{5.9}$$

In a setting with deterministic allocations but with uncertain preference transitions, Equation 5.9 shows the expected number of times, looking forward into the future, wherein agent i dominates agent j. For example, consider a situation where in each period agents' preferences evolve such that all agents have different top choices. In this case, no agent dominates another agent in this trajectory of matching decisions. Likewise, if agents have identical preferences in all possible states (homogeneous preferences), then starting from the initial period and given sequential RSD, each agent gets the same chance of dominating another agents.

5.8.5 Manipulation With Dominance Increasing Strategies

We now reach our main results in this section, which depend on how the preferences evolve to different profiles in the future. We first show that given a transition function, under Adaptive RSD no agent can benefit from a dominance increasing strategy if, in the future, preferences evolve such that a dominated agent gets the chance to dominate the dominating agent. In the next section, we show that even when the above assumption does not hold, if an agent's misreport does not cause other agents' preferences to evolve to more conflicting trajectories, then a successful manipulation improves the allocation of all agents in expectation.

Theorem 5.7. *Given history-dependent transitions and agent i's dominance increasing report $\hat{\succ}_i^t$, when other agents report truthfully, Adaptive RSD prevents strong manipulation if for any j such that $\omega_{ij}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) > \omega_{ij}^t(\pi(\succ_i^t, \succ_{-i}^t))$ we have*

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{ji}^{\tau}(\pi(\succ^{\tau}, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}_i^t, \succ_{-i}^t)\right] \geq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{ji}^{\tau}(\pi(\succ^{\tau}, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right] \tag{5.10}$$

Proof. Consider history-dependent transition dynamics. For weak gsd-strategyproofness, we must prove that agent i's misreport does not improve its overall allocation outcome, considering both immediate and future possible outcomes. The lsd-strategyproofness of

Adaptive RSD in Lemma 5.4 implies that no agent can immediately benefit from misreporting. Thus, we only need to worry about future possible outcomes.

Let $\hat{\succ}_i^t$ denote agent i's strategic report at time t. According to Lemma 5.2, if agent i's misreport does not affect the random decision at the current state, then the overall trajectory remains unchanged, looking forward into the future. Thus, agent i can only change the decision trajectory by changing the random outcome at t. Formally, if $\hat{\succ}_i^t$ results in $\pi(\hat{h}^t) \neq \pi(h^t)$, then by lsd-strategyproofness (Lemma 5.4) and non-bossiness (Proposition 5.3) of Adaptive RSD it must be that $\pi_i(h^t) \succ_i^t \pi_i(\hat{h}^t)$. Note that strict inequality is the direct implication of Adaptive RSD's non-bossiness in Proposition 5.3. Thus, we can focus on agent i's strategic report which results in a different random outcome at the current step.

The state at time t is denoted by $h^t = (\succ^t, \boldsymbol{d}^{t-1})$, and assume that $\boldsymbol{d}^{t-1}$ is balanced. Let $\boldsymbol{d}^t$ and $\hat{\boldsymbol{d}}^t$ denote the dominance history reporting according to $\succ_i^t$ and $\hat{\succ}_i^t$ respectively. For simplicity, we use $\hat{\succ}^t = (\hat{\succ}_i^t, \succ_{-i}^t)$ to denote the preference profile at which agent i is misreporting but all other agents are reporting truthfully. Suppose agent i's misreport does not change its dominance on a subset of agents while it is dominance increasing for another subset of agents.

Consider a subset of agents $N_d \subset N$ such that for each $k \in N_d$, after misreporting agent i has less than or equal dominance, that is, $\omega_{ik}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) \leq \omega_{ik}^t(\pi(\succ^t))$. By Lemma 5.5 we have $\omega_{ki}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) \leq \omega_{ki}^t(\pi(\succ^t))$ yielding that $\hat{\boldsymbol{d}}_{ki}^t \leq \boldsymbol{d}_{ki}^t$. Thus, in all future steps, agent i does not benefit by receiving any priority over k.

We focus on those agents $j \in N \setminus N_d$ that agent i increases its dominance on them at time t. By assumption, for each j, $\omega_{ij}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) > \omega_{ij}^t(\pi(\succ_i^t, \succ_{-i}^t))$ we have

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{ji}^{\tau}(\pi(\succ^{\tau}, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}_i^t, \succ_{-i}^t)\right] \geq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{ji}^{\tau}(\pi(\succ^{\tau}, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right]$$

which means that if an agent is dominated by i its preferences (along with agent i's preferences) evolve such that they do compete over resources in the future. Adaptive RSD ensures that in all future steps $\tau > t$, a dominated agent j is prioritized over the dominating agent i, which implies that agent j would prefer its own outcome to i in the future periods. By Equation 5.7, we know that $\omega_{ij}^t(\pi(h^t)) = 1 - \bar{\omega}_{ij}^t(\pi(h^t))$. Applying this to the above inequality we have

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} 1 - \bar{\omega}_{ji}^{\tau}(\pi(\succ^{\tau}, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}_i^t, \succ_{-i}^t)\right] \geq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} 1 - \bar{\omega}_{ji}^{\tau}(\pi(\succ^{\tau}, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right]$$

which further simplifies to

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}_{ji}^{\tau}(\pi(\succ^{\tau}, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}_i^t, \succ_{-i}^t)\right] \leq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}_{ji}^{\tau}(\pi(\succ^{\tau}, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right] \tag{5.11}$$

By definition, $\bar{\omega}_{ji}$ is the probability that agent i prefers its own outcome to agent j's outcome. Thus, a dominated agent will be prioritized over the dominating agent, and given the assumption in Equation 5.10, it has equal chance of benefiting from this priority and reciprocating in the future periods.

Since agent i cannot gain immediate benefit at t by misreporting, and by Proposition 5.3 and lsd-strategyproofness of Adaptive RSD we have $\bar{\omega}_{ji}^t(\pi(\hat{\succ}^t, \boldsymbol{d}^{t-1})) < \bar{\omega}_{ji}^t(\pi(\succ^t, \boldsymbol{d}^{t-1}))$. Adding this inequality to Equation 5.11, shows that for all agents j, agent i prefers its outcome to agent j's under truthful report. This inequality holds for all agents. Therefore, there exists at least one ranking position o_ℓ such that agent i strictly prefers its outcome under truthfulness, implying that Adaptive RSD is weakly gsd-strategyproof. □

Theorem 5.7 assumes history-dependent transitions. Because Markovian transitions can be considered as a special case of history-dependent transitions with only one step memory, Theorem 5.7 similarly holds for Markovian transitions.

Corollary 5.1. *Theorem 5.7 holds for Markovian transition dynamics.*

Note that Equation 5.10 prevents the transition dynamics to evolve to states that are non-conflicting for agent i but may result in more conflicting trajectories for all other agents. This is because a strategic agent can report a preference such that preferences evolve to states where the dominated agents would not have a chance to reciprocate (due to completely different and non-conflicting preferences); henceforth, manipulation is still possible without the assumption in Equation 5.10. It is important to note that Adaptive RSD only prevents certain types of manipulation (dominance increasing) and cannot prevent all types of manipulations: An agent may still manipulate the outcome by reducing its dominance on other agents and influencing the evolution of preferences such that its preferences do not conflict with those of other agents in the future periods.

5.8.6 Pareto Improving Manipulations

We show that even though Adaptive RSD cannot prevent all types of manipulations, under some mild assumption on preference dynamics, any successful manipulation – whether dominance increasing or reducing – that improves an agent's outcome is harmless to other agents and leads to a Pareto improvement.

Theorem 5.8. *Given history-dependent transitions and Adaptive RSD, agent i's strategic report $\hat{\succ}_i^t$ resulting in successful strong manipulation results in a Pareto improvement in expectation if for all $j, k \in N \setminus i$ we have*

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{jk}^{\tau}(\pi(\succ^{\tau}, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}_i^t, \succ_{-i}^t)\right] \leq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega_{jk}^{\tau}(\pi(\succ^{\tau}, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right] \tag{5.12}$$

Proof. To show that any strong manipulation that improves a strategic agent's outcome would cause Pareto improvement, we must show that if $\forall o_\ell \in M,\ W_i^\pi(\hat{h}^t, o_\ell) \geq W_i^\pi(h^t, o_\ell)$ with possibly one strict inequality, then for all $j \in N \setminus i$ we have $\forall o_\ell \in M,\ W_j^\pi(\hat{h}^t, o_\ell) \geq W_j^\pi(h^t, o_\ell)$. Note that the assumption in Equation 5.12 prevents the transition dynamics to evolve to states that are non-conflicting for agent i but may result in more conflicting trajectories for all other agents.

Let $\hat{\succ}_i^t$ denote agent i's strategic report at time t. Recall from Section 5.8.1 that history at time t can be presented as $h^t = (\succ^t, \boldsymbol{d}^{t-1})$. By lsd-strategyproofness of Adaptive RSD, agent i's misreport at t cannot improve its immediate outcome. According to Lemma 5.2, if agent i's misreport does not influence the random decision at the current state, then the overall trajectory remains unchanged, looking forward into the future. Thus, agent i can only change the trajectory by changing the random outcome at t. By Proposition 5.3 we have that $\pi_i(h^t) \succ_i^t \pi_i(\hat{h}^t)$.

Assume that agent i's misreport has either increased or has not changed its dominance on some agents while strictly reducing its dominance on a subset of agents $N_d \subseteq N$. Formally, $\forall j \in N_d, \omega_{ij}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) < \omega_{ij}^t(\pi(\succ^t))$, where $N_d \subseteq N$ denotes a subset of agents on which i's dictatorial dominance has strictly reduced, and for all other agents $\forall j' \in N \backslash N_d, \omega_{ij'}^t(\pi(\hat{\succ}_i^t, \succ_{-i}^t)) \geq \omega_{ij'}^t(\pi(\succ^t))$. Let $\boldsymbol{d}^t$ and $\hat{\boldsymbol{d}}^t$ denote the dominance history after reporting $\succ_i^t$ and $\hat{\succ}_i^t$ respectively, and assume that the dominance history up until now is balanced, *i.e.*, $\forall i, j,\ \boldsymbol{d}_{i,j}^{t-1} = 0$.

The idea is that an agent's misreport may lead the evolution of preferences to a sequence where agent i's allocation improves while (some) other agents get into more conflicting states with one another (but not agent i), and thus, the misreport may be harmful to other agents. In the following steps, we show that given our assumption on preference dynamics (Equation 5.8), such harmful manipulation is impossible.

Step (*i*): $\forall j' \in N \backslash N_d, \omega^t_{ij'}(\pi(\hat{\succ}^t_i, \succ^t_{-i})) \geq \omega^t_{ij'}(\pi(\succ^t))$, meaning that agent i's misreport increases its dominance on some agents. Since the dominance history up until t is the same, by the exclusive disjunction for the dominance profile we can immediately write $\hat{\boldsymbol{d}}^t_{ij'} > \boldsymbol{d}^t_{ij'}$.

For Pareto improvement, assume that $\hat{\boldsymbol{d}}^t_{ij'} > \boldsymbol{d}^t_{ij'}$, but agent i has changed the decision trajectory such that $\forall o_\ell \in M,\ W^\pi_i(\hat{h}^t, o_\ell) \geq W^\pi_i(h^t, o_\ell)$. This means that the expected dominance of other agents on i is reduced, which implies that the new decision trajectory consists of profiles, $\succ^\tau$, where agent i's dominance on other agents is decreased (profiles where agent i's preference is different from others) such that $\omega^\tau_{ij'}(\pi((\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1}))) < \omega^\tau_{ij'}(\pi((\succ^\tau, \boldsymbol{d}^{\tau-1})))$. Formally for all $j' \in N \setminus N_d$,

$$\sum_{\succ^\tau} \sum_{\mu \in \mathcal{M}} T(\succ^\tau |(\succ^t, \hat{\boldsymbol{d}}^t), \mu)\pi(\mu|h^t)\ \omega^\tau_{ij'}(\pi((\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1}))) < \sum_{\succ^\tau} \sum_{\mu \in \mathcal{M}} T(\succ^\tau |(\succ^t, \boldsymbol{d}^t), \mu)\pi(\mu|\hat{h}^t)\ \omega^\tau_{ij'}(\pi((\succ^\tau, \boldsymbol{d}^{\tau-1})))$$

which for all periods we can write as

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega^\tau_{ij'}(\pi(\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}^t_i, \succ^t_{-i})\right] < \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega^\tau_{ij'}(\pi(\succ^\tau, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right]$$

Using Equation 5.7, it can be written as

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}^\tau_{j'i}(\pi(\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}^t_i, \succ^t_{-i})\right] \leq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}^\tau_{j'i}(\pi(\succ^\tau, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right] \tag{5.13}$$

Applying Lemma 5.5 to $\tau > t$, it must be the case that for all $j' \in N \setminus N_d$, $\omega^\tau_{j'i}(\pi((\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1}))) < \omega^\tau_{j'i}(\pi((\succ^\tau, \boldsymbol{d}^{\tau-1})))$. Given the inequality in Equation 5.12, for all two agents $j', k \in N \setminus i$ we can write

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}^\tau_{j'k}(\pi(\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}^t_i, \succ^t_{-i})\right] \geq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}^\tau_{j'k}(\pi(\succ^\tau, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right]$$

which means that all other agents prefer their sequence of outcomes (in expectation) under agent i's misreport.

Step (*ii*): $\forall j \in N_d, \omega^t_{ij}(\pi(\hat{\succ}^t_i, \succ^t_{-i})) < \omega^t_{ij}(\pi(\succ^t))$, that is, agent i's strategic misreport reduces its dominance on a subset of agents. By Lemma 5.5, reducing dominance on an agent reduces the expected domination on agent i, thus, $\forall j \in N_d, \omega^t_{ji}(\pi(\hat{\succ}^t_i, \succ^t_{-i})) < \omega^t_{ji}(\pi(\succ^t))$, and $\hat{\boldsymbol{d}}^t_{ji} < \boldsymbol{d}^t_{ji}$. This implies that the immediate dominance between agents i and j has strictly decreased.

Given that agent i has improved its overall assignment in the future periods (and not immediately) by misreporting, implies

$$\mathbb{E}\left[\sum_{\tau=t} \bar{\omega}^\tau_{ji}(\pi((\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1}))) \mid (\hat{\succ}^t_i, \succ^t_{-i})\right] > \mathbb{E}\left[\sum_{\tau=t} \bar{\omega}^\tau_{ji}(\pi((\succ^\tau, \boldsymbol{d}^{\tau-1}))) \mid \succ^t\right]$$

Given the assumption, for all $j, k \in N_d$ we have

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega^\tau_{jk}(\pi(\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}^t_i, \succ^t_{-i})\right] \leq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \omega^\tau_{jk}(\pi(\succ^\tau, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right]$$

which can be written as

$$\mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}^\tau_{jk}(\pi(\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1})) \mid (\hat{\succ}^t_i, \succ^t_{-i})\right] \geq \mathbb{E}\left[\sum_{\tau=t+1}^{\lambda} \bar{\omega}^\tau_{jk}(\pi(\succ^\tau, \boldsymbol{d}^{\tau-1})) \mid \succ^t\right]$$

Adding the inequalities in Step I and Step II, with those of agent i after manipulation, for all agents in $j \in N$ we have

$$\mathbb{E}\left[\sum_{\tau=t}^{\lambda} \sum_{i \in N} \bar{\omega}^\tau_{ji}(\pi((\succ^\tau, \hat{\boldsymbol{d}}^{\tau-1}))) | (\hat{\succ}^t_i, \succ^t_{-i})\right] \geq \mathbb{E}\left[\sum_{\tau=t}^{\lambda} \sum_{i \in N} \bar{\omega}^\tau_{ji}(\pi((\succ^\tau, \boldsymbol{d}^{\tau-1}))) | \succ^t\right] \tag{5.14}$$

For each agent j, by Equation 5.7 we know that

$$\bar{\omega}^t_{ji}(\pi(h^t)) = \sum_{\mu \in \mathcal{M}: \mu(j) \succ^t_j \mu(i)} \pi(\mu \mid h^t)$$

which implies that agent j prefers its outcome to agent i. By inequality 5.14, each agent prefers its outcome to all other agents' allocations. Therefore, for any o_ℓ, it must be that $W^\pi_j(((\hat{\succ}^t_i, \succ^t_{-i}), \boldsymbol{d}^{t-1}), o_\ell) \geq W^\pi_j((\succ^t, \boldsymbol{d}^{t-1}), o_\ell)$. This holds for all agents, which means that in a sequence of random matchings such that at some $\tau > t$ every agent's assignment (including agent i) improves when agent i successfully manipulates the evolution of preferences, implying that for all agents $i \in N$, $W^\pi_i(((\hat{\succ}^t_i, \succ^t_{-i}), \boldsymbol{d}^{t-1}), o_\ell) \geq W^\pi_i((\succ^t, \boldsymbol{d}^{t-1}), o_\ell), \forall o_\ell \in M$. □

5.8.7 Strategyproofness Guarantee

Theorem 5.7 and 5.8 showed that under certain preference dynamics a strategic agent cannot benefit from misreporting, and in situations where successful manipulation is possible, often this manipulation can lead to a Pareto improvement. In this section, we reach an interesting result, which is independent of the previous assumptions on how preferences evolve in the future. Without any assumption on transition dynamics, we show that Adaptive RSD is weakly gsd-strategeyproof if the planning horizon is at most equal to the number of agents – no strong manipulation is possible.

Theorem 5.9. *Adaptive RSD is weakly gsd-strategyproof for all $\lambda \leq n$.*

Proof. Consider n agents and n objects and a strategic agent i with a misreport $\hat{\succ}_i^t$ at time t when its true preference is $\succ_i^t$. To prove weak gsd-strategyproofness, we need to show that strong manipulability is impossible under the Markovian assumption. Formally, we must show that there exists some ranking position o_ℓ such that $W_i^\pi(\succ^t, o_\ell) > W_i^\pi((\hat{\succ}_i^t, \succ_{-i}^t), o_\ell)$.

Fixing the dominance history up to time t, we assume that the history $\boldsymbol{d}^{t-1}$ is balanced. Lemma 5.4 implies that agent i's misreport cannot improve its immediate outcome. By Lemma 5.2, an agent can only influence the evolution of states by changing the decision at the current time t. If $\hat{\succ}_i^t$ results in $\pi(\hat{h}^t) \neq \pi(h^t)$, then by non-bossiness of Adaptive RSD (Proposition 5.3) agent i cannot change the allocation without changing its own outcome. By lsd-strategyproofness of Adaptive RSD (Lemma 5.4) it must be that $\pi_i(h^t) \succ_i^t \pi_i(\hat{h}^t)$. Thus, agent i can change the sequence of outcomes (trajectory) by changing the matching at time t.

We show that agent i can only weakly manipulate the outcome by potentially changing the decision trajectory. We provide a proof by construction and show that even if there exist non-conflicting trajectories, strong manipulation is not possible for $\lambda < n + 1$.

Suppose that agent i's strategic report takes the agent to a trajectory where in all steps agent i receives its top choice. Thus, for agent i we can write,

$$W_i^\pi(\hat{\succ}^t) = \mathring{\mu}^t(i) + (\lambda - 1)\overset{*}{\mu}(i), \tag{5.15}$$

where $\mathring{\mu}^t(i)$ is agent i's allocation at the time t when it misreports, and $\overset{*}{\mu}(i)$ is a best-case outcome where agent i receives its top choice according to Lemma 5.1.

