One-Sided Matching Under Strategic Reporting

INSIGHTS FROM AMSTERDAM SCHOOL CHOICE



Mayesha Tasnim

One-Sided Matching Under Strategic Reporting

Insights from Amsterdam School Choice

Mayesha Tasnim

Cover design: © Tahsin Adiba Rahman, 2026
Copyright © 2026 by Mayesha Tasnim
ISBN: 978-94-6534-579-6
Printed by: ProefschriftMaken | www.proefschriftmaken.nl

One-Sided Matching Under Strategic Reporting
Insights from Amsterdam School Choice

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op vrijdag 4 september 2026, te 11.00 uur

door Mayesha Tasnim
geboren te Dhaka

Promotiecommissie

Promotores:	prof. dr. S. Ghebreab	Universiteit van Amsterdam
	prof. dr. T. Blanke	Universiteit van Amsterdam
Copromotores:	dr. M.A. Baak	ING Bank
	dr. rer. nat. E. Acar	Universiteit van Amsterdam
Overige leden:	prof. dr. U. Endriss	Universiteit van Amsterdam
	prof. dr. M.V. Dignum	Umeå Universitet
	prof. dr. E.M. Heemskerk	Universiteit van Amsterdam
	prof. dr. J.A. Good	Universiteit van Amsterdam
	dr. R.T. Rădulescu	Utrecht University

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

CONTENTS

1 INTRODUCTION 1
1.1 Amsterdam School Choice . . . 2
1.2 Research Questions . . . 3
1.3 Thesis Contributions . . . 5
1.4 List of publications . . . 7

2 STRATEGIC REPORTING IN RANK MINIMIZATION 9
2.1 Introduction . . . 10
2.2 Related Work . . . 12
2.3 Rank Minimization using Hungarian Algorithm (RM-HAL) . . . 13
2.4 Strategies for RM-HAL . . . 18
2.5 Datasets . . . 22
2.6 Experimental results . . . 23
2.7 Discussion . . . 30
2.8 Limitations . . . 31
2.9 Conclusion . . . 31

3 STRATEGIC REPORTING IN PLACEMENT GUARANTEE 33
3.1 Introduction . . . 34
3.2 Amsterdam School Choice and the Placement Guarantee . . . 36
3.3 Related Work . . . 36
3.4 School Choice Model . . . 37
3.5 Strategic Reporting under Placement Guarantee . . . 38
3.6 Simulation Environment and Data . . . 40
3.7 Experimental Design . . . 41
3.8 Results . . . 42
3.9 Discussion . . . 44
3.10 Limitations . . . 46
3.11 Conclusion . . . 46

4 STRATEGIC REPORTING IN PRACTICE 49
4.1 Introduction . . . 50
4.2 Related Work . . . 51
4.3 Methods . . . 54
4.4 Results . . . 57
4.5 Discussion . . . 61
4.6 Conclusion . . . 62

5 SEQUENTIAL LEARNING FOR ONE-SIDED MATCHING 65
5.1 Introduction . . . 66
5.2 Related Work . . . 68
5.3 Preliminaries . . . 69
5.4 Matching as a Sequential Decision Problem . . . 71
5.5 Evaluating Robustness to Preference Misreporting . . . 74
5.6 Experiments . . . 77

5.7 Results . . . 79
5.8 Discussion . . . 83
5.9 Conclusion . . . 85

6 REFLECTIONS ON CIVIC AI RESEARCH 87
6.1 Introduction . . . 88
6.2 Our Research Principles . . . 89
6.3 The Case of Amsterdam School Choice . . . 90
6.4 Reflecting on Our Experience Building Civic-Centered AI . . . 92
6.5 Conclusion . . . 95

7 CONCLUSION 97
7.1 Summary of Findings . . . 97
7.2 Main Conclusions . . . 100
7.3 Future Directions . . . 103
7.4 Final Reflection . . . 105

A APPENDIX: CHAPTER 2 107
A.1 RM using Hungarian algorithm . . . 107
A.2 Example Application of Best Response Strategy . . . 107
A.3 Acceptance of top-n Matches . . . 108
A.4 Amsterdam Dataset . . . 108
A.5 Additional Results . . . 109

B APPENDIX: CHAPTER 3 113
B.1 RSD with Placement Guarantee Algorithm . . . 113
B.2 Additional Experimental Results . . . 113

C APPENDIX: CHAPTER 4 117
C.1 Survey Questionnaire . . . 117
C.2 SP/A Modelling Details . . . 120
C.3 Baseline Models . . . 120
C.4 Estimating Security and Potential . . . 121
C.5 Correlation Among Covariates . . . 122

D APPENDIX: CHAPTER 5 123
D.1 Appendix A: GFlowNet Implementation . . . 123
D.2 Appendix B: DQN Implementation Details . . . 125
D.3 Robustness under Correlated Preferences . . . 125
D.4 Population-Adoption Results . . . 127

SUMMARY 131

NEDERLANDSE SAMENVATTING 133

BIBLIOGRAPHY 135

ACKNOWLEDGEMENTS 145

For my family,
who made everything possible

1

INTRODUCTION

This thesis centers around a deceptively simple problem: how can we assign people to resources on the basis of their preferences? This is called a *one-sided matching problem for indivisible resources*, also known as an assignment problem between people and items [1]. Despite its relatively simple formal structure, this is also an important administrative problem, especially in settings where publicly available resources must be allocated in a fair and transparent way. Examples include house allocation, course allocation, and school choice, where people submit rankings over the resource in question, and a centralized