Now we consider agent i's allocation when reporting truthfully. Suppose that agent i's allocation at time t consists of a worst-case matching outcome (as introduced in Lemma 5.1). Lemma 5.3 states that in the first period there exists a subset of matchings $\mathcal{M}' \subseteq \mathcal{M}$

with $\sum_{\mu \in \mathcal{M}'} \pi(\mu|\hat{\succ}^t) > 0$ such that $\sum_{\mu \in \mathcal{M}'} \pi(\mu| \succ^t) \geq \frac{1}{n}$. Thus, under truthfulness there is at least $1/n$ chance that the policy chooses a matching decision precisely as the decision under misreport and continue into the same trajectory. Therefore, with the probability of at least $\frac{1}{n}$ the evolution of preferences follow the best-case trajectory, while with probability of $\frac{n-1}{n}$ the decision trajectory falls into the worst-case outcome for agent i in all future steps. Thus, agent i's allocation for λ planning steps can be written as,

$$W_i^\pi(\succ^t) = \frac{1}{n}(o_1 + \ldots + o_n) + (\frac{1}{n})\overset{*}{\mu}(i) + \frac{(\lambda-1)(n-1)}{n}[\frac{1}{n}(o_1 + \ldots + o_n)]$$

which can be written for every ranking position o_ℓ as

$$W_i^\pi(\succ^t, o_\ell) = \sum_1^\ell \frac{1}{n} + (\frac{\lambda - 1}{n}) + (\frac{(\lambda-1)(n-1)}{n}) \sum_1^\ell \frac{1}{n} \tag{5.16}$$

For weak gsd-strategyproofness, we must show that at least for one ranking position o_ℓ, the inequality $W_i^\pi(\succ^t, o_\ell) > W_i^\pi((\hat{\succ}_i^t, \succ_{-i}^t), o_\ell)$ holds. Given equations 5.15 and 5.16, and by replacing $\overset{*}{\mu}(i)$ with top choice o_1, we must show that

$$\sum_1^\ell \frac{1}{n} + (\frac{\lambda - 1}{n}) + (\frac{(\lambda-1)(n-1)}{n}) \sum_1^\ell \frac{1}{n} > \mathring{\mu}^t(i) + (\lambda - 1)$$

By lsd-strategyproofness of Adaptive RSD we know that agent i's allocation under truthfulness stochastically dominates its allocation under misreport. Proposition 5.3 implies that $\pi(\hat{h}^t) \neq \pi(h^t)$. Thus, agent i's outcome at time t, is always worse under misreport. Note that in general a misreport may result in an allocation that is as good as the allocation under a truthful report; however, by Lemma 5.2 such misreport does not affect the evolution of preferences at all. Therefore, for $\lambda = 1$ we will have

$$\sum_1^\ell \frac{1}{n} > \mathring{\mu}^t(i)$$

We first show that for o_1, manipulation is always possible when $\lambda > 1$. For manipulability, We replace $\ell = 1$ and write

$$\begin{aligned}\frac{1}{n} + (\frac{\lambda-1}{n}) + (\frac{(\lambda-1)(n-1)}{n})\frac{1}{n} &< (\lambda-1) \Leftrightarrow \\ \frac{1}{n} + (\frac{(\lambda-1)(n-1)}{n})\frac{1}{n} &< \frac{(\lambda-1)(n-1)}{n} \Leftrightarrow \\ 1 + (\frac{(\lambda-1)(n-1)}{n}) &< (\lambda-1)(n-1) \Leftrightarrow \\ \frac{n}{(n-1)^2} &< (\lambda-1)\end{aligned}$$

which holds for any $\lambda \geq 2$ and $n > 2$, implying that the score of obtaining the first rank object under the misreport is strictly greater that the score on under truthfulness.

To show that the mechanism is not strongly manipulable, we need to show such inequality does not hold for all ranking positions. For all $\lambda > 1$, we show that there exists a ranking position ℓ where the following inequality strictly holds

$$\sum_{1}^{\ell}\frac{1}{n} + (\frac{\lambda-1}{n}) + (\frac{(\lambda-1)(n-1)}{n})\sum_{1}^{\ell}\frac{1}{n} > \mathring{\mu}^t(i) + (\lambda-1)$$

Assume $\ell = n-1$, then we can simplify the above inequality and write:

$$\begin{aligned}\frac{1}{n}(n-1) + (\frac{\lambda-1}{n}) + (\frac{(\lambda-1)(n-1)}{n})\frac{1}{n}(n-1) &> \lambda-1 \Leftrightarrow \\ \frac{1}{n}(n-1) + (\frac{\lambda-1}{n}) + (\frac{(\lambda-1)(n-1)^2}{n^2}) &> \lambda-1 \Leftrightarrow \\ \frac{1}{n}(n-1) + (\frac{(\lambda-1)(n-1)^2}{n^2}) &> \frac{(\lambda-1)(n-1)}{n}\end{aligned}$$

Multiplying both sides by n and by algebraic simplification we have,

$$\begin{aligned}(n-1) + (\frac{(\lambda-1)(n-1)^2}{n}) &> (\lambda-1)(n-1) \Leftrightarrow \\ 1 + (\frac{(\lambda-1)(n-1)}{n}) &> \lambda-1 \Leftrightarrow \\ 1 &> \lambda-1-(\frac{(\lambda-1)(n-1)}{n}) \Leftrightarrow \\ 1 &> \frac{\lambda-1}{n}\end{aligned}$$

which implies $\lambda < n + 1$, and states that there is a ranking position at which the allocation under truthful is strictly better than that of misreporting, implying that the Adaptive RSD mechanism is weakly gsd-strategyproof. □

Under any type of transition dynamics, a strong manipulation by moving the evolution of preferences to more desired states is impossible if $\lambda \leq n$. Note that our analysis in this theorem provides a lower bound for λ, because it poses no assumption on how preferences evolve and assumes the discounting factor of $\gamma = 1$, *i.e.*, future gains are as important as immediate allocations. Therefore, $\lambda \leq n$ provides the minimum planning horizon at which the Adaptive RSD mechanism guarantees weak gsd-strategyproofness. It is still possible to improve this result for larger planning horizons by considering a diminishing discount factor for future allocations or by making further assumptions on preference dynamics and the population of agents with similar preferences.

While an agent can weakly manipulate and receive an incomparable sequence of outcomes in the future, the manipulation is only harmful to some other agents only if the strategic agent dominates another agents and can guarantee no conflict of preferences with those dominated agents in the future. Therefore, the Adaptive RSD mechanism provides incentives for truthfulness if agents' preferences dynamics are similar. In many markets, agents tend to express similar preferences due to similar or cohesive beliefs about the alternatives [1, 9]. In contrast, as we discussed in this section, in markets with diverse preferences, a strategic agent may still be able to benefit from misreporting by exploiting the structure and dynamics of other agents' preferences.

5.9 Fairness in Sequential Matchings

Fairness is a vital desirable property in designing allocation mechanisms in multiagent settings. Mechanism designers often seek to deploy allocation mechanisms that treat the participating agents fairly while guaranteeing strategyproofness and efficiency.

In this section, we argue that in contrast to sequential RSD that may result in unbounded envy among agents, as a result of balancing priorities, Adaptive RSD satisfies some desirable notions of local and global fairness.

We consider two notions of fairness in sequential matching problems; a local fairness notion of *equity* (equal treatment of equals), and a global notion of *ex post envyfreeness.*

Definition 5.17. *A sequential matching mechanism is* ***equitable*** *at each time step t if and only if $\forall i, j \in N$ with identical dominance histories $d_{ij}^{t-1} = d_{ji}^{t-1}$, if $\succ_i^t = \succ_j^t$ then*

$$\forall y \in M, \sum_{x \in M: x \succ_i^t y} p_i^t(x|(\succ^t, \boldsymbol{d}^{t-1})) = \sum_{x \in M: x \succ_j^t y} p_j^t(x|(\succ^t, \boldsymbol{d}^{t-1}))$$

In sequential mechanisms, envyfreeness relies on balancing priorities over the course of the assignment sequence. Ex post envyfreeness is the strongest notion of fairness that states that at each time step no pair of agents should be envious of one another. We consider ex post envyfreeness at each time based on the sequence of decisions up to and including the current period.

Definition 5.18. *Given a sequence of matchings $(\mu^1, \ldots \mu^t)$, a matching mechanism is* ***periodic ex post envyfree*** *(PEF) if at all times t, every two agents have dominated one another in equal number of matchings, i.e., $\forall i, j \in N$, $\sum_{s=1}^{t} \ \omega_{ji}^s(\mu^s) = \sum_{s=1}^{t} \ \omega_{ij}^s(\mu^s)$.*

The PEF notion of fairness is a strict requirement that is hard to guarantee in most small markets for allocating indivisible alternatives without the use of monetary compensation.

Proposition 5.4. *Adaptive RSD does not satisfy PEF, but does yield a sequence of equitable local matchings.*

Proof. Consider two agents i, j with identical preferences over two objects at $t = 1$, *i.e.*, $\succ_i^1 = \succ_j^1 = a \succ b$. Assume that Adaptive RSD assigns the more preferred object to $\mu^1(i) = a$, leaving j envious. For all future times, assume the following preferences $\forall t > 1, \succ_i^t = b \succ_i a$ and $\succ_j^t = a \succ_j b$. Thus, agent j will never dominate i and will remain envious of i.

To prove the equitability of Adaptive RSD at each round, we need only to show that Adaptive RSD treats all agents with identical preferences and dominance histories equally. By Proposition 5.2, we know that Adaptive RSD is equivalent to RSD when $d_{ij}^{t-1} = 0, \forall i, j \in N$, and thus satisfies equitability. By Algorithm 3, Adaptive RSD only adjusts the priority orderings when agents' dominance histories are unbalanced. Thus, it assigns the same weight to the priority orderings of agents with equal dominance histories, that is, for all t when $d_{ij}^t = d_{ji}^t$. □

5.9.1 Degree of Envy

The nonexistence of ex post envyfreeness for $n > 2$, raises a crucial question of whether our mechanism ensures some degree of envy. Among several plausible ways of defining envy [45], we consider a natural notion of envy; the envy of a single agent towards another agent. Given a history of assignments h^t, agent i's *degree of envy* with respect to agent j is

$$e_{ij}(h^t) = \sum_{s=1}^{t} \left[\omega_{ji}^s(\mu^s) - \omega_{ij}^s(\mu^s)\right]$$

Definition 5.19. *A sequential mechanism is c-envious if for all times t, $\forall i \in N$ agent i is envious to agent j for at most c assignments. That is, $c = \max_{ij} e_{ij}(h^t), \forall t$.*

Theorem 5.10. *An Adaptive RSD matching mechanism is* 1*-envious.*

Proof. By dominance history of Adaptive RSD at all times $\forall i, j \in N$, $d_{ij}^t = 1 \rightarrow d_{ji}^t = 0$, and if $d_{ij}^t = 1$, then for all future times $\tau > t$ when the dominance history is unbalanced, $f'(j) > f'(i)$. Thus, agent i cannot dominate j for all future transitions until it gets dominated by j at least once. Moreover, agent i's preference dynamic may evolve so that it never conflicts with agent j's preference. Thus, $\forall t, e_{ji}(h^t) \leq 1$, implying that Adaptive RSD is 1-envious. □

In fact, Adaptive RSD interplays between random assignments in repeated decisions to maintain an approximately fair global policy. It is easy to see that the maximum envy of a society of agents with Adaptive RSD mechanism is $\frac{n(n-1)}{2}$.

5.10 Discussion

In this chapter, we studied dynamic matching problems without the use of transferable currencies. We investigated the incentive and fairness properties of sequential matching mechanisms when agents have dynamic ordinal preferences. Our contributions in this chapter are:

A model for dynamic matching

We provided a model for reasoning over and analyzing ordinal matching decisions in dynamic settings. We formulated a generic dynamic matching problem using a *history-dependent matching process*, with states of the matching process corresponding to a history of preference profiles and matching decisions, and developed a scoring function for evaluating sequences of matching decisions. Finally, we introduced a number of properties we argue are important for matching mechanisms in dynamic settings.

Extending conventional matching mechanisms

Focusing on a strategyproof matching mechanisms in one-shot settings, namely Random Serial Dictatorship (RSD), we showed that it is susceptible to manipulation in repeated allocations with dynamic preferences. Hence, one cannot simply run a sequence of RSD assignments in dynamic settings. Furthermore, we showed that in general, this manipulation result is strong, that is, for any utility functions consistent with agents' underlying preferences, manipulation of the mechanism can be beneficial.

The Adaptive RSD mechanism

We proposed a new mechanism (Adjusted RSD) that maintains a history of pairwise interactions between agents, and adapts the priority orderings of agents in each period based on this history. We showed that our mechanism is locally strategyproof, and that there are situations where it is globally strategyproof (*e.g.* when there are 2 agents, when agents have similar preference dynamics, and when the planning horizon is bounded), and even when the mechanism is manipulable, the manipulative actions taken by an agent will often result in a Pareto improvement in general. Thus, we make the argument that while manipulative behavior may still be unavoidable, it is not necessarily at the cost to other agents in the system.

5.11 Future Work

The model and results in this chapter raise a number of interesting future research directions. Although our proposed mechanism provides a solution for achieving strategyproof and fair allocations in certain type of situations, designing a matching mechanism that satisfies truthful incentives for any population of agents is still an open problem. Moreover, characterizing the set of truthful matching policies in dynamic settings is an interesting future direction, particularly when agents are capable of learning after each matching decision. One possible first direction is to study *deterministic sequential matching* mechanisms and their extensions for repeated allocations when preferences evolve dynamically over time.

To circumvent the issues of incentive design in dynamic settings, one may consider matching problems with agents that have particular underlying utility functions (*e.g.* linear positional utility functions) or consider problems with a homogeneous set of agents with restricted preference dynamics. In Chapter 6, we focus on these restricted cases, and introduce particular classes of problems for which positive results exist in dynamic settings.

The incompatibility of ordinal efficiency and strategyproofness in static settings [30] prevents us from designing truthful optimal policies in dynamic settings. However, there may exist some approximately efficient random policies in the policy space that incentivizes truthfulness in sequential settings, perhaps, by renouncing the ex post efficiency or other local requirements. An important open question is whether in the ordinal domain one can design a matching policy with desirable incentive properties that guarantees a sequence from the set of Pareto frontier matchings in expectation.

Finally, in large one-shot markets where there are large number of copies of each object (such as assigning students to housing), the stochastic inefficiency of the RSD mechanism vanishes [44]. One potential direction would be to study the efficiency and envyfreeness of sequential matching problems in markets with multiple capacities and various agent to object ratios.

Chapter 6

Matching with Restricted Preference Dynamics

In Chapter 5, we proposed and analyzed a sequential matching model for dynamic preferences where monetary payment is not permitted. We showed that simply repeatedly using the Random Serial Dictatorship (RSD) mechanism (a simple and well-studied matching mechanism with a number of desirable properties when used in static domains) could lead to complex strategic behavior by agents. We then proposed a new, history-dependent, mechanism which guaranteed global strategyproofness under some mild assumptions.

In this chapter, we revisit this general result for a single class of utility functions and look at overcoming this impossibility result by restricting the preferences. Inspired by the mechanism design literature on quasi-linear utilities [67] and single-peaked preferences [107, 125], we exploit a particular utility function from the literature that allows us to formulate the matching problem as a planning problem and leverage Markov Decision Process (MDP) models. We show that the more general impossibility result still holds in this setting. However, if we place additional restrictions on the *dynamics* of the agents' preferences, then there are interesting subclasses for which possibility results exist.

6.1 The Model

We first describe the matching model for dynamic preferences, one that extends that of Chapter 5, by showing that under certain assumptions the model can be encoded as a

Markov Decision Process, and thus the (sequential) matching problem can be modeled as a planning problem.

Let $N = \{1, ..., n\}$ be a set of agents, and $M = \{1, ..., m\}$ be the set of alternatives to be assigned to the agents, where $n = m$.[1] Each agent has a strict preference ordering over the set of items at time t denoted by $\succ_i^t$, where $a \succ_i^t b$ means that agent i strictly prefers item a to item b at time t. Let $\mathcal{P}(M)$ or $\mathcal{P}$ denote the class of all strict linear preferences over M where $|\mathcal{P}| = m!$. We let $\succ^t = (\succ_1^t, \ldots, \succ_n^t) \in \mathcal{P}^n$ denote the preference orderings of the agents at time t, and refer to $\succ^t$ as the *preference profile*. We write $\succ_{-i}^t$ to denote $(\succ_1^t, \ldots, \succ_{i-1}^t, \succ_{i+1}^t, \ldots, \succ_n^t)$, and thus $\succ^t = (\succ_i^t, \succ_{-i}^t)$. In addition, we assume that each agent is endowed with a private utility function, u_i, that is consistent with its preference ordering, that is, $u_i(\succ_i^t, a) > u_i(\succ_i^t, b)$ at time t if and only if $a \succ_i^t b$. Throughout this chapter, we often instantiate homogeneous utility functions, that is, agents share the same form of the utility functions but may have different underlying ordinal preferences.

A matching $\mu^t : N \to M$ is a mapping that assigns a unique item to each agent. A matching is *feasible* iff for all $i, j \in N$, $\mu^t(i) \neq \mu^t(j)$ when $i \neq j$. We let $\mathcal{M}$ denote the set of feasible matchings.

A *matching mechanism* $\pi : \mathcal{P}^n \to \Delta(\mathcal{M})$ returns a probability distribution over the set of all possible matchings for each time t. Thus, $\pi(\succ^t)$ is the induced policy under preference profile $\succ^t$. We also write $\pi(\mu^t | \succ^t)$ to denote the probability of a deterministic matching μ^t given the preference profile $\succ^t$. Note that a matching mechanism is based on ordinal preferences (as opposed to cardinal utilities) and returns a (randomized) matching decision given a profile of reported preferences. Since we assume that agents have utility functions, given a preference profile $\succ^t$ the expected utility for agent i at time t under matching mechanism π is

$$\mathbb{E}_\pi[u_i | \succ^t] = \sum_{\mu \in \mathcal{M}} u_i(\succ_i^t, \mu(i)) \; \pi(\mu | \succ^t) \tag{6.1}$$

The basic setup in this chapter is same as the model in Chapter 5, with the addition of specific utility functions. We are interested in settings where there is a sequence of matching decisions and agents' preferences *evolve* over time. In particular, we assume that the preference held by an agent at time t depends on the preferences it held earlier along with allocations (*i.e.* matchings) made previously. In sequential settings, agent's preferences may evolve or change based on idiosyncratic preferences or previous matching outcomes. Therefore, it could be the case that $\succ_i^t \neq \succ_i^{t+1}$, *i.e.* $a \succ_i^t b$ while $b \succ_i^{t+1} a$. We

[1] For $n > m$, we add $n - m$ *dummy items* corresponding to a null assignment to the set M.

assume there is an underlying stochastic transition model $T(\succ^{t+1} \mid \succ^t, \mu^t)$ which denotes the probability that agents will transition to a state where they have joint preference $\succ^{t+1}$ after matching decision μ^t in a state with joint preference $\succ^t$. Furthermore, we assume the stochastic transition model is *common knowledge* and it is independent across agents, meaning that $T(\succ^{t+1} \mid \succ^t, \mu^t) = \prod_{i \in N} T_i(\succ_i^{t+1} \mid \succ_i^t, \mu^t(i))$.

Given the matching problem description above, we can model it as a multiagent Markov Decision Process (MDP) $\langle S, A, T, R, \gamma \rangle$ where S is the state space, A is the action space, $T(\cdot|\cdot,\cdot)$ is the transition function between states given some action and $R : S \mapsto \mathbb{R}_{\geq 0}$ is the reward function. In particular: The state space S is the set of preference profiles, *i.e.* $S = \mathcal{P}^n$, the action space A is the set of feasible matchings, *i.e.* $A = \mathcal{M}$, the transition function between states is precisely defined by the stochastic transition function, *i.e.* $T(\succ^{t+1} \mid \succ^t, \mu^t)$, the reward function at state $\succ^t$ given matching μ^t is $R(\succ^t |\mu^t) = \sum_{i \in N} u_i(\succ_i^t, \mu^t(i))$, and γ is a discount factor such that $0 < \gamma \leq 1$. The reward function, R, in each state is the (utilitarian) social welfare at that state given the utility functions.

Abusing notation somewhat, a *policy*, for the above MDP, returns a (randomized) matching for each state (preference profile). Note that a policy, along with the transition function, implicitly specifies the sequences of future decisions.

For a particular policy π, given a planning horizon λ and assuming that agents have revealed $\succ^t$ at time t, the expected utility of agent i is

$$V_i(\succ^t, \pi) = \mathbb{E}\left[\sum_{k=t}^{\lambda} \gamma^{k-t} \sum_{\mu \in \mathcal{M}} u_i(\succ_i^k, \mu(i)) \pi(\mu | \succ^k)\right]$$

The *expected value* of policy π for preference profile (state) $\succ^t$ is thus

$$V(\succ^t, \pi) = \sum_{i \in N} V_i(\succ^t, \pi).$$

6.1.1 Properties

In Chapter 5, we studied the dynamic matching problem for general preferences, without any assumption on the underlying utility models. In this chapter, since we are interested in finding and evaluating matching policies for particular utility models, the problem is a special case of those studied in Chapter 5. Thus, we are looking for solution concepts that guarantee equilibria given specific utility models and populations with homogeneous utility functions and satisfy weaker notions of truthfulness.

One of the most important goals in sequential matching is to find policies that maximize the overall utility of the participating agents. In addition, a desirable policy should prescribe an efficient matching decision in each period. An (ex ante) optimal policy maximizes the expected sum of utilities and is always guaranteed to exist [115].

Definition 6.1 (Optimal Policy)**.** *A randomized policy is (ex ante) optimal if,*

$$\pi^* = \arg\max_{\pi \in \Pi} V(\succ^t, \pi), \forall \succ^t \in \mathcal{P}^n.$$

Given a utility model, a randomized matching prescribed by policy π at time t is (ex ante) Pareto efficient if no other randomized matching increases the expected utility of one agent without decreasing the expected utility of all others.

Definition 6.2 (Pareto Efficiency)**.** *Given agents' utility models, a randomized matching prescribed by π at time t is (ex ante) Pareto efficient if for all other randomized policies π', for every agent $i \in N$ we have*

$$\mathbb{E}_{\pi}[u_i | \succ^t] \geq \mathbb{E}_{\pi'}[u_i | \succ^t] \tag{6.2}$$

Since the preferences of the agents are private, the matching policy relies on having the agents reveal their preferences at each time step. However, agents may be strategic and given the matching policy and the preferences of other agents, a strategic agent reports a preference ordering $\sigma_i((\succ_i^t, \succ_{-i}^t)|\pi)$ so as to maximize its own expected utility according to

$$\sigma_i((\succ_i^t, \succ_{-i}^t)|\pi) \in \arg\max_{\hat{\succ}_i^t \in \mathcal{P}} V_i(\succ^t, \pi((\hat{\succ}_i^t, \succ_{-i}^t))).$$

where $\pi((\hat{\succ}_i^t, \succ_{-i}^t))$ is the induced policy when agent i misreports at time t (while all other agents are truthful), assuming that all agents (including agent i) report truthfully in all future periods.

In dynamic settings, we assume that we have access to the agents' utility models that are consistent with their underlying private preferences. The goal is to achieve an equilibrium in which each agent's best response is to play the equilibrium strategy, knowing the current private preferences of other agents and the expectation over future preferences. Namely, we are interested in *within-period ex post Nash equilibrium* [15, 25, 42] as a refinement of Bayes-Nash equilibrium where a best-response strategy is solely based on the revelation of other agents' preferences (or types) at the current period and not the realization of their future preferences.

A mechanism (policy) is within-period stochastic Ex Post Incentive Compatible (w.p.s EPIC) if each agent maximizes its utility by reporting truthfully when all other agents are truthful in the current period and all future periods.

Definition 6.3 (w.p.s EPIC). *A matching mechanism π is within-period stochastic ex post incentive compatible or w.p.s EPIC if at all periods t, for every agent $i \in N$, for all true preferences $\succ^t \in \mathcal{P}^n$, and for any misreport $\hat{\succ}_i^t = \sigma_i(\succ^t | \pi)$,*

$$V_i(\succ^t, \pi(\succ^t)) \geq V_i(\succ^t, \pi((\hat{\succ}_i^t, \succ_{-i}^t))) \tag{6.3}$$

Given a randomized matching mechanism, we define the notion of incentive compatibility when agents play in equilibrium going forward from the current period, after seeing the preferences of other agents at the current period and before the realization of the randomized matching decision. Note that this is a weaker notion compared to the global stochastic strategyproofness used in Chapter 5, where truthfulness is a dominant strategy for all possible utility functions.

6.1.2 Utility Function Model

While none of the formalisms so far has relied on the particular instantiation of utility functions, in this section we introduce the utility functions we use for the rest of the chapter. In particular, we use a simple linear positional utility function, based on the Borda score used in both the social choice and matching literature [12, 43, 53].

Definition 6.4 (Linear Utility). *The linear utility of agent i with preference ordering $\succ_i^t$ for matching $\mu^t(i)$ is*

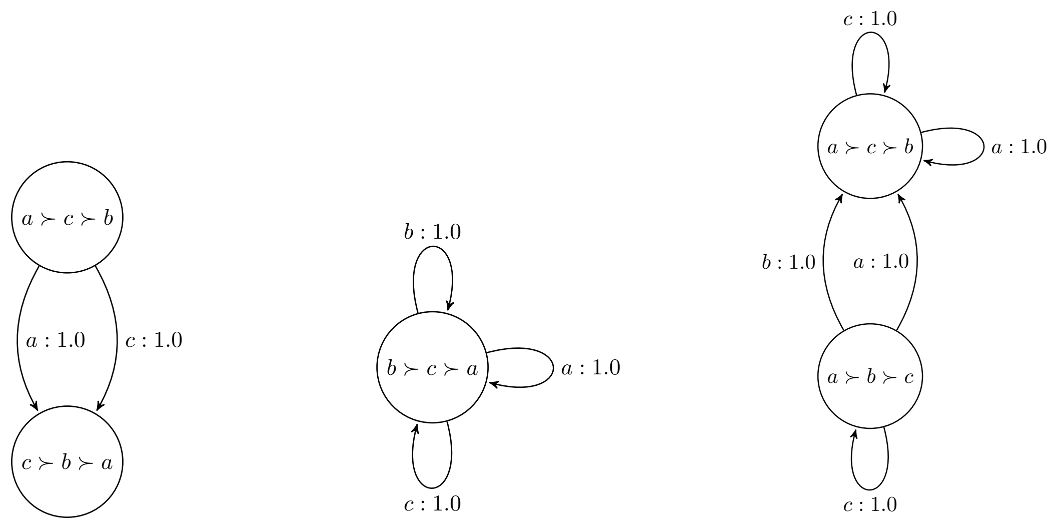
$$u_i(\succ_i^t, \mu^t(i)) = m - rank(\succ_i^t, \mu^t(i)),$$

where the function $rank : \mathcal{P} \times M \to \mathbb{N}$ returns the rank of item $\mu^t(i)$ in agent's preference $\succ_i^t$ at time t.

6.2 The Sequential RSD Mechanism

In single step matching problems efficiency and incentive compatibility constraints are incompatible whether in cardinal domains [82, 145] or ordinal settings [30]. This incompatibility persists for optimal matching policies in sequential settings.

Theorem 6.1. *Let $\langle \mathcal{P}^n, \mathcal{M}, P, R \rangle$ be a matching MDP. There is no optimal (deterministic or randomized) matching policy that satisfies w.p.s EPIC.*



(a) Agent 1 *(b) Agent 2* *(c) Agent 3*

Figure 6.1: Three local MDPs.

Proof. Proof by counterexample (Figure 6.1): consider the problem of assigning 3 items $\{a, b, c\}$ to 3 agents $\{1, 2, 3\}$ in 2 time periods, with 3 local MDPs representing each agent's preference model (starting at top states). The optimal policy is a uniform random tie-breaking over π:(abc, cba) and π':(cba, cba), with $V(\succ^1, \pi)$=$V(\succ^1, \pi')$=11. If agent 3 misreports $\hat{\succ}_3^1 = a \succ b \succ c$, using backward induction the optimal policy of the new joint MDP would be π^*:(cba, cba) with value 11. Thus, agent 3's expected utility improves **from** $\frac{1}{2}V_3(\succ^1, \pi) + \frac{1}{2}V_3(\succ^1, \pi') = 3.5$ **to** $V_3(\succ^1, \pi^*) = 4$. □

This impossibility leads us to consider a well-known strategyproof matching mechanism, Random Serial Dictatorship, as a potential incentive compatible matching policy for dynamic settings. Note that this impossibility result under a particular utility function immediately applies to the generic setting we studied in Chapter 5 without any assumption on the utility models.

In this section, we first revisit the background on Random Serial Dictatorship (RSD) as a strategyproof mechanism in static settings, and then focus attention on the game-theoretic properties of RSD in sequential matching settings.

Definition 6.5 (Serial Dictatorship)**.** *Given a preference profile $\succ$ and an ordering of agents f, simple serial dictatorship (SD) is defined as follows: agent $f(1)$ receives its best* *object m_1 in M according to $\succ_{f(i)}$; agent $f(2)$ receives its best object m_2 in $M \backslash \{m_1\}$; agent $f(n)$ receives its best object m_n in $M \backslash \{m_1, \ldots, m_{n-1}\}$.*

The simple serial dictatorship is the only deterministic mechanism that satisfies strategyproofness, non-bossiness, and neutrality [134]. Thus, randomization over all possible simple serial dictatorships preserves strategyproofness while satisfying the fair treatment of equals. Formally, let $\mathcal{F}$ be the class of all orderings. An ordering is chosen randomly from a uniform distribution over all possible orderings $\mathcal{F}$. Given a preference profile $\succ^t$, the probability of matching decision μ prescribed by RSD is

$$\tilde{\pi}(\mu | \succ^t) = \frac{1}{n!} |\{f \in \mathcal{F} : \mathcal{SD}(f, \succ^t) = \mu\}| \tag{6.4}$$

RSD satisfies strategyproofness for all possible utility models that are consistent with preference profiles. That is, for all agents $i \in N$ and for any u_i we have $\mathbb{E}_{\tilde{\pi}}[u_i | \succ^t] \geq \mathbb{E}_{\tilde{\pi}}[u_i | (\hat{\succ}^t_i, \succ^t_{-i})]$. In fact, RSD is the only mechanism that is ex post efficient, strategyproof, and fair [22].

The *sequential RSD* is a matching policy that simply prescribes an RSD matching at each period. A preference ordering that remains unchanged in all periods is called *fixed* or *time-invariant* preference ordering. In sequential settings with fixed underlying preferences, *i.e.* for all agents $i \in N$, $\forall t, \succ^t_i = \succ_i$, sequential RSD is incentive compatible.

Theorem 6.2. *Given fixed preferences, for any utility model consistent with the preferences, sequential RSD is w.p.s. EPIC.*

Proof. By the stochastic strategyproofness of RSD [6], no agent can receive a better random outcome from misreporting at the current state. Since the preferences are fixed (time-invariant), by forward induction no agent would be able to gain a better outcome in any of the future sequences. □

6.2.1 Truthfulness Under Sequential RSD With Particular Utility Models

In this section, we show that sequential RSD is not w.p.s EPIC even in our restricted setting with linear utility functions and independent Markovian transitions.

It is possible for an agent to report a preference ordering at time t, not equal to its true preferences, that does not change the matching in that period. Our first lemma shows that given a fixed deterministic matching, an agent cannot change the underlying preference dynamics by misreporting it's true preferences.

Lemma 6.1 (Lemma 5.2 from Chapter 5)**.** *Fixing $\mu \in \mathcal{M}$, for any $\succ^{t+1} \in \mathcal{P}^n$ given agent i's misreport $\hat{\succ}_i^t$,*

$$T(\succ^{t+1} \mid \succ^t, \mu) = T(\succ^{t+1} \mid (\hat{\succ}_i^t, \succ_{-i}^t), \mu)$$

Two Agents: With only two agents, each agent's strategic behavior is restricted only to misreporting its top choice in the current period in order to gain expected reward in the future. By RSD induced decisions, no agent can have an immediate expected gain by misreporting. With linear utilities, the *expected immediate cost of misreporting* to agent i at time t is defined by the following lemma.

Lemma 6.2. *With linear utility functions, for $n = 2$,*

$$\mathbb{E}_{\tilde{\pi}}[u_i(\succ_i^t, \mu) \mid \succ^t] - \mathbb{E}_{\tilde{\pi}}[u_i(\succ_i^t, \mu) \mid (\hat{\succ}_i^t, \succ_{-i}^t)] = \frac{1}{2}$$

Proof. Let $\succ^t = (\succ_1^t, \succ_2^t)$ be the preferences of the two agents. Consider first the case where $\succ_1^t = \succ_2^t$. It is clear that expected utility under RSD assignment is $\frac{1}{2}$ for both agents. Deviating from a truthful report will result in a deterministic allocation where the strategic agent's utility reduces to 0, while the other agent's utility improves to 1. Now consider the case where $\succ_1^t \neq \succ_2^t$. In this case, both agents will receive their top choices, and thus, any misreporting would reduce the expected utility for both agents from 1 to $\frac{1}{2}$. □

The sequential RSD yields a sequence of Pareto efficient matchings while satisfying incentive compatibility.

Theorem 6.3. *The sequential RSD mechanism is w.p.s EPIC for $n = 2$ when agents have linear utility functions.*

Proof. We must show that $V_i(\succ^t, \tilde{\pi}(\succ^t)) \geq V_i(\succ^t, \tilde{\pi}((\hat{\succ}_i^t, \succ_{-i}^t)))$, for any $\hat{\succ}_i^t = \sigma_i(\succ^t \mid \tilde{\pi})$ when the other agent reports truthfully. Note that the transition function T is known and common knowledge. We start by proving the case with only 2 time steps ($\lambda \leq 2$) and then the general case ($\lambda > 2$).

Case $\lambda \leq 2$: For contradiction, we assume that $V_i(\succ^t, \tilde{\pi}((\hat{\succ}_i^t, \succ_{-i}^t))) > V_i(\succ^t, \tilde{\pi}(\succ^t))$, and expand this inequality as follows:

$$\mathbb{E}_{\tilde{\pi}}[u_i(\succ_i^t, \mu) + \gamma \textstyle\sum_{\succ^{t+1}} T(\succ^{t+1} \mid \hat{\succ}^t, \mu) u_i(\succ_i^{t+1}, \mu) \mid \hat{\succ}^t] >$$
$$\mathbb{E}_{\tilde{\pi}}[u_i(\succ_i^t, \mu) + \gamma \textstyle\sum_{\succ^{t+1}} T(\succ^{t+1} \mid \succ^t, \mu) u_i(\succ_i^{t+1}, \mu) \mid \succ^t]$$

where $\hat{\succ}^t = (\hat{\succ}^t_i, \succ^t_{-i})$ is the preference profile at time t when i misreports and all other agents are truthful. For a stronger result we assume no discounting ($\gamma = 1$). By rearranging the immediate utilities and future expected utilities and using Lemma 6.1 we have

$$\mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} \mid \succ^t, \mu) u_i(\succ^{t+1}_i, \mu) \mid \hat{\succ}^t] -$$
$$\mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} \mid \succ^t, \mu) u_i(\succ^{t+1}_i, \mu) \mid \succ^t] >$$
$$\mathbb{E}_{\tilde{\pi}}[u_i(\succ^t_i, \mu)) \mid \succ^t] - \mathbb{E}_{\tilde{\pi}}[u_i(\succ^t_i, \mu) \mid (\hat{\succ}^t_i, \succ^t_{-i})]$$

Lemma 6.2 implies that the immediate expected cost of misreporting at any period is $\frac{1}{2}$. The agent must report truthfully in the last step $(t+1)$, thus a misreport could only affect the expectation over rewards at $t+1$. Applying Lemma 6.2 to time $t+1$, the best strategy only improves the future reward by at most $\frac{1}{2}$, which is less than or equal the immediate cost of misreporting, contradicting the assumption.

Case $\lambda > 2$: Each misreport imposes an expected cost of 1/2. The agent must be truthful in any step that he wishes to gain more reward, after some number of misreports. We exploit the 2 time-step model by backward reasoning: at λ agent must reveal its true preference; otherwise there will be an expected loss. At $\lambda - 1$, similar to the previous case with 2 time steps, the best response is to report truthfully. Misreporting at $\lambda - 2$ and then being truthful from then on yields the expected utility of

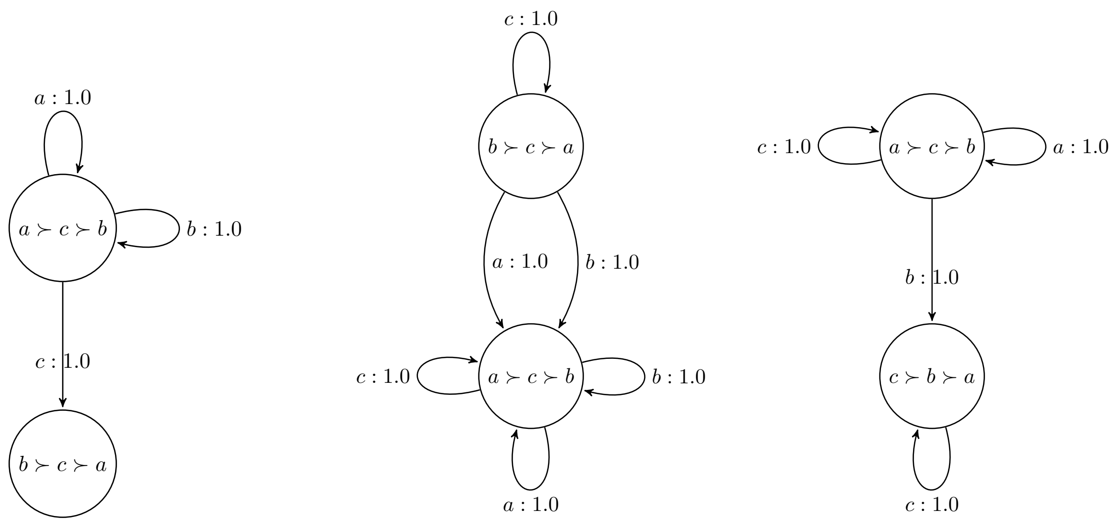
$$\mathbb{E}_{\tilde{\pi}}[u_i(\succ^\lambda_i, \mu) | \succ^\lambda] + \mathbb{E}_{\tilde{\pi}}[u_i(\succ^{\lambda-1}_i, \mu) | \succ^{\lambda-1})] - \frac{1}{2}$$

which is less than $\mathbb{E}_{\tilde{\pi}}[\sum_{t=\lambda-2}^{\lambda} u_i(\succ^t_i, \mu) | \succ^t]$, and so on for all $t \in \lambda$. Therefore, in no period the agent can gain more than its immediate cost of misreporting. □

Three Agents (and more): In the general problem of matching with dynamic preferences for $n \geq 3$, sequential RSD does not satisfy the strong incentive requirement of strategyproofness in the global sense (Chapter 5, Theorem 5.1). The global strategyproofness requires incentive compatibility for the space of all utilities consistent with preferences. We show that sequential RSD is still manipulable when the problem is restricted to the weaker solution concept of within-period stochastic ex post equilibrium.

Theorem 6.4. *For agents with dynamic preferences and linear utilities, sequential RSD is not w.p.s EPIC when $n \geq 3$.*

Proof. Consider the problem of assigning 3 items $\{a, b, c\}$ to 3 agents $\{1, 2, 3\}$ in 3 time periods, with local MDPs representing agents' preference models (Figure 6.2). Only reachable states are shown. The multiagent MDP is shown in Figure 6.3, with "start" showing



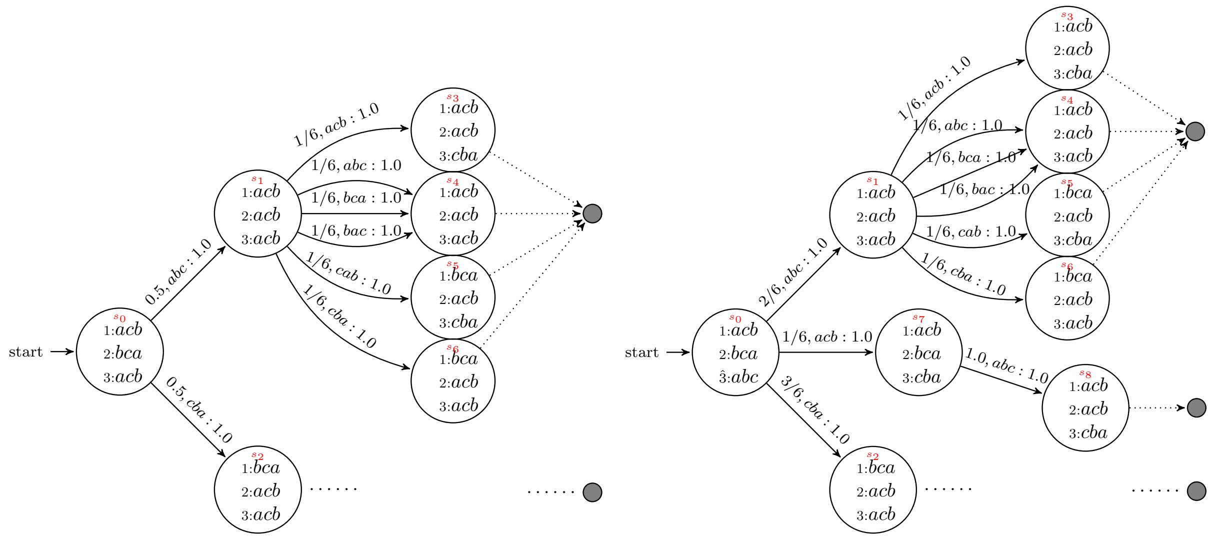
(a) Agent 1 *(b) Agent 2* *(c) Agent 3*

Figure 6.2: Three dynamic preferences represented via 3 local MDPs, showing only relevant actions and transitions. Each state is an ordering, and the transitions are deterministic.

the reported preferences at the beginning, and gray nodes indicating final states. The edges x,act:y represent the probability x of action 'act' (assigning items a,c,t to agents 1,2,3 respectively) according to RSD and transitioning to the next state with probability y. Using backward induction we compute the value of sequential RSD policy for each agent. The expected utility for agent 3 is $V_3(\succ^1, \tilde{\pi}) = \frac{1}{2}[1 + V_3(s_1, \tilde{\pi})] + \frac{1}{2}[2 + V_3(s_2, \tilde{\pi})]$ for both trajectories of decisions. However, agent 3 can improve its expected value by (mis)reporting $\hat{\succ}_3^1 : a \succ b \succ c$ at the beginning, and then being truthful thereafter (illustrated in Figure 6.3b). Note that the decision $\mu^1 = cba$ is chosen with the exact same probability in both cases, so any progress from there would be identical. Agent 3's expected utility after misreporting would be $V_3(\succ^1, \tilde{\pi}(\hat{\succ}_i^1, \succ_{-i}^1)) = \frac{2}{6}[1+V_3(s_1, \tilde{\pi})] + \frac{1}{6}[0+V_3(s_7, \tilde{\pi})] + \frac{3}{6}[2+V_3(s_2, \tilde{\pi})]$. Agent 3's expected utility improves from $\frac{25.08}{6}$ to $\frac{25.38}{6}$, implying that sequential RSD is not w.p.s EPIC.

□

This impossibility result (Theorem 6.4) raises the question *Is there ever any special circumstance where it is possible to use a policy based on RSD for sequential matching?*. Inspired by research in mechanism design for overcoming impossibility results like the Gibbard-Satterthwaite Theorem [63,123], we considered two directions. First, Theorem 6.4 used the solution concept w.p.s. EPIC, and thus it might be possible to use a weaker



(a) Truthful *(b) Misreporting*

Figure 6.3: A three-agent sequential matching problem in three decision periods, with the joint MDP representing the induced social problem. Small nodes represent the absorbing states: no decision is made after reaching these states.

solution concept and derive possibility results (see, for example [98]). However, it is unclear what would be an appropriate weakening of w.p.s. EPIC, which is already a weakening of global strategyproofness. The second alternative is to study whether there are interesting constraints that can be made on the agents preferences (think, for example, of the positive results for the class of single-peaked preferences [107, 125] and quasi-linear utilities [67] in the mechanism design literature). In the next section we pursue this direction and highlight several interesting subclasses of preferences for which interesting results are achievable.

6.3 Restricting Preference Dynamics

We consider two restricted cases, namely, *myopic* agents and *rotational* preferences. We show that even though sequential RSD does not generally satisfy the w.p.s EPIC property, it is incentive compatible under some restriction on agents' preference dynamics. In the next section, we focus attention on another special case called *single-minded* preferences.

6.3.1 Myopic agents

An agent is called *myopic* (short-sighted) if it is only concerned with immediate expected rewards, *i.e.* its best-response strategy is only dependent on current preferences and decisions. For a myopic agent, the discount factor is $\gamma = 0$. Thus, we have

$$\begin{aligned} V_i(\succ^t, \tilde{\pi}) &= \mathbb{E}\left[\sum_{k=t}^{\lambda} \gamma^{k-t} \sum_{\mu \in \mathcal{M}} u_i(\succ_i^k, \mu(i))\tilde{\pi}(\mu| \succ^k)\right] \\ &= \mathbb{E}\left[\sum_{k=t}^{\lambda} 0^{k-t} \sum_{\mu \in \mathcal{M}} u_i(\succ_i^k, \mu(i))\tilde{\pi}(\mu| \succ^k)\right] \\ &= \mathbb{E}\left[\sum_{\mu \in \mathcal{M}} u_i(\succ_i^t, \mu(i))\tilde{\pi}(\mu| \succ^t)\right] \\ &= \mathbb{E}_{\tilde{\pi}}[u_i(\succ_i^t, \mu(i))| \succ^t] \end{aligned}$$

When an agent is myopic, the expected utility is dependent on the immediate matching, and not the future possible allocations. The following theorem follows exactly from this fact.

Theorem 6.5. *Given myopic agents, sequential RSD is w.p.s ex post incentive compatible.*

Proof. Myopic agents only consider the immediate outcomes. Since RSD is strategyproof for a single matching decision, an agent does not benefit from misreporting its preferences at the current period. Thus, sequential RSD is w.p.s ex post incentive compatible when agents are myopic. □

In this special case, each agent only cares about its current assignment, and thus, does not benefit from possible gains in the future steps.

6.3.2 Rotational preferences

Often alternatives lose their desirability after being assigned to an agent. People may lose interest in rereading a book, travelers may prefer a new location after a trip to a specific location, team members may exchange the tasks among themselves in a new project, *etc.* A preference is *rotational* if for all times, the allocated item at time t becomes the least preferred item at $t+1$. Let $\varphi(\succ_i^t, j)$ be a function that prescribes agent i's preference after receiving item j when its preference is $\succ_i^t$. The preference of agent i, $\succ_i^{t+1}$ after being assigned item j is defined as:

$$\varphi(\succ_i^t, j) = \begin{cases} rank(\succ_i^{t+1}, k) = rank(\succ_i^t, k) - 1, & j \succ_i^t k \\ rank(\succ_i^{t+1}, k) = m, & \text{if } k = j \\ rank(\succ_i^{t+1}, k) = rank(\succ_i^t, k) & Otherwise. \end{cases}$$

where k denotes the item in ranking k of preference $\succ_i^t$, and $m = |M|$ is the last possible ranking (see Figure 6.4 for an example). The restriction on the reported preferences limits each agent's reporting strategy as (i) no agent can insist on the same assignment, (ii) the only misreport is through claiming an incorrect preference ordering, thus, reporting a low ranking for an item which would have a high ranking in future steps. With deterministic transition of preferences, agents are only required to submit a single report at the starting state.

Theorem 6.6. *Given rotational preferences, the sequential RSD mechanism is w.p.s EPIC.*

Proof. We must show that for all possible revealing strategies $\sigma_i(\succ^t |\tilde{\pi}) = \hat{\succ}_i$, $V_i(\succ^t, \tilde{\pi}(\succ^t)) \geq V_i(\succ^t, \tilde{\pi}((\hat{\succ}_i^t, \succ_{-i}^t)))$. Since preference changes are deterministic and rotational, the center plans all future decisions based on a single initial claim. Suppose agent i misreports its private preference. Since RSD is strategyproof in each period, for the immediate utility at time t, $\sum_\mu u_i(\succ_i^t, \mu)\tilde{\pi}(\succ^t, \mu) \geq \sum_\mu u_i(\succ_i^t, \mu)\tilde{\pi}((\hat{\succ}_i^t, \succ_{-i}^t))$. For the next time steps, by rotational preferences $\mathbb{E}[V_i((\varphi(\succ_i^t, \mu(i)), \succ_{-i}^{t+1}), \tilde{\pi})| \succ^t] \geq \mathbb{E}[(V_i(\varphi(\succ_i^t, \mu(i)), \succ_{-i}^{t+1}), \tilde{\pi})|(\hat{\succ}_i^t, \succ_{-i}^t)]$, implying that at no time agent's misreport would increase its expected utility, thus, the RSD is truthful for rotational preferences. □

One might conjecture that restricting preferences to rotational models results in truthful implementation of an optimal policy. However, by constructing a single-state MDP, it is easy to see that an optimal policy does not guarantee incentive compatibility.

Theorem 6.7. *An optimal matching policy with rotational preferences is not w.p.s EPIC.*

Proof. Consider a set of agents ($n \geq 3$) with rotational preferences and assume a one-period matching decision. In one-shot settings, given linear utilities an optimal random matching is prescribed by the Probabilistic Serial rule [30], which guarantees stochastic dominance efficiency, but is highly prone to manipulation. Thus, given rotational preferences, an agent can still manipulate the outcome and immediately benefit from misreporting its preferences. □

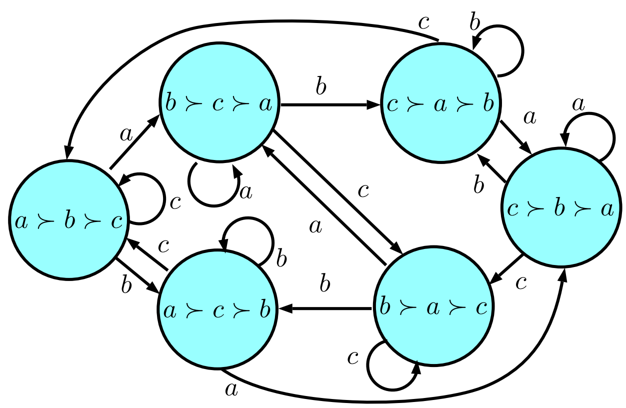


Figure 6.4: A rotational preference with 3 alternatives.

Stochastic Rotational Preferences: A stochastic preference is rotational when the probability of transition to the deterministic rotational preference is higher than all other preferences. Formally, for agent i, the transition model is defined as $T_i(\succ_i^{t+1} \mid \succ_i^t, \mu^t(i))$, such that

$$T_i(\varphi(\succ_i^t, \mu^t(i))| \succ_i^t, \mu^t(i)) > T_i(\succ_i^{t+1}|\succ_i^t, \mu^t(i)), \forall \succ_i^{t+1} \in \mathcal{P},$$

where $\succ_i^{t+1} \neq \varphi(\succ_i^t, \mu^t(i))$. With stochastic preferences, at each time step agents are required to report their preference to the center.

When preferences are rotational, after the realization of each matching decision, agents' preferences evolve such that an allocated item becomes the least desired item. Therefore, preferences diverge after each matching decision, *i.e.*, competition over similar items decreases over time, and preferences get aligned such that agents rank items differently.

Example 6.1. Suppose there are three agents with identical preferences as shown in Table 6.1a. The sequential RSD mechanism assigns equal probabilities to all $n!$ Pareto efficient allocations. Suppose that the realized matching decision is (a, b, c), *i.e.*, agents 1, 2, and 3 receive items a, b, and c respectively. Under rotational dynamic assumption, the preferences of agents (are more likely) to evolve to the profile as shown in Table 6.1b, and subsequently after matching (b, c, a) to the profile as shown in Table 6.1c.

This observation about how preference profiles change over time prevents strategic agents to benefit from misreporting in each step. The incentive compatibility property states that given a common probabilistic model of transitions, an agent's best response is to reveal its private preferences truthfully at each step, looking forward into the future.

$$\begin{pmatrix} \succ_1^1\colon \boxed{a} \succ b \succ c \\ \succ_2^1\colon a \succ \boxed{b} \succ c \\ \succ_3^1\colon a \succ b \succ \boxed{c} \end{pmatrix} \quad \begin{pmatrix} \succ_1^2\colon \boxed{b} \succ c \succ a \\ \succ_2^2\colon a \succ \boxed{c} \succ b \\ \succ_3^2\colon \boxed{a} \succ b \succ c \end{pmatrix} \quad \begin{pmatrix} \succ_1^3\colon c \succ a \succ b \\ \succ_2^3\colon a \succ b \succ c \\ \succ_3^3\colon b \succ c \succ a \end{pmatrix}$$

(a) Initial profile. *(b) Profile after (a, b, c).* *(c) Profile after (b, c, a).*

Table 6.1: A sample rotational profile and how it evolves in three periods after the realization of matching decisions. The realized matchings at each period are indicated by boxes.

The next theorem shows that when preferences are rotational, no agent can increase its expected utility by misreporting its preferences.

Theorem 6.8. *Given stochastic rotational preferences, sequential RSD satisfies w.p.s incentive compatibility.*

Proof. Let $\hat{\succ}^t = (\hat{\succ}_i^t, \succ_{-i}^t)$ and assume that all agents are truthful going forward after the current period. For incentive compatibility, we must show that $V_i(\succ^t, \tilde{\pi}(\succ^t)) \geq V_i(\succ^t, \tilde{\pi}((\hat{\succ}_i^t, \succ_{-i}^t)))$. Expanding this inequality we can write

$$\underbrace{\mathbb{E}_{\tilde{\pi}}[u_i(\succ^t, \mu) |\succ^t]}_{\text{A}} + \underbrace{\mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} |\succ^t, \mu) V_i(\succ^{t+1}, \tilde{\pi}) |\succ^t]}_{\text{B}} \geq$$
$$\underbrace{\mathbb{E}_{\tilde{\pi}}[u_i(\succ^t, \mu) |\hat{\succ}^t]}_{\text{C}} + \underbrace{\mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} |\hat{\succ}^t, \mu) V_i(\succ^{t+1}, \tilde{\pi}) |\hat{\succ}^t]}_{\text{D}}$$

By the dominant strategy incentive compatibility of RSD misreporting never improves the immediate outcome, and a misreport only influences the evolution of preferences if it changes the current allocation, thus we have $A > C$.

By the definition of stochastic rotational preferences for each agent i we have $T_i(\varphi(\succ_i^t, \mu(i)) | \succ^t, \mu(i)) > T_i(\varphi(\succ_i^t, \mu(i)) |\hat{\succ}^t, \mu(i))$. If agent i benefits from misreporting, for expressions B and D we must have $D > B$. Using the transition independence assumption and Lemma 6.1, we can compute the transition of preference profile by Cartesian product and write

$$\mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} |\hat{\succ}^t, \mu) V_i(\succ^{t+1}, \tilde{\pi}) |\hat{\succ}^t] > \mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} |\succ^t, \mu) V_i(\succ^{t+1}, \tilde{\pi}) |\succ^t]$$

In each priority ordering, agent i may only benefit by choosing a less desired item and changing the allocation of some other agent j such that j's outcome improves. Otherwise, since preferences are rotational, agent j's more preferred items remain the same while agent i does not immediately benefit from misreporting. Therefore, agent i does not also gain in the future periods, since both agents will compete over the similar items in the next periods.

Since after each matching, the preferences converge to less competing states (due to rotational dynamics), after agent i's strategic report its expected utility gain is as most equal to its immediate loss, that is,

$$\mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} |\hat{\succ}^t, \mu)V_i(\succ^{t+1}, \tilde{\pi})|\hat{\succ}^t] - \mathbb{E}_{\tilde{\pi}}[\sum_{\succ^{t+1}} T(\succ^{t+1} \mid \succ^t, \mu)V_i(\succ^{t+1}, \tilde{\pi})| \succ^t] \leq$$
$$\mathbb{E}_{\tilde{\pi}}[u_i(\succ^t, \mu)| \succ^t] - \mathbb{E}_{\tilde{\pi}}[u_i(\succ^t, \mu)|\hat{\succ}^t]$$

The immediate expected utility for agent i decreases as agent i's misreport influences more agents' preference dynamics. In addition, for agent i, the difference between expected utility after truthfulness and misreporting decreases after each rotational update. This implies that agent i's future gain cannot exceed its immediate loss, proving that $V_i(\succ^t, \tilde{\pi}(\succ^t)) \geq V_i(\succ^t, \tilde{\pi}((\hat{\succ}^t_i, \succ^t_{-i})))$. Therefore, sequential RSD is w.p.s EPIC when agents have stochastic rotational preferences. □

6.4 Single-Minded Preferences

In this section, we restrict ourselves to players with simple preferences called "single-minded" agents. Single-minded agents have been extensively studied in resource allocation markets such as combinatorial auction design [28]. We first focus on one-shot settings and show that even though the RSD mechanism is strategyproof, it does not always guarantee Pareto efficiency when agents are single-minded. We propose a randomized mechanism that satisfies a set of desirable properties in static and one-shot settings. Then, we study the incentive compatibility of our randomized mechanisms in dynamic settings and show how a unilateral deviation by a strategic agent would impact the social welfare.

A single-minded agent is indifferent between all objects except only one object as its top choice. Given a utility model, a single-minded agent receives a positive utility for its top choice and zero otherwise.

Definition 6.6 (Single Minded Preferences)**.** *Agent i is single minded if there exists an item $o_k \in M$ such that $\succ_i = o_k \succ o_1 \sim \ldots \sim o_{k-1} \sim o_{k+1} \sim \ldots \sim o_m$.*

An important axiomatic property in designing truthful matching mechanisms is *non-bossiness* due to Satterthwaite and Sonnenschein [122]. A mechanism is non-bossy if an agent cannot change the random allocation without changing the random allocation for itself.[2] When a mechanism allows agents to be indifferent among different assignments, then non-bossiness is a strong condition, since it requires that the allocation exactly remains the same even when some agents are indifferent between two equally good Pareto efficient allocations. Henceforth, we define a weaker notion of non-bossiness for agents with single-minded preferences.

Definition 6.7 (Weak Non-Bossiness (WNB))**.** *A mechanism is weakly non-bossy if for all $\succ \in \mathcal{P}^n$ and agent $i \in N$, for all $\hat{\succ}_i$ such that $\mathbb{E}_\pi[u_i| \succ] = \mathbb{E}_\pi[u_i|(\hat{\succ}_i, \succ_{-i})]$, for all agents $j \in N$ we have $\mathbb{E}_\pi[u_j| \succ] = \mathbb{E}_\pi[u_j|(\hat{\succ}_i, \succ_{-i})]$.*

The strategyproofness and weak non-bossiness of a mechanism together with Pareto efficiency guarantee the desirability of matching mechanisms. Strategyproofness and non-bossiness prevent any strategic behavior by agents while Pareto efficiency ensures that the outcome of a mechanism is stable and desirable by all agents.

When agents are single-minded, even though the RSD mechanism guarantees strategyproofness, it may prescribe Pareto dominated matchings due to underlying indifferences within agents' preference orderings. Moreover, under single-mindedness RSD no longer satisfies non-bossiness nor weak non-bossiness.

Theorem 6.9. *Under single-minded preferences, RSD is strategyproof but fails to guarantee Pareto efficiency and weak non-bossiness.*

Proof. We need only to show that there exists a serial dictatorship in the support of RSD that fails to satisfy Pareto efficiency and WNB.

Consider three agents with preferences as following $\succ_1 = a \succ b \sim c$, $\succ_2 = a \succ b \sim c$, and $\succ_3 = b \succ a \sim c$. Given the following priority ordering $f = (1, 2, 3)$, the serial dictatorship runs as follows: agent 1 chooses a, agent 2 is indifferent between b and c so it may choose b and agent 3 gets the remaining object c. Since agent 2 is indifferent between b and c (meaning that $u_2(\succ_2, b) = u_2(\succ_2, c)$, for all utility models), but agent 3 prefers object b to c, *i.e.* $u_3(\succ_3, b) > u_3(\succ_3, c)$. Thus, a matching that assigns c to agent 2 and b to agent 3 strictly improves agent 3's utility without making agent 1 worse off, implying that even though RSD is strategyproof, the induced matching is not Pareto efficient.

[2]Formally, a mechanism is non-bossy if for all $\succ \in \mathcal{P}^n$ and agent $i \in N$, for all $\hat{\succ}_i$ such that $\pi_i(\succ) = \pi_i((\hat{\succ}_i, \succ_{-i}))$ we have $\pi(\succ^t) = \pi((\hat{\succ}_i, \succ_{-i}))$.

For WNB, in the above serial dictatorship, agent 2's decision to pick object b (as opposed to c) does not change its utility but strictly decreases agent 3's utility. RSD is a uniform randomization over all serial dictatorship mechanisms, implying that RSD is neither Pareto efficient nor WNB. □

Inspired by the RSD mechanism, we propose a mechanism, called Random Equivalence Class Assignment (RECA) that ensures our desirable properties. Before describing the mechanism, we define equivalence classes for single-minded agents, where agents with similar top choices are assigned to the same equivalence classes.

Definition 6.8. *Given a set of agents N, an equivalence class for object $x \in M$ is defined as follows*

$$C_x = \{i \in N : top(\succ_i) = x\} \tag{6.5}$$

Given the definition of equivalence classes, Algorithm 4 shows the steps of our RECA mechanism. The RECA mechanism (Algorithm 4) allocates objects to agents with single-minded preferences as follows:

- Initialization: the mechanism creates equivalent classes based on the objects.
- Each agent is assigned to an equivalence class that corresponds to its top choice item.
- For each equivalence class, the mechanism assigns the object associated to that equivalence class to a randomly chosen agent from the set of agents in that equivalence class.
- While there are unassigned objects, the mechanism randomly assigns an object from the set of remaining objects to an agent from the set of remaining agents.

We let π be the random matching prescribed by RECA where $\pi(\succ)$ denotes the fractional probabilities before the realization of the matching decision and $\pi_{i,j}(\succ)$ denotes the probability that agent i receives object j under RECA.

We show that our algorithm for matching objects to a set of agents with single-minded preferences satisfies strategyproofness, Pareto efficiency, and weak non-bossiness.

Theorem 6.10. *Random Equivalence Class Assignment (RECA) is strategyproof, Pareto efficient, and weakly non-bossy.*

Algorithm 4: Random Equivalence Class Assignment (RECA)

Input: A preference profile $\succ$ of single-minded agents
Output: A matching μ according to reported preferences $\succ$

```
// Setting up equivalence classes
A ← ∅ ;                          // Set of assigned agents, initially empty.
B ← ∅ ;                          // Set of assigned objects, initially empty.
foreach (x ∈ M) do
    C_x ← {} ;
foreach (i ∈ N) do
    C_top(≻_i) ← C_top(≻_i) ∪ {i} ;
foreach ( Equivalence class C_x) do
    i ← randomly choose an agent from C_x ;
    μ(i) = x ;                   // Assign object x to the randomly chosen agent.
    A ← A ∪ {i} ;                // Add agent to the set of assigned agents.
    B ← B ∪ {x} ;                // Add object to the set of assigned objects.
foreach ( object x ∈ M \ B) do
    i ← randomly choose an agent from N \ A ;
    μ(i) = x ;
    A ← A ∪ {i} ;                // Add agent to the set of assigned agents.
    B ← B ∪ {x} ;                // Add object to the set of assigned objects.
return μ
```

Proof. **Strategyproofness**: since agents are single-minded, they are indifferent between all objects that are not ranked first. Consider a case where agent i misreports its top object, and thus, $top(\succ_i) \neq top(\hat{\succ}_i)$. If no other agent has the same top object, then agent i is in $C_{top(\hat{\succ}_i)}$ and receives that object with certainty. By Algorithm 4 it's clear that $\mathbb{E}_\pi[u_i|\succ] = \mathbb{E}_\pi[u_i|(\hat{\succ}_i, \succ_{-i})]$. If there exists some other agent j with $top(\succ_j) = top(\hat{\succ}_i)$, but no other agent's top choice is $top(\succ_i)$, *i.e.* no other agent wants agent i's top choice according to its truthful preference, then agent i has some probability of receiving its top choice. However, since no one else wants agent i's top choice then $C_{top(\succ_i)}$ consists of only agent i, which means that agent i would receive $top(\succ_i)$ with certainty under a truthful report. This shows that under no condition a strategic agent could benefit from misreporting.

Pareto efficiency: Let μ denote the assignment returned by RECA, and for simplicity let $u_i(\mu(i)) = u_i(\succ_i, \mu(i))$ denote agent i's utility for $\mu(i)$. Suppose that RECA is not Pareto efficient. This means that there are at least two agents i and j such that $u_i(\mu(j)) > u_i(\mu(i))$ and $u_j(\mu(i)) \geq u_j(\mu(j))$.

The inequality $u_j(\mu(i)) \geq u_j(\mu(j))$ implies that either agent j is indifferent between $\mu(i) \sim \mu(j)$ or $\mu(i) \succ_j \mu(j)$. If $\mu(i) \succ_j \mu(j)$ then changing the assignment of agent j with i would make agent j strictly worse off, which contradicts the assumption. Let's assume that agent j is indifferent between $\mu(i)$ and $\mu(j)$, *i.e.* $u_j(\mu(i)) = u_j(\mu(j))$. Thus, it must be the case that $top(\succ_j) \neq \mu(i) \neq \mu(j)$.

The strict inequality in $u_i(\mu(j)) > u_i(\mu(i))$ implies that $\mu(j) \succ_i \mu(i)$, and thus by single-mindedness of preferences it must be the case that $\mu(j) = top(\succ_i)$. Therefore, we can conclude that agent i belongs to the equivalence class of $C_{\mu(j)}$. Since $\mu(i) \neq \mu(j)$ by definition, it's clear that Algorithm 4 has not randomly selected agent i from $C_{\mu(j)}$, and agent i was assigned an object from the set of unassigned objects. Thus, $\mu(j)$ must be assigned to an agent with k with $top(\succ_k) = \mu(j)$ who strictly prefers $\mu(j)$ to $\mu(i)$. However, we know that $\mu(j)$ is assigned to j and we showed that $u_j(\mu(i)) = u_j(\mu(j))$, which contradicts the assumption.

Weak non-bossiness: The proof follows from the steps in Algorithm 4. A mechanism is weakly bossy, if an agent reports a preference $\hat{\succ}_i$ such that its expected utility remains the same $\mathbb{E}_\pi[u_i|\succ] = \mathbb{E}_\pi[u_i|(\hat{\succ}_i, \succ_{-i})]$ while at least one another agent j's expected utility is $\mathbb{E}_\pi[u_j|\succ] > \mathbb{E}_\pi[u_j|(\hat{\succ}_i, \succ_{-i})]$.

If agent i reports a preference such that $top(\hat{\succ}_i) \neq top(\succ_i)$, then by Algorithm 4 it loses it chance of getting selected from $C_{top(\succ_i)}$, and thus $\mathbb{E}_\pi[u_i|\succ] < \mathbb{E}_\pi[u_i|(\hat{\succ}_i, \succ_{-i})]$. For all other objects, either the object is a top choice of agent j which would be randomly assigned to an agent from $C_{top(\succ_j)}$, or the object is not a top choice of agent j, implying that $\mathbb{E}_\pi[u_j|\succ] = \mathbb{E}_\pi[u_j|(\hat{\succ}_i, \succ_{-i})]$. Thus, no agent can reduce the expected utility of another agent while receiving the same expected utility, which shows that RECA is weakly non-bossy. □

Computing RSD probabilities for general preferences is #P-complete [16], because the associated counting problem is intractable and requires $n!$ deterministic matchings to be computed. In the next theorem, we show that although computing RSD probabilities is intractable, computing RECA fractional probabilities can be done in polynomial time.

Theorem 6.11. *Given single-minded preferences, computing RECA probabilities can be done in polynomial time.*

Proof. It is sufficient to show that the probability that an agent receives each of the objects can be written in a closed form. Let c denote the number of non-empty equivalence classes.

Thus, given m objects, according to Algorithm 4 the number of objects that are ranked first by no agent is $m - c'$.

Let $\pi_{i,j} = \pi_{i,j}(\succ)$ denote the probability that agent i receives object j under RECA. For each agent, the probability of receiving the objects in M according to RECA is:

- In each equivalence class, the mechanism assigns the associated object to a randomly chosen agent from the set of agents in that class. Thus, $\pi_{i,top(\succ_i)} = \frac{1}{|C_{top(\succ_i)}|}$.
- All objects that are ranked first by at least one agent get assigned to a member of their equivalence classes. Therefore, for $m - c'$ remaining objects, the probability that agent i receives object j', is $\pi_{i,j'} = \frac{(1-\pi_{i,top(\succ_i)})}{m-c'}$.
- Lastly, no agent can have a chance to receive an object that is not in its own equivalence class, but it is ranked first by another agent. For each object j which is ranked first by at least one agent, assign probability zero to all agents that $top(\succ_i) \neq j$, *i.e.*, for each j where $\exists i' \in N$ such that $j = top(\succ_{i'})$, $\pi_{i,j} = 0$ for all i such that $j \neq top(\succ_i)$.

Each of the above steps can be done in at most $O(n^2)$. Thus, RECA probabilities for single-minded agents can be computed in polynomial time, which completes the proof. □

6.4.1 Dynamic Single-Minded Preferences

Now we are ready to extend our strategyproof, Pareto efficient, and WNB mechanism to sequential matching problems with dynamic single-minded preferences. In dynamic settings, a single-minded agent is interested in a single specific item at each period, with some positive cardinal value if its top item is assigned, and is indifferent about all the other items. This is different from our initial setting as it permits a special case of indifference between all objects except the top choice. In dynamic settings, a single-minded agent's top choice is time variant, that is, in each period there exists an item $o_k \in M$ such that $\succ_i^t = o_k \succ o_1 \sim \ldots \sim o_{k-1} \sim o_{k+1} \sim \ldots \sim o_m$.

The next lemma states that, fixing the preference of all agents except agent i, if there is a profile in which agent i's expected utility improves, then the (ex ante) social welfare of all agents also improves.

Lemma 6.3. *Consider the RECA mechanism, π, and single-minded agents. Given any two preference profiles $\succ^t$ and $\hat{\succ}^t$ such that $\hat{\succ}^t = (\hat{\succ}^t_i, \succ^t_{-i})$, if agent i strictly prefers the matching $\pi(\succ^t)$ to $\pi(\hat{\succ}^t)$, that is $\mathbb{E}_\pi[u_i| \succ^t] > \mathbb{E}_\pi[u_i|\hat{\succ}^t]$, then*

$$\sum_{i\in N} \mathbb{E}_\pi[u_i| \succ^t] \geq \sum_{i\in N} \mathbb{E}_\pi[u_i|\hat{\succ}^t]$$

Proof. Note that we are focusing on two different preference profiles where all agents are truthful (including agent i). Fixing the preferences of other agents $\succ^t_{-i}$, agent i's preference must be $\hat{\succ}^t_i \neq \succ^t_i$ in profile $\hat{\succ}^t$. Clearly a change within those objects that are not ranked first would not make an agent strictly better off. Thus, given any two preference profiles, $\succ^t$ and $\hat{\succ}^t$, if an agent is strictly better off under $\succ^t$ compared to $\hat{\succ}^t$, it has higher probability of receiving its top choice. Thus, by strictness of the inequality $\mathbb{E}_\pi[u_i| \succ^t] > \mathbb{E}_\pi[u_i|\hat{\succ}^t]$, we know that agent i has higher probability of receiving its top choice under $\succ^t$.

If $\succ^t_i \neq \hat{\succ}^t_i$ then $|C_{top(\succ^t_i)}| < |C_{top(\hat{\succ}^t_i)}|$, otherwise if $|C_{top(\succ^t_i)}| \geq |C_{top(\hat{\succ}^t_i)}|$ then there are more than or equal agents in the equivalence class that agent i belongs to, and thus, agent i will not be better off. Since agent i's expected utility improves, then for all agents j such that $top(\succ^t_j) = top(\succ^t_i)$, we have $\sum_j \mathbb{E}_\pi[u_j| \succ^t] > \sum_j \mathbb{E}_\pi[u_j|\hat{\succ}^t]$ because $|C_{top(\succ^t_j)}|$ is smaller than $|C_{top(\hat{\succ}^t_j)}|$. Note that even though agent j's expected utility may decrease, the expected sum of utilities does not decrease, and in each equivalence class C_x, there always exists $|C_x|$ agents with exactly $\frac{1}{|C_x|}$ chance of receiving the top choice. For all other agents j' such that $top(\succ^t_{j'}) \neq top(\succ^t_i)$, RECA chooses an agent from each equivalent class uniformly at random, there are $|C_{top(\succ^t_{j'})}|$ agents with $\frac{1}{|C_{top(\succ^t_{j'})}|}$ chance of receiving the top choice. Since the preferences are fixed for all agents, for all such equivalence classes C_x, the sum of utilities will be $|C_x|\frac{1}{|C_x|}$, which is greater than or equal to the expected sum under $\hat{\succ}^t$. Adding this to the sum of utilities of agents in equivalence class $C_{top(\succ^t_i)}$ results in $\sum_{i\in N} \mathbb{E}_\pi[u_i| \succ^t] \geq \sum_{i\in N} \mathbb{E}_\pi[u_i|\hat{\succ}^t]$, implying that the sum of expected utilities does not decrease if at least one agent's allocation strictly improves. □

An interesting consequence of the above lemma is that if a single-minded agent benefits from misreporting its preferences, changing the evolution of preference dynamics would improve the overall expected sum of utilities for all agents. In other words, an agent's misreport only results in a more efficient matching policy. That is, even if manipulation in a sequence of RECA assignments (sequential RECA) is possible, the manipulation will always benefit all agents.

Theorem 6.12. *Given sequential RECA, π, with single-minded agents, for any agent i with misreport $\hat{\succ}_i^t$ where $\hat{\succ}^t = (\hat{\succ}_i^t, \succ_{-i}^t)$, if $V_i(\succ^t, \pi(\hat{\succ}^t)) > V_i(\succ^t, \pi(\succ^t))$ when all others are truthful we have*

$$\sum_{i \in N} V_i(\succ^t, \pi(\hat{\succ}^t)) > \sum_{i \in N} V_i(\succ^t, \pi(\succ^t)) \tag{6.6}$$

Proof. It is easy to see that under sequential RECA even when agents are single-minded, a strategic agent can benefit from misreporting when $n \geq 3$. Consider agent i with misreport $\hat{\succ}_i^t$ at time t such that $V_i(\succ^t, \pi(\hat{\succ}^t)) > V_i(\succ^t, \pi(\succ^t))$. For all $j \in N$ if $top(\succ_j^t) \neq top(\hat{\succ}_i^t)$ then either $|C_{top(\succ_j^t)}|$ is decreased when $top(\succ_j^t) = top(\succ_i^t)$ or agent i's misreport does not change other agents' allocations because of the independence of transitions. Thus, according to RECA these agents' expected utilities may improve, that is, $\mathbb{E}_\pi[u_j|\hat{\succ}^t] \geq \mathbb{E}_\pi[u_j| \succ^t]$, and the misreport does not change the evolution of preferences for these agents, thus, $V_j(\succ^t, \pi(\hat{\succ}^t)) > V_j(\succ^t, \pi(\succ^t))$.

We need to focus on all $j' \in N$ such that $top(\succ_{j'}^t) = top(\hat{\succ}_i^t)$. If agent i successfully gains from manipulating the allocation then we must have $V_i(\succ^t, \pi(\hat{\succ}^t)) > V_i(\succ^t, \pi(\succ^t))$, which can be written as

$$\mathbb{E}_\pi[u_i(\succ^t, \mu)|\hat{\succ}^t] + \mathbb{E}_\pi[\sum_{\succ^{t+1}} T(\succ^{t+1} |\hat{\succ}^t, \mu) V_i(\succ^{t+1}, \pi)|\hat{\succ}^t] >$$
$$\mathbb{E}_\pi[u_i(\succ^t, \mu)| \succ^t] + \mathbb{E}_\pi[\sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) V_i(\succ^{t+1}, \pi)| \succ^t]$$

However, by strategyproofness of RECA in one-shot settings, the immediate expected utility of misreporting is strictly less than being truthful $\mathbb{E}_\pi[u_i(\succ^t, \mu)|\hat{\succ}^t] < \mathbb{E}_\pi[u_i(\succ^t, \mu)| \succ^t$]. This implies that agent i must have gained in the future periods. For all agents j' such that $top(\succ_{j'}^t) = top(\hat{\succ}_i^t)$, the $|C_{top(\succ_{j'}^t)}|$ now includes agent i, thus the sum over the immediate expected utility is

$$\sum_{j' \in C_{top(\succ_{j'}^t)}} \mathbb{E}_\pi[u_{j'}|\succ^t] = |C_{top(\succ_{j'}^t)}| \frac{1}{|C_{top(\succ_{j'}^t)}| + 1}$$
$$\sum_{j' \in C_{top(\succ_{j'}^t)}} \mathbb{E}_\pi[u_{j'}|\succ^t] \geq \frac{1}{2}$$

Therefore, agent i's misreport decreases the social welfare of all other agents by at most $1 - \frac{1}{2}$, which is exactly equal to the utility loss by agent i. However, by $V_i(\succ^t, \pi(\hat{\succ}^t)) >$

$V_i(\succ^t, \pi(\succ^t))$, agent i's allocation has improved in the future by at least the same amount. Thus, for the future expected value we should have

$$\mathbb{E}_\pi[\sum_{\succ^{t+1}} T(\succ^{t+1} | \hat{\succ}^t, \mu) V_i(\succ^{t+1}, \pi) | \hat{\succ}^t] > \mathbb{E}_\pi[\sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) V_i(\succ^{t+1}, \pi) | \succ^t]$$

By Lemma 6.1 and since the expectation is over all matchings, we can write the future transitions as

$$\sum_\mu \sum_{\succ^{t+1}} T(\succ^{t+1} | \hat{\succ}^t, \mu) \pi(\mu | \hat{\succ}^t) = \sum_\mu \sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) \pi(\mu | \hat{\succ}^t) \tag{6.7}$$

Thus, for all future utilities we have

$$\sum_\mu \sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) \pi(\mu | \hat{\succ}^t) V_i(\succ^{t+1}, \pi) > \sum_\mu \sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) \pi(\mu | \succ^t) V_i(\succ^{t+1}, \pi)$$

By Lemma 6.3, fixing the preferences of all agents $j' \in N$, for preference profile $\succ^{t+1}$ at which agent i's matching is strictly improved, the sum of expected utilities for all agents gets improved. Agent i's misreport at time t can change the allocation of *only* one another agent in equivalence class $C_{top(\hat{\succ}_i^t)}$. We already know that for all j where $top(\succ_j^t) \neq top(\hat{\succ}_i^t)$ the allocation, and consequently transitions to the next profile, does not change.

Let $C_a^{t'}, C_b^{t'}, \ldots, C_m^{t'}$ denote the equivalence classes in the next periods $t' > t$. Given RECA, under truthful reporting the sum of utilities is $\sum_{x \in M} |C_x^{t'}| \frac{1}{|C_x^{t'}|}$. Let j' be an agent in $C_{top(\hat{\succ}_i^t)}$ where its allocation changes after misreport and its preference evolve such that $top(\succ_{j'}^{t+1}) \neq top(\succ_i^{t+1})$. Thus, there exists an equivalence class $C_x^{t'} = C_{top(\succ_{j'}^{t+1})}$ with one extra agent j' in it. But since agent i's allocation improves, it cannot be in this equivalence class, that is, $top(\succ_{j'}^{t+1}) \neq top(\succ_i^{t+1})$. If no other agent is in $C_{top(\succ_{j'}^{t+1})}$, then there exists a class $C_y^{t'}$ with one less member, and in each future period the expected sum of utilities for all agents improves. Otherwise, the sum of utilities for all agent in $C_y^{t'}$ is $|C_y^{t'}| \frac{1}{|C_y^{t'}|}$. Thus, the sum of utilities for all agents except i remains the same. Given the expected transition according to Equation 6.7, the expected value for all agents is

$$\sum_{i \in N} \sum_\mu \sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) \pi(\mu | \hat{\succ}^t) V_i(\succ^{t+1}, \pi) >$$
$$\sum_{i \in N} \sum_\mu \sum_{\succ^{t+1}} T(\succ^{t+1} | \succ^t, \mu) \pi(\mu | \succ^t) V_i(\succ^{t+1}, \pi)$$

This similarly holds for all periods $t' > t$. Therefore, agent i's misreport has improved the evolution of matching decisions, and thus, the social welfare for all agents improves. □

6.5 Concluding Remarks

We considered matching with evolving preferences and formulated this as a planning problem by leveraging MDP models. In this framework, we showed that the sequential RSD mechanism does not satisfy within-period ex post incentive compatibility even when agents are endowed with linear positional scoring utilities.

In many real-life problems, agents often have structured preferences and these preferences do not radically change from one period to the next. Motivated by this observation, we studied subclasses of matching problems with restricted preference *dynamics*. By examining some additional mild restrictions on the *dynamics* of the agents' preferences, we were able to create interesting subclasses for which a simple sequential RSD mechanism satisfies the notion of within-period ex post incentive compatibility. Moreover, we investigated the problems with single-minded preferences and showed that RSD matchings do not necessarily satisfy Pareto efficiency even in one-shot settings. To overcome this issue, we proposed a randomized matching mechanism that satisfies strategyproofness, Pareto efficiency, and weak non-bossiness in one-shot settings. Even though our mechanism for assigning objects to single-minded agents in dynamic settings is manipulable, we showed that any unilateral deviation would benefit all participating agents.

The dynamic matching problem gives rise to several intriguing questions. Although an optimal policy does not guarantee truthfulness, finding and characterizing truthful matching policies (from the space of all randomized policies) that maximize the social welfare is still an open problem. In static settings, manipulating random matching mechanisms (for example, Probablistic Serial Rule [30]) is shown to be NP-hard [20]. One interesting direction is to study the complexity of manipulating the matching policies in dynamic environments, and to characterize players' strategic behaviors under various populations and utility models. Finally, it would be interesting to further study the incentive and efficiency properties of dynamic matching policies on the full preference domain with indifferences [87] and also in large markets [97].

Chapter 7

Conclusions

The overarching theme of this thesis was investigating ways to efficiently and fairly distribute a set of discrete resources to a set of self-interested agents, in the absence of monetary transfers. To accomplish this goal, we focused on one-shot matching decisions wherein players compete over scarce available resources in multiagent settings as well as mechanisms and strategies that incentivize truthful behavior of agents in repeated matching problems with uncertainties about preferences. More specifically, we addressed the following questions:

- What deterministic or randomized mechanisms are appropriate to assign multiple objects to agents based on a given quota?
- Which randomized matching mechanism should a social planner adopt in practice to ensure a desirable level of social welfare under various risk attitudes?
- Do the desirable properties of traditional matching mechanisms carry over to dynamic settings with repeated allocations?
- Can we design truthful mechanisms for dynamic matching problems?

Our findings shed light onto challenges of ensuring desired economic properties such as fairness and efficiency while guaranteeing strategyproofness. These findings can help mechanism designers and social planners when considering which mechanisms and approaches to deploy in competitive markets. For example, an institution may decide to adopt our proposed RSDQ mechanism in Chapter 3 to fairly assign teaching loads to faculty under a certain teaching quota while ensuring Pareto efficiency and preventing agent manipulation. In a different scenario, a social planner who wants to fairly assign a set of shifts (or tasks) repeatedly over time to a set of social workers, nursing staff, or any other groups of players should be cautious about the possibility of manipulation and may decide to adopt the Adaptive RSD mechanism that we proposed in Chapter 5.

Working in the one-shot settings, in Chapter 3, we investigated the class of deterministic matching mechanisms when there is a quota to be fulfilled. We showed that serial dictatorship mechanisms (and their sequential dictatorship counterparts) characterize the set of all possible matching mechanisms with desirable economic properties, enabling social planners to remedy the inherent unfairness in deterministic allocation mechanisms by assigning quotas according to some fairness criteria (such as seniority, priority, etc.). Moreover, we generalized our findings to randomized mechanisms for lexicographic preferences, expanding random serial dictatorship mechanisms to quota mechanisms. Our proposed mechanism satisfies our desired properties such as strategyproofness and ex post efficiency, while guaranteeing a stronger notion of fairness called envyfreeness. These findings, prove that the well-known Random Serial Dictatorship (RSD) mechanism in standard assignment settings satisfies envyfreeness when preferences are lexicographic. Random quota mechanisms provide a rich and extended class for object allocation with no restriction on the market size nor quota structure while providing envyfreeness in lexicographic domains, justifying the use of such mechanisms in many practical applications.

In Chapter 4, we provided a systematic empirical study of two seminal random matching mechanisms, namely Random Serial Dictatorship (RSD) and Probabilistic Serial Rule (PS). Our main goal was to provide better empirical insights to the theoretical findings so that mechanism designers of multiagent systems can decide which mechanism to adopt in practice. In the space of general ordinal preferences, we showed that while RSD does not theoretically guarantee stochastic dominance efficiency, in most cases RSD and PS allocations are incomparable under stochastic (or lexicographic) dominance. When instantiating the preferences with actual utility functions, PS allocations are only slightly better than RSD allocations in terms of social welfare, particularly under risk averse utilities, while in some cases RSD allocations are superior in terms of social welfare. In addition, we showed that not only PS is manipulable in most cases, this manipulation and the possible gain from manipulation become more severe when agents are risk averse, and designers need to take this into consideration.

Our empirical findings in this thesis complement the theoretical guarantees of matching mechanisms, shedding light on practical implications of deploying each of the given mechanisms. Another interesting future direction is to investigate how these results compare to settings where the objectives are expressed based on an egalitarian social welfare function [71, 127] as opposed to our utilitarian approach for social welfare.

Focusing on settings with repeated matching, in Chapter 5 we studied the incentive and fairness properties of sequential matching with dynamic ordinal preferences. We first proposed a model, based on stochastic decision processes, for analyzing and reasoning about stochastic sequences of allocations under ordinal assignment of objects to agents. We showed that in contrast to one-shot settings, a sequence of RSD-induced matchings is prone to manipulation under both history-dependent and Markovian transitions. Subsequently, we proposed a history-dependent matching mechanism that satisfies strategyproofness if preference dynamics follow a certain type of trajectory, and showed that our proposed mechanism, called Adaptive RSD, prevents harmful manipulations under a mild assumption while sustaining the local properties of the sequential RSD. Nonetheless, we showed that a strong manipulation is still possible, but a unilateral manipulation often leads to a Pareto improving sequence of decisions for all agents. Moreover, removing all the assumptions on how preferences evolve over time, we showed that Adaptive RSD is weakly gsd-strategyproof if the planning horizon is bounded.

Finally, in Chapter 6 we restricted attention to subclasses of matching problems where sequential RSD can still be used. We formulated the sequential matching problem as a Multiagent MDP where agents have linear positional scoring utilities. We showed that even under this restriction, sequential RSD is only strategyproof when there are two agents and does not preclude strategic reporting when there are more than two agents. In a number of real-life applications, agents often have structured preferences and these preferences do not radically change from one period to the next. Motivated by this observation, we studied subclasses of matching problems with restricted preference *dynamics*, and showed that when agents are myopic or when they have rotational preferences a simple sequential RSD mechanism satisfies the notion of within-period ex post incentive compatibility. Moreover, we investigated the problems with *single-minded preferences* and showed that RSD matchings do not necessarily satisfy Pareto efficiency even in one-shot settings. To overcome this issue, we proposed a randomized matching mechanism that satisfies strategyproofness, Pareto efficiency, and weak non-bossiness in one-shot settings. Even though our mechanism for assigning objects to single-minded agents in dynamic settings is manipulable, we showed that any unilateral deviation would benefit all participating agents.

7.1 Future Work

The decision problem in most multiagent settings is dynamic, rather than static, and agents' preferences may change or evolve depending on previous allocations or some uncertain events. In dynamic settings, we investigated the desirability of different matching mechanisms for resource allocation that make decisions by eliciting agents' preferences, and showed that the traditional matching mechanisms fail to prevent this type of manipulation. An interesting research direction in dynamic matching is to provide an axiomatic approach for comparing sequences of allocations by defining new axioms beyond the traditional notion of stochastic dominance in ordinal domains.

In recent years, there have been efforts in designing approximate mechanisms that partially satisfy the desirable properties such as strategyproofness and efficiency in one-shot settings [105]. An interesting future direction is study and design sequential mechanisms in dynamic settings with modifiable parameters that can approximately satisfy truthfulness, efficiency, or fairness given a set of desired parameters.

In this thesis, we studied the settings where an agent unilaterally misreports its private underlying preferences. An intriguing future direction concerns with characterizing the set of matching mechanisms for repeated allocations that globally satisfy strategyproofness when agents can learn from previous allocations and decisions and their best-response strategies is dependent on the acquired endogenous information.

Mechanism design approaches for matching in dynamic environments are important in multiagent resource allocation. Our work in this domain assumed that the population of agents is fixed over time and only their preferences may evolve. However, in most multiagent settings, agents arrive and depart stochastically over time and the outcome of a mechanism depends heavily on the reported preferences of agents as well as private information about agents' departure and arrival times [2, 8, 68, 96, 132]. An intriguing research direction is to study a more realistic multiagent setting by combining the two notions for uncertainty over the population of agents and over agents' preferences.

Finally, throughout this work we restricted our attention to settings with no positive or negative externalities. Nonetheless, in many situations agents not only possess an idiosyncratic preferences over alternatives, but they may also have preferences over the outcomes of other agents as soft or hard constraints [37, 70]. A promising future direction is to study such externalities in dynamic matching when agents' preferences change over time and develop algorithmic techniques to ensure the desirable properties of resource allocation in these multiagent settings.

Appendices

Appendix A

Numerical Results of Chapter 4

The following table shows the results of comparing RSD and PS under ordinal preferences for various combinations of agents and objects. Note that in most instances, RSD and PS do not induce the same random allocation.

			Dominance		RSD	PS manipulability		
n	**m**	**Equal**	**SD**	**LD**	**weakEnvy**	**weak**	**SD**	**LD**
2	2	100%	0%	0%	0%	0%	0%	0%
2	3	27%	18%	29%	23%	31%	31%	31%
2	4	10%	36%	60%	20%	53%	53%	53%
2	5	3%	39%	78%	16%	78%	78%	78%
2	6	1%	45%	90%	13%	87%	87%	87%
2	7	0%	46%	95%	12%	95%	95%	95%
2	8	0%	45%	96%	11%	97%	97%	97%
2	9	0%	47%	96%	11%	100%	100%	100%
2	10	0%	48%	99%	9%	99%	99%	99%
3	2	100%	0%	0%	0%	0%	0%	0%
3	3	67%	0%	0%	11%	24%	0%	0%
3	4	3%	5%	40%	47%	77%	5%	5%
3	5	0%	4%	75%	46%	96%	26%	27%
3	6	0%	6%	84%	42%	95%	53%	54%
3	7	0%	5%	90%	41%	100%	68%	69%
3	8	0%	5%	93%	39%	100%	80%	83%

Continued on the next page

			Dominance		RSD	PS manipulability		
n	m	Equal	SD	LD	weakEnvy	weak	SD	LD
3	9	0%	9%	96%	35%	100%	90%	92%
3	10	0%	7%	95%	34%	100%	94%	94%
4	2	62%	38%	38%	0%	0%	0%	0%
4	3	33%	34%	46%	21%	42%	0%	0%
4	4	21%	3%	8%	27%	72%	0%	0%
4	5	0%	0%	48%	61%	96%	1%	1%
4	6	0%	0%	76%	62%	98%	17%	18%
4	7	0%	0%	84%	62%	100%	33%	35%
4	8	0%	1%	93%	61%	99%	52%	54%
4	9	0%	1%	94%	60%	100%	65%	69%
4	10	0%	2%	95%	56%	100%	79%	85%
5	2	39%	61%	61%	0%	0%	0%	0%
5	3	8%	34%	83%	27%	66%	0%	0%
5	4	3%	19%	53%	42%	94%	0%	0%
5	5	6%	1%	7%	42%	90%	0%	0%
5	6	0%	0%	58%	69%	100%	0%	0%
5	7	0%	0%	84%	71%	100%	4%	4%
5	8	0%	0%	91%	71%	100%	18%	18%
5	9	0%	0%	94%	71%	100%	32%	36%
5	10	0%	0%	97%	70%	100%	49%	55%
6	2	21%	79%	79%	0%	0%	0%	0%
6	3	2%	71%	96%	31%	59%	0%	0%
6	4	0%	22%	88%	52%	90%	0%	0%
6	5	0%	9%	46%	59%	98%	0%	0%
6	6	3%	1%	7%	54%	96%	0%	0%
6	7	0%	0%	62%	74%	100%	0%	0%
6	8	0%	0%	89%	74%	100%	1%	1%
6	9	0%	0%	95%	75%	100%	8%	9%
6	10	0%	0%	97%	75%	100%	23%	25%
7	2	12%	88%	88%	0%	0%	0%	0%
7	3	1%	64%	99%	33%	83%	0%	0%
7	4	0%	26%	97%	57%	99%	0%	0%
7	5	0%	8%	87%	66%	100%	0%	0%

Continued on the next page

			Dominance		RSD	PS manipulability		
n	**m**	**Equal**	**SD**	**LD**	**weakEnvy**	**weak**	**SD**	**LD**
7	6	0%	2%	41%	69%	100%	0%	0%
7	7	1%	1%	6%	61%	99%	0%	0%
7	8	0%	0%	71%	79%	100%	0%	0%
7	9	0%	0%	93%	79%	100%	0%	0%
7	10	0%	0%	96%	78%	100%	5%	6%
8	2	8%	92%	92%	0%	0%	0%	0%
8	3	0%	63%	100%	34%	76%	0%	0%
8	4	0%	33%	99%	60%	95%	0%	0%
8	5	0%	10%	97%	70%	100%	0%	0%
8	6	0%	4%	83%	74%	100%	0%	0%
8	7	0%	1%	29%	74%	100%	0%	0%
8	8	0%	0%	5%	69%	99%	0%	0%
8	9	0%	0%	70%	81%	100%	0%	0%
8	10	0%	0%	93%	82%	100%	0%	0%
9	2	3%	97%	97%	0%	0%	0%	0%
9	3	0%	76%	100%	35%	70%	0%	0%
9	4	0%	33%	100%	62%	100%	0%	0%
9	5	0%	19%	99%	72%	100%	0%	0%
9	6	0%	6%	98%	76%	100%	0%	0%
9	7	0%	2%	78%	78%	100%	0%	0%
9	8	0%	0%	26%	78%	100%	0%	0%
9	9	0%	0%	4%	71%	100%	0%	0%
9	10	0%	0%	69%	84%	100%	0%	0%
10	2	2%	99%	99%	0%	0%	0%	0%
10	3	0%	70%	100%	37%	79%	0%	0%
10	4	0%	46%	100%	63%	98%	0%	0%
10	5	0%	17%	100%	73%	97%	0%	0%
10	6	0%	10%	99%	77%	100%	0%	0%
10	7	0%	2%	95%	79%	100%	0%	0%
10	8	0%	1%	77%	80%	100%	0%	0%
10	9	0%	0%	21%	79%	100%	0%	0%
10	10	0%	0%	4%	73%	100%	0%	0%

Table A.1: Experimental results over the space of preference profiles.

Appendix B

Proof of Theorem 5.4

Proof. The proof follows by constructing an instance of the matching process where an agent's strategic report causes the evolution of preferences (and subsequently the trajectory of decisions) to result in the best expected sequence of decisions for a strategic agent. There are three steps in the proof: (i) construct a best possible sequence of outcomes for agent i under her misreport, (ii) construct the worst possible sequence under the truthful report, (iii) use Definition 5.11 to find the minimum required horizon for achieving strong manipulation.

Let π denote a sequential RSD policy, and T be a history-dependent transition function. Let $\hat{\succ}_i^t$ be agent i's best strategic report at time t, given its true preference $\succ_i^t$. Let $o_1, \ldots, o_n$ denote a rank-ordered list of alternatives at each time step, that is, o_1 represents the top choice alternative according to $\succ_i^t$.

In each matching period, let $\overset{*}{\succ}$ denote a preference profile where agent i receives its top choice, and $\overset{\circ}{\succ}$ be a profile with the worst possible outcome for agent i. By Lemma 5.1, the best possible allocation for an agent is to receive its top choice with certainty, and the worst possible allocation is to receive every object with the probability of $\frac{1}{n}$. Note that under sequential RSD policy, any other randomized outcome would be an improvement over the worst-case matching. By local strategyproofness of sequential RSD, agent i does not receive any immediate gain by misreporting. Without loss of generality, we can shift the time and assume that the matching process starts at time $t = 1$. For simplicity, let us assume that agent i receives its last choice after misreporting at t, that is, $\mu^t(i) = \tilde{\mu} = o_n$. Starting at time t, there is no information in history so we can write

$$T(h^{t+1}|\hat{h}^t, \mu^t) = T(\succ^{t+1} |\hat{\succ}^t, \mu^t)$$

Constructing the best possible outcome for agent i from time t onwards, we can assume that for all future time periods we have

$$T(h^{t+2}|h^{t+1}, \mu^{t+1}) = 1$$

where $\pi(h^{t+2})$ induces the best outcome for agent i, that is, $\pi(h^{t+2}) = \overset{*}{\mu}$ where $\overset{*}{\mu}$ is a best outcome for agent i according to Lemma 5.1 where it receives its top choice with certainty. Now, we can compute the score of assigning all alternatives for λ steps:

$$\begin{aligned} W_i^\pi(\hat{h}^t) &= \mathring{\mu}^t(i) + (\lambda - 1)\overset{*}{\mu}(i) \\ &= o_3 + (\lambda - 1)o_1 \end{aligned} \tag{B.1}$$

We now consider the worst sequence of outcomes for agent i when reporting truthfully at t. Let $\mathring{\mu} = \pi(\succ^t)$ be the worst-case random matching at t. Since the transition function is history-dependent, it may be the case that agents always transition to a worst-case profile, and hence, receive the worst-case random allocation. In other words, for all matchings $\mu \in \mathcal{M}$ and times $t' > t$, the system transitions to $h^{t'}$ where $\succ^{t'}$ is a profile with worst-case outcome for agent i:

$$T(h^{t'}|h^t, \mu) = 1$$

where $\pi(h^{t'}) = \mathring{\mu}$. Thus, for λ planning horizon we can write

$$\begin{aligned} W_i^\pi(h^t) &= \mathring{\mu}^1(i) + \frac{2}{3}(\lambda - 1)\mathring{\mu}(i) + \frac{2}{3}(\lambda - 1)\overset{*}{\mu}(i) \\ &= \frac{1}{3}(o_1 + o_2 + o_3) + \frac{2}{3}(\lambda - 1)(\frac{1}{3}(o_1 + o_2 + o_3)) + \frac{1}{3}(\lambda - 1)o_1 \end{aligned} \tag{B.2}$$

To prove the possibility of manipulation, we need to show that for some ranking position ℓ the following holds:

$$\exists\ o_\ell \in M,\ W_i^\pi(\hat{h}^t, o_\ell) > W_i^\pi(h^t, o_\ell) \tag{B.3}$$

Replacing $W_i^\pi(h^t)$ with equations solved for truthful report (Equation B.2) and $W_i^\pi(\hat{h}^t)$ with equations solved for the strategic report (Equation B.1), we have

$$\begin{aligned} \tilde{\mu} + (\lambda - 1)\overset{*}{\mu}(i) &> \mathring{\mu}^1(i) + \frac{2}{3}(\lambda - 1)\mathring{\mu}(i) + \frac{2}{3}(\lambda - 1)\overset{*}{\mu}(i) \\ \tilde{\mu} + \frac{2}{3}(\lambda - 1)\overset{*}{\mu}(i) &> \frac{2(\lambda - 1) + 3}{3}\mathring{\mu} \end{aligned}$$

For simplicity, we replace the term $\lambda - 1$ with λ'. Now, writing random matchings according to Equation 5.3 over alternatives and using Lemma 5.1 we have

$$o_3 + \frac{2}{3}\lambda' o_1 > \frac{2\lambda' + 3}{9} o_1 + \frac{2\lambda' + 3}{9} o_2 + \frac{2\lambda' + 3}{9} o_3$$

For $\ell = 1$ we have

$$\frac{2}{3}\lambda' o_1 > \frac{2\lambda' + 3}{9} o_1$$

Thus, strong manipulation *may* be possible if $\lambda' > \frac{3}{4}$, and thus, $\lambda > \frac{7}{4}$. Because of the discrete time assumption, we can write $\lambda \geq 2$. To ensure strong manipulability, we similarly need to show that this inequality holds (at least weakly) for other ranks. For $\ell = 2$ we write

$$\frac{2}{3}\lambda' \geq \frac{2\lambda' + 3}{9} + \frac{2\lambda' + 3}{9}$$

which results in $\lambda' \geq 3$. Thus, sequential RSD is strongly manipulable when $\lambda \geq 4$. It is easy to see that sequential RSD is only weakly manipulable for $1 < \lambda \leq 3$. This concludes our proof. □

Appendix C

Proof of Theorem 5.5

Proof. Let π be a sequential RSD mechanism, and let $\succ^t = (\succ_i^t, \succ_{-i}^t)$ and $\hat{\succ}^t = (\hat{\succ}_i^t, \succ_{-i}^t)$ be the profiles where agent i reports truthfully and non-truthfully respectively. We construct two cases, one where agent i receives a worst sequence under a truthful report and one where agent i receives a best possible (expected) sequence of allocations under a misreport. Let $\overset{*}{\succ}$ denote a state (preference profile) where agent i receives its top choice, and $\overset{\circ}{\succ}$ be a profile with the worst possible random matching for agent i.

Step (*i*): Agent i's strategic report will yield an immediate random allocation of $\pi(\hat{\succ}^t)$ where by local strategyproofness $\pi_i(\succ^t) \succeq_i^t \pi_i(\hat{\succ}^t)$. By Lemma 5.2, agent i can only affect the decision trajectory by changing the current random allocation. Assuming the best possible scenario, we assume that the misreport takes agent i to the best state $\overset{*}{\succ}$, that is,

$$\sum_{\mu \in \mathcal{M}} T(\overset{*}{\succ} | \hat{\succ}^{t'}, \mu) \pi(\mu | \hat{\succ}^{t'}) = 1$$

for all $t < t' \leq \lambda$, that is, all matchings prescribed by π will take the system to state $\overset{*}{\succ}$. Thus, we have

$$\begin{aligned} W_i^\pi(\hat{\succ}^t) &= \pi_i(\hat{\succ}^t) + \sum_{t'=2}^{\lambda} \sum_{\mu \in \mathcal{M}} T(\overset{*}{\succ} | \hat{\succ}^{t'}, \mu) \pi(\mu | \hat{\succ}^{t'}) \pi_i(\overset{*}{\succ}) \\ &= \pi_i(\hat{\succ}^t) + \sum_{t'=2}^{\lambda} \pi_i(\overset{*}{\succ}) \\ &= \pi_i(\hat{\succ}^t) + (\lambda - 1)\pi_i(\overset{*}{\succ}) \end{aligned}$$

Since $\pi_i(\overset{*}{\succ})$ assigns agent i's top choice with certainty (best case), we can rewrite the above equation in the following closed form:

$$\begin{aligned} W_i^{\pi}(\hat{\succ}^t) &= \pi_i(\hat{\succ}^t) + (\lambda - 1)\pi_i(\overset{*}{\succ}) \\ &= \pi_i(\hat{\succ}^t) + (\lambda - 1)\ o_1 \end{aligned}$$

Step *(ii)*: This step follows by constructing the worst-case sequence of outcomes under agent i's truthful report. Let $\overset{\circ}{\succ}$ be a profile with the worst possible outcome for agent i under truthfulness. Then, according to Lemma 5.1, $\pi_i(\succ^t) = \frac{1}{3}(o_1 + o_2 + o_3)$. Considering the worst-case scenario for agent i, we assume that every decision at t will keep agent i in state $\overset{\circ}{\succ}$ for all future steps $t' > t$.

$$\sum_{\mu \in \mathcal{M}} T(\overset{\circ}{\succ} |\succ^{t'}, \mu)\pi(\mu| \succ^{t'}) = 1$$

By Lemma 5.3, there exists a subset of Pareto efficient matchings that are in $\pi(\succ^t)$ as well as $\pi(\hat{\succ}^t)$. We denote these two sets as $\mathcal{M}^{\succ^t} \subset \mathcal{M}$ and $\mathcal{M}^{\hat{\succ}^t} \subset \mathcal{M}$. Thus, the next state can be written as

$$\underbrace{\sum_{\mu \in \mathcal{M}^{\succ^t}} T(\overset{\circ}{\succ} |\succ^t, \mu)\pi(\mu| \succ^t)W_i^{\pi}(\overset{\circ}{\succ}) +}_{\text{worst}} \underbrace{\sum_{\mu \in \mathcal{M}^{\hat{\succ}^t}} T(\overset{*}{\succ} |\hat{\succ}^t, \mu)\pi(\mu| \succ^t)W_i^{\pi}(\overset{*}{\succ})}_{\text{best}}$$

Using Lemma 5.2, we can further simplify the second statement and write

$$\sum_{\mu \in \mathcal{M}^{\succ^t}} T(\overset{\circ}{\succ} |\succ^t, \mu)\pi(\mu| \succ^t)W_i^{\pi}(\overset{\circ}{\succ}) + \sum_{\mu \in \mathcal{M}^{\hat{\succ}^t}} T(\overset{*}{\succ} |\succ^t, \mu)\pi(\mu| \succ^t)W_i^{\pi}(\overset{*}{\succ})$$

With three agents and assuming the worst-case scenario, by Lemma 5.1 and Lemma 5.3 the probability of transitioning to the next state can be written as

$$\frac{2}{3}\pi(\overset{\circ}{\succ}) + \frac{1}{3}\pi(\overset{*}{\succ})$$

Thus, by Lemma 5.3 there is $\frac{1}{3}$ probability of choosing a matching that guides the trajectory to the best scenario, from where the evolution of preferences continues as in Step (i). This is because of the memorylessness property of Markovian transitions, that is, it is not important how the system ends up in a state and the next-state transitions only depend on the currents state and decision.

Because of the Markovian assumption, by forward induction, in every next state the same assumptions hold, meaning that in every future step there is a probability of transitioning to the *best* state, which takes agent i to the best scenario. Thus, we can recursively write

$$W_i^{\pi}(\succ^t) = \underbrace{\frac{2}{3}\left[\underbrace{\frac{2}{3}\Big(\underbrace{\frac{2}{3}[\underbrace{\frac{2}{3}\underbrace{\pi_i(\overset{\circ}{\succ})}_{\text{t}}+\frac{1}{3}\pi_i(\overset{*}{\succ})]}_{\text{t+1}}+\frac{1}{3}\pi_i(\overset{*}{\succ})}_{\text{t+2}}\Big)+\frac{1}{3}\pi_i(\overset{*}{\succ})}_{\text{t+3}}\right]+\frac{1}{3}\pi_i(\overset{*}{\succ})}_{\text{t+4}}$$

Recall that $\pi_i(\overset{*}{\succ}) = o_1$. For three objects and λ planning steps, we can write the above equation in the following closed forms:

$$W_i^{\pi}(\succ^t, o_1) = \frac{1}{3}\ \lambda + \frac{1}{3}\sum_{k=1}^{\lambda-1}(\frac{2}{3})^k(\lambda - k)$$

$$W_i^{\pi}(\succ^t, o_2) = \frac{1}{3}\sum_{k=1}^{\lambda}(\frac{2}{3})^{k-1}$$

$$W_i^{\pi}(\succ^t, o_3) = \frac{1}{3}\sum_{k=1}^{\lambda}(\frac{2}{3})^{k-1}$$

Note that the feasibility condition states that for any λ, the sum of above must guarantee the following:

$$\frac{1}{3}\ \lambda + \frac{1}{3}\sum_{k=1}^{\lambda-1}(\frac{2}{3})^k(\lambda - k) + \frac{1}{3}\sum_{k=1}^{\lambda}(\frac{2}{3})^{k-1} + \frac{1}{3}\sum_{k=1}^{\lambda}(\frac{2}{3})^{k-1} = \lambda.$$

Step (*iii*): For strong manipulation, using steps (i) and (ii) we need to show that

$$\forall o_\ell \in M,\ W_i^\pi(\hat{h}^t, o_\ell) \geq W_i^\pi(h^t, o_\ell)$$

with one strict inequality. For o_1, we should have

$$\pi_i(\hat{\succ}^t) + (\lambda - 1) > \frac{1}{3}\ \lambda + \frac{1}{3}\sum_{k=1}^{\lambda-1}(\frac{2}{3})^k(\lambda - k)$$

By local strategyproofness of sequential RSD, $\pi_i(\succ^t)$ is preferred to $\pi_i(\hat{\succ}^t)$, implying that for $\lambda = 1$, strong (or even weak) manipulablity is impossible. However, for $\lambda = 2$ because $\pi_i(\hat{\succ}^t) \neq \pi_i(\succ^t)$, we have $1 > \frac{2}{3} + (\frac{1}{3})\frac{2}{3}$, which also holds for any $\lambda \geq 2$, meaning that misreport will improve the allocation of ranking position o_1.

For o_2, we must have $W_i^\pi(\hat{\succ}^t, o_2) \geq W_i^\pi(\succ^t, o_2)$, that is,

$$\pi_i(\hat{\succ}^t) + (\lambda - 1) \geq \frac{1}{3}\ \lambda + \frac{1}{3}\sum_{k=1}^{\lambda-1}(\frac{2}{3})^k(\lambda - k) + \frac{1}{3}\sum_{k=1}^{\lambda}(\frac{2}{3})^{k-1}$$

For o_2, agent i's strategic report that changes the evolution of preferences for other agents would result in a randomized decision $\pi_i(\hat{\succ}^t)$ with at least $\frac{1}{n!} = \frac{1}{6}$ probability of receiving objects ranked first or second. Thus,

$$\frac{1}{6} + (\lambda - 1) \geq \frac{1}{3}\ \lambda + \frac{1}{3}\sum_{k=1}^{\lambda-1}(\frac{2}{3})^k(\lambda - k) + \frac{1}{3}\sum_{k=1}^{\lambda}(\frac{2}{3})^{k-1}$$

It is easy to verify that for any $\lambda \geq 4$ the above inequality holds, implying that sequential RSD is not even weakly gsd-strategyproof for $\lambda \geq 4$, while it is only weakly manipulable for $1 < \lambda < 4$. This concludes our proof. □

Appendix D

Proof of Proposition 5.1

Proof. For contradiction, assume that $\omega^t_{ij}(\pi(\succ^t)) > \frac{1}{2}$ for some $\succ^t$. By Equation 5.6 we write

$$\sum_{\mu \in \mathcal{M}: \mu(i) \succ^t_j \mu(j)} \pi(\mu \mid h^t) > \frac{1}{2}$$

Since RSD induces a probability distribution over all matching $\mu \in \mathcal{M}$,

$$\sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ^t_j \mu(j)}} \pi(\mu \mid h^t) + \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(j) \succ^t_j \mu(i)}} \pi(\mu \mid h^t) = 1$$

Therefore, by subtracting the above inequalities we have

$$\sum_{\substack{\mu \in \mathcal{M}: \\ \mu(j) \succ^t_j \mu(i)}} \pi(\mu \mid h^t) < \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ^t_j \mu(j)}} \pi(\mu \mid h^t)$$

which implies that agent j strictly prefers agent i's assignment to its own assignment. This immediately contradicts the uniformity of distribution over priority orderings and the weakly envy-freeness of the RSD mechanism. $\square$

Appendix E

Proof of Lemma 5.6

Proof. The proof follows directly from Algorithm 3. Here we provide a constructive proof. Take any two agents $i, j \in N$ at time $t+1$. Consider two different sequence of matchings resulting in two dominance histories $\hat{\boldsymbol{d}}^t$ and $\boldsymbol{d}^t$ such that agent i has dominated agent j in the trajectory leading to $\hat{\boldsymbol{d}}^t$ compared to $\boldsymbol{d}^t$. That is, $\hat{d}^t_{ij} > d^t_{ij}$.

At time $t+1$, as shown in Algorithm 3, for each priority ordering where i is ordered before j the priority ordering changes such that j selects an alternative before agent i, hence, for agent j we can write

$$\sum_{\substack{\mu \in \mathcal{M}: \\ \mu(j) \succ_j^{t+1} \mu(i)}} \pi(\mu \mid \hat{h}^{t+1}) > \sum_{\substack{\mu \in \mathcal{M}: \\ \mu(i) \succ_j^{t+1} \mu(j)}} \pi(\mu \mid h^{t+1})$$

Agent j's assignment improves given $\hat{h}^{t+1}$. Since the random matching induced by Adaptive RSD is a probability function, for agent i the probability of receiving alternatives that are preferred to agent j's assignment is less than or equal the same probability under h^{t+1}. Note that the equal part in the inequality is only valid in situations where the two agents are not competing for any of the alternatives at $t+1$. $\square$

References

[1] Atila Abdulkadiroğlu, Yeon-Koo Che, and Yosuke Yasuda. Resolving conflicting preferences in school choice: The Boston mechanism reconsidered. *American Economic Review*, 101(1):399–410, February 2011.

[2] Atila Abdulkadiroğlu and Simon Loerscher. Dynamic house allocations. *Working papar*, 2007.

[3] Atila Abdulkadiroğlu, Parag A. Pathak, and Alvin E. Roth. Strategy-proofness versus efficiency in matching with indifferences: redesigning the new york city high school match. Technical report, National Bureau of Economic Research, 2009.

[4] Atila Abdulkadiroğlu, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez. The boston public school match. *American Economic Review*, pages 368–371, 2005.

[5] Atila Abdulkadiroğlu, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez. Changing the boston school choice mechanism. Technical report, National Bureau of Economic Research, 2006.

[6] Atila Abdulkadiroğlu and Tayfun Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.

[7] Atila Abdulkadiroğlu and Tayfun Sönmez. House allocation with existing tenants. *Journal of Economic Theory*, 88(2):233–260, 1999.

[8] Mohammad Akbarpour, Shengwu Li, and Shayan Oveis Gharan. Thickness and information in dynamic matching markets. *Available at SSRN 2394319*, 2014.

[9] Jorge Alcalde-Unzu and Marc Vorsatz. Do we agree? measuring the cohesiveness of preferences. *Theory and Decision*, pages 1–27.

[10] Abdullah Alsheddy and Edward P.K. Tsang. Empowerment scheduling for a field workforce. *Journal of Scheduling*, 14(6):639–654, 2011.

[11] Kenneth Joseph Arrow. Essays in the theory of risk-bearing. 1974.

[12] Kenneth Joseph Arrow, Amartya Sen, and Kōtarō Suzumura. *Handbook of Social Choice and Welfare*, volume 1. Gulf Professional Publishing, 2002.

[13] Itai Ashlagi, Felix Fischer, Ian A. Kash, and Ariel D. Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 2013.

[14] Itai Ashlagi, Patrick Jaillet, and Vahideh H. Manshadi. Kidney exchange in dynamic sparse heterogenous pools. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 25–26, New York, NY, USA, 2013. ACM.

[15] Susan Athey and Ilya Segal. An efficient dynamic mechanism. *Econometrica*, 81(6):2463–2485, 2013.

[16] Haris Aziz, Felix Brandt, and Markus Brill. The computational complexity of random serial dictatorship. *Economics Letters*, 121(3):341–345, 2013.

[17] Haris Aziz, Felix Brandt, and Markus Brill. On the tradeoff between economic efficiency and strategy proofness in randomized social choice. In *Proceedings of the 12th international conference on Autonomous agents and multi-agent systems*, AAMAS 2013, pages 455–462. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[18] Haris Aziz, Jiashu Chen, Aris Filos-Ratsikas, Simon Mackenzie, and Nicholas Mattei. Egalitarianism of random assignment mechanisms. *arXiv preprint arXiv:1507.06827*, 2015.

[19] Haris Aziz, Serge Gaspers, Simon Mackenzie, Nicholas Mattei, Nina Narodytska, and Toby Walsh. Equilibria under the probabilistic serial rule. In *Proceedings of the 24th International Conference on Artificial Intelligence*, IJCAI 2015, pages 1105–1112. AAAI Press, 2015.

[20] Haris Aziz, Serge Gaspers, Simon Mackenzie, Nicholas Mattei, Nina Narodytska, and Toby Walsh. Manipulating the probabilistic serial rule. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, pages 1451–1459. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[21] Haris Aziz, Serge Gaspers, Simon Mackenzie, and Toby Walsh. Fair assignment of indivisible objects under ordinal preferences. *Artificial Intelligence*, 227:71–92, 2015.

[22] Sophie Bade. Random serial dictatorship: The one and only. Max Planck Institute for Research on Collective Goods (Unpublished), 2014.

[23] Sophie Bade. Serial dictatorship: the unique optimal allocation rule when information is endogenous. *Theoretical Economics*, To appear., 2014.

[24] Jonathan F. Bard and Hadi W. Purnomo. Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164(2):510–534, 2005.

[25] Dirk Bergemann and Juuso Välimäki. The dynamic pivot mechanism. *Econometrica*, 78(2):771–789, 2010.

[26] Anand Bhalgat, Deeparnab Chakrabarty, and Sanjeev Khanna. Social welfare in one-sided matching markets without money. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 87–98. Springer, 2011.

[27] Francis Bloch and David Cantala. Markovian assignment rules. *Social Choice and Welfare*, 40(1):1–25, 2013.

[28] Liad Blumrosen and Noam Nisan. Combinatorial auctions. *Algorithmic Game Theory*, 267:300, 2007.

[29] Anna Bogomolnaia and Eun Jeong Heo. Probabilistic assignment of objects: Characterizing the serial rule. *Journal of Economic Theory*, 147(5):2072–2082, 2012.

[30] Anna Bogomolnaia and Hervé Moulin. A new solution to the random assignment problem. *Journal of Economic Theory*, 100(2):295–328, 2001.

[31] Craig Boutilier. Planning, learning and coordination in multiagent decision processes. In *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge*, pages 195–210. Morgan Kaufmann Publishers Inc., 1996.

[32] Craig Boutilier, Moises Goldszmidt, and Bikash Sabata. Sequential auctions for the allocation of resources with complementarities. In *International Joint Conference on Artificial Intelligence*, volume 16, pages 527–534. Citeseer, 1999.

[33] Sylvain Bouveret and Jérôme Lang. Manipulating picking sequences. In *In Proceedings of the 21st European Conference on Artificial Intelligence (ECAI14)*, pages 141–146, Prague, Czech Republic, 2014. IOS Press.

[34] Steven J. Brams and Daniel L. King. Efficient fair division help the worst off or avoid envy? *Rationality and Society*, 17(4):387–421, 2005.

[35] Steven J. Brams and Philip D. Straffin. Prisoners' dilemama and professional sports drafts. *American Mathematical Monthly*, pages 80–88, 1979.

[36] Steven J. Brams and Alan D. Taylor. *The Win-Win solution: guaranteeing fair shares to everybody*. WW Norton & Company, 2000.

[37] Simina Brânzei, Tomasz Michalak, Talal Rahwan, Kate Larson, and Nicholas R. Jennings. Matchings with externalities and attitudes. In *Proceedings of the 12th international conference on Autonomous agents and multi-agent systems*, AAMAS 2013, pages 295–302. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[38] Eric Budish and Estelle Cantillon. The multi-unit assignment problem: Theory and evidence from course allocation at harvard. *The American economic review*, 102(5):2237–71, 2012.

[39] Eric Budish, Yeon-Koo Che, Fuhito Kojima, and Paul Milgrom. Designing random allocation mechanisms: Theory and applications. *The American Economic Review*, 103(2):585–623, 2013.

[40] Edmund K. Burke, Patrick De Causmaecker, Greet Vanden Berghe, and Hendrik Van Landeghem. The state of the art of nurse rostering. *Journal of scheduling*, 7(6):441–499, 2004.

[41] Ruggiero Cavallo. Mechanism design for dynamic settings. *ACM SIGecom Exchanges*, 8(2):7, 2009.

[42] Ruggiero Cavallo, David C. Parkes, and Satinder Singh. Efficient mechanisms with dynamic populations and dynamic types. Harvard University. Division of Engineering and Applied Physics, 2009.

[43] John R. Chamberlin and Paul N. Courant. Representative deliberations and representative decisions: Proportional representation and the Borda rule. *American Political Science Review*, 77(03):718–733, 1983.

[44] Yeon-Koo Che and Fuhito Kojima. Asymptotic equivalence of probabilistic serial and random priority mechanisms. *Econometrica*, 78(5):1625–1672, 2010.

[45] Yann Chevaleyre, Ulle Endriss, Sylvia Estivie, and Nicolas Maudet. Reaching envy-free states in distributed negotiation settings. In *Proceedings of the 20th International Conference on Artificial Intelligence*, IJCAI 2007, pages 1239–1244.

[46] Wonki Jo Cho. Incentive properties for ordinal mechanisms. *Games and Economic Behavior*, 95:168–177, 2016.

[47] Andrew M. Colman and Jonathan A. Stirk. Singleton bias and lexicographic preferences among equally valued alternatives. *Journal of Economic Behavior & Organization*, 40(4):337–351, 1999.

[48] Partha Dasgupta, Peter Hammond, and Eric Maskin. The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, pages 185–216, 1979.

[49] Melanie L. De Grano, D.J. Medeiros, and David Eitel. Accommodating individual preferences in nurse scheduling via auctions and optimization. *Health Care Management Science*, 12(3):228–242, 2009.

[50] John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. Dynamic matching via weighted myopia with application to kidney exchange. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, AAAI 2012, pages 1340–1346, 2012.

[51] Carmel Domshlak, Eyke Hüllermeier, Souhila Kaci, and Henri Prade. Preferences in AI: An overview. *Artificial Intelligence*, 175(7):1037–1052, 2011.

[52] Aimee Drolet and Mary Frances Luce. The rationalizing effects of cognitive load on emotion-based trade-off avoidance. *Journal of Consumer Research*, 31(1):63–77, 2004.

[53] Joanna Drummond and Craig Boutilier. Elicitation and approximately stable matching with partial preferences. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pages 97–105. AAAI Press, 2013.

[54] Umut Dur. Dynamic school choice problem. Technical report, 2012. Working paper.

[55] Lars Ehlers and Bettina Klaus. Coalitional strategy-proof and resource-monotonic solutions for multiple assignment problems. *Social Choice and Welfare*, 21(2):265–280, 2003.

[56] Özgün Ekici and Onur Kesten. An equilibrium analysis of the probabilistic serial mechanism. *International Journal of Game Theory*, pages 1–20, 2015.

[57] Aris Filos-Ratsikas, Søren Kristoffer Stiil Frederiksen, and Jie Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *Algorithmic Game Theory*, pages 1–12. Springer, 2014.

[58] Peter C. Fishburn. Lexicographic orders, utilities and decision rules: A survey. *Management Science*, pages 1442–1471, 1974.

[59] Peter C. Fishburn. Axioms for lexicographic preferences. *The Review of Economic Studies*, pages 415–419, 1975.

[60] Daniel Fragiadakis, Atsushi Iwasaki, Peter Troyan, Suguru Ueda, and Makoto Yokoo. Strategyproof matching with minimum quotas. *ACM Transactions on Economics and Computation*, 4(1):6, 2016.

[61] David Gale. College course assignments and optimal lotteries. *University of California at Berkeley*, 1987.

[62] David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

[63] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica: Journal of the Econometric Society*, pages 587–601, 1973.

[64] Judy Goldsmith and Ulrich Junker. Preference handling for artificial intelligence. *AI Magazine*, 29(4):9, 2009.

[65] Judy Goldsmith and Robert H. Sloan. The AI conference paper assignment problem. In *Proceedings of the AAAI Workshop on Preference Handling for Artificial Intelligence*, pages 53–57, 2007.

[66] Masahiro Goto, Naoyuki Hashimoto, Atsushi Iwasaki, Yujiro Kawasaki, Suguru Ueda, Yosuke Yasuda, and Makoto Yokoo. Strategy-proof matching with regional minimum quotas. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, pages 1225–1232. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[67] Theodore Groves. Incentives in teams. *Econometrica: Journal of the Econometric Society*, pages 617–631, 1973.

[68] Sujit Gujar and David C. Parkes. Dynamic matching with a fall-back option. In *Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, pages 263–268. IOS Press, 2010.

[69] Josef Hadar and William R. Russell. Rules for ordering uncertain prospects. *The American Economic Review*, pages 25–34, 1969.

[70] Isa E. Hafalir. Stability of marriage with externalities. *International Journal of Game Theory*, 37(3):353–369, 2008.

[71] John C. Harsanyi. *Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility*. Springer, 1980.

[72] John W. Hatfield. Strategy-proof, efficient, and nonbossy quota allocations. *Social Choice and Welfare*, 33(3):505–515, 2009.

[73] Greg Hines and Kate Larson. Preference elicitation for risky prospects. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 889–896. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[74] Hadi Hosseini. One-sided matching with dynamic preferences. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, pages 2005–2006. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[75] Hadi Hosseini and Kate Larson. Strategyproof quota mechanisms for multiple assignment problems. *arXiv preprint arXiv:1507.07064*, 2015.

[76] Hadi Hosseini, Kate Larson, and Robin Cohen. Matching with dynamic ordinal preferences. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 936–943, 2015. To appear.

[77] Hadi Hosseini, Kate Larson, and Robin Cohen. On manipulability of random serial dictatorship in sequential matching with dynamic preferences. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*, pages 4168–4169, 2015.

[78] Hadi Hosseini, Kate Larson, and Robin Cohen. Random serial dictatorship versus probabilistic serial rule: A tale of two random mechanisms. *CoRR*, abs/1503.01488, 2015.

[79] Hadi Hosseini, Kate Larson, and Robin Cohen. An empirical comparison of one-sided matching mechanisms. In *The 3rd Workshop on Exploring Beyond the Worst Case in Computational Social Choice (EXPLORE) at AAMAS 2016*, 2016.

[80] Hadi Hosseini, Kate Larson, and Robin Cohen. Investigating the characteristics of one-sided matching mechanisms: (extended abstract). In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, pages 1443–1444. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[81] David Hugh-Jones, Morimitsu Kurino, and Christoph Vanberg. An experimental study on the incentives of the probabilistic serial mechanism. Technical report, Discussion Paper, Social Science Research Center Berlin (WZB), Research Area Markets and Politics, Research Unit Market Behavior, 2013.

[82] Aanund Hylland and Richard Zeckhauser. The efficient allocation of individuals to positions. *The Journal of Political Economy*, pages 293–314, 1979.

[83] Robert W. Irving, Telikepalli Kavitha, Kurt Mehlhorn, Dimitrios Michail, and Katarzyna Paluch. Rank-maximal matchings. In *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 68–75. Society for Industrial and Applied Mathematics, 2004.

[84] Barbara E. Kahn and Jonathan Baron. An exploratory study of choice rules favored for high-stakes decisions. *Journal of Consumer Psychology*, 4(4):305–328, 1995.

[85] Thomas Kalinowski, Nina Narodytska, and Toby Walsh. A social welfare optimal sequential allocation procedure. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pages 227–233. AAAI Press, 2013.

[86] Thomas Kalinowski, Nina Narodytska, Toby Walsh, and Lirong Xia. Strategic behavior when allocating indivisible goods sequentially. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI 2013)*, pages 452–458, 2013.

[87] Akshay-Kumar Katta and Jay Sethuraman. A solution to the random assignment problem on the full preference domain. *Journal of Economic Theory*, 131(1):231–250, 2006.

[88] John Kennes, Daniel Monde, and Norovsambuu Tumennasan. Dynamic matching markets and the deferred acceptance mechanism. Technical report, School of Economics and Management, University of Aarhus, 2015.

[89] John Kennes, Daniel Monte, and Norovsambuu Tumennasan. The day care assignment: A dynamic matching problem. *American Economic Journal: Microeconomics*, 6(4):362–406, 2014.

[90] David A. Kohler and R. Chandrasekaran. A class of sequential games. *Operations Research*, 19(2):270–277, 1971.

[91] Rajeev Kohli and Kamel Jedidi. Representation and inference of lexicographic preference models and their variants. *Marketing Science*, 26(3):380–399, 2007.

[92] Fuhito Kojima. Random assignment of multiple indivisible objects. *Mathematical Social Sciences*, 57(1):134–142, 2009.

[93] Fuhito Kojima and Mihai Manea. Incentives in the probabilistic serial mechanism. *Journal of Economic Theory*, 145(1):106–123, 2010.

[94] Piotr Krysta, David Manlove, Baharak Rastegari, and Jinshan Zhang. Size versus truthfulness in the house allocation problem. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 453–470. ACM, 2014.

[95] Ryoji Kurata, Masahiro Goto, Atsushi Iwasaki, and Makoto Yokoo. Controlled school choice with soft bounds and overlapping types. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 951–957, 2015.

[96] Morimitsu Kurino. House allocation with overlapping generations. *American Economic Journal: Microeconomics*, 6(1):258–289, 2014.

[97] Qingmin Liu and Marek Pycia. Ordinal efficiency, fairness, and incentives in large markets. *Unpublished mimeo*, 2013.

[98] Dipjyoti Majumdar and Arunava Sen. Ordinally Bayesian incentive compatible voting rules. *Econometrica*, 72(2):523–540, 2004.

[99] Mihai Manea. Asymptotic ordinal inefficiency of random serial dictatorship. *Theoretical Economics*, 4(2):165–197, 2009.

[100] David Manlove. *Algorithmics of matching under preferences*. World Scientific Publishing, 2013.

[101] Andrew McLennan. Ordinal efficiency and the polyhedral separating hyperplane theorem. *Journal of Economic Theory*, 105(2):435–449, 2002.

[102] Timo Mennle and Sven Seuken. An axiomatic approach to characterizing and relaxing strategyproofness of one-sided matching mechanisms. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 37–38. ACM, 2014.

[103] Timo Mennle and Sven Seuken. Hybrid mechanisms: Trading off strategyproofness and efficiency in one-sided matching. Technical report, 2014. Working Paper.

[104] Timo Mennle and Sven Seuken. Hybrid mechanisms: Trading off strategyproofness and efficiency of random assignment mechanisms. Technical report, 2015. Working Paper.

[105] Timo Mennle and Sven Seuken. Partial strategyproofness: An axiomatic approach to relaxing strategyproofness for assignment mechanisms. Technical report, 2015. Working paper.

[106] Timo Mennle, Michael Weiss, Basil Philipp, and Sven Seuken. The power of local manipulation strategies in assignment mechanisms. In *Proceedings of the 24th International Conference on Artificial Intelligence*, IJCAI'15, pages 82–89. AAAI Press, 2015.

[107] Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.

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

[109] Szilvia Pápai. Strategyproof multiple assignment using quotas. *Review of Economic Design*, 5(1):91–105, 2000.

[110] Szilvia Pápai. Strategyproof and nonbossy multiple assignments. *Journal of Public Economic Theory*, 3(3):257–271, 2001.

[111] David C. Parkes. Online mechanisms. *Algorithmic Game Theory, ed. N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani*, pages 411–439, 2007.

[112] David C. Parkes and Ariel D. Procaccia. Dynamic social choice with evolving preferences. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, AAAI 2013, pages 767–773, 2013.

[113] Parag A. Pathak. Lotteries in student assignment. *Unpublished mimeo, Harvard University*, 2006.

[114] Parag A. Pathak and Jay Sethuraman. Lotteries in student assignment: An equivalence result. *Theoretical Economics*, 6(1):1–17, 2011.

[115] Martin L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming.* John Wiley & Sons, Inc., 1994.

[116] Alvin E. Roth and John H. Kagel. *The handbook of experimental economics*, volume 1. Princeton university press Princeton, 1995.

[117] Alvin E. Roth, Tayfun Sönmez, and M. Utku Ünver. Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488, 2004.

[118] Alvin E. Roth and Marilda A. Oliveira Sotomayor. *Two-sided matching: A study in game-theoretic modeling and analysis*, volume 18. Cambridge University Press, 1992.

[119] Daniela Saban and Jay Sethuraman. A note on object allocation under lexicographic preferences. *Journal of Mathematical Economics*, 2013.

[120] Daniela Saban and Jay Sethuraman. A note on object allocation under lexicographic preferences. *Journal of Mathematical Economics*, 50:283–289, 2014.

[121] Daniela Saban and Jay Sethuraman. The complexity of computing the random priority allocation matrix. *Mathematics of Operations Research*, 40(4):1005–1014, 2015.

[122] Mark A. Satterthwaite and Hugo Sonnenschein. Strategy-proof allocation mechanisms at differentiable points. *The Review of Economic Studies*, pages 587–597, 1981.

[123] Mark Allen Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

[124] Leonard J. Schulman and Vijay V. Vazirani. Allocation of divisible goods under lexicographic preferences. *arXiv preprint arXiv:1206.4366*, 2012.

[125] James Schummer and Rakesh V. Vohra. Mechanism design without money. *Algorithmic Game Theory*, 10:243–299, 2007.

[126] Lloyd Shapley and Herbert Scarf. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1):23–37, 1974.

[127] Lloyd S. Shapley. Utility comparison and the theory of games. *La Decision. Paris: Edition du Centre de la Recherche Scientifique, France.*, pages 251–263, 1969.

[128] Yoav Shoham and Kevin Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.

[129] Sue Perrott Siferd and W.C. Benton. Workforce staffing and scheduling: Hospital nursing specific models. *European Journal of Operational Research*, 60(3):233–246, 1992.

[130] Paul Slovic. Choice between equally valued alternatives. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3):280, 1975.

[131] Tayfun Sönmez and M. Utku Ünver. Course bidding at business schools. *International Economic Review*, 51(1):99–123, 2010.

[132] Tayfun Sönmez and M. Utku Ünver. House allocation with existing tenants: A characterization. *Games and Economic Behavior*, 69(2):425–445, 2010.

[133] Lars-Gunnar Svensson. Queue allocation of indivisible goods. *Social Choice and Welfare*, 11(4):323–330, 1994.

[134] Lars-Gunnar Svensson. Strategy-proof allocation of indivisible goods. *Social Choice and Welfare*, 16(4):557–567, 1999.

[135] Katia P. Sycara. Multiagent systems. *AI magazine*, 19(2):79, 1998.

[136] Robert M. Townsend. Optimal multiperiod contracts and the gain from enduring relationships under private information. *The Journal of Political Economy*, pages 1166–1186, 1982.

[137] Amos Tversky, Shmuel Sattath, and Paul Slovic. Contingent weighting in judgment and choice. *Psychological Review*, 95(3):371, 1988.

[138] M. Utku Ünver. Dynamic kidney exchange. *The Review of Economic Studies*, 77(1):372–414, 2010.

[139] Hal R. Varian. Equity, envy, and efficiency. *Journal of Economic Theory*, 9(1):63–91, 1974.

[140] Rakesh V. Vohra. Dynamic mechanism design. *Surveys in Operations Research and Management Science*, 17(1):60–68, 2012.

[141] John Von Neumann. A certain zero-sum two-person game equivalent to the optimal assignment problem. *Contributions to the Theory of Games*, 2:5–12, 1953.

[142] Gerhard Weiss. *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press, 1999.

[143] Michael Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, 2009.

[144] Michael Yee, Ely Dahan, John R. Hauser, and James Orlin. Greedoid-based non-compensatory inference. *Marketing Science*, 26(4):532–549, 2007.

[145] Lin Zhou. On a conjecture by gale about one-sided matching problems. *Journal of Economic Theory*, 52(1):123–135, 1990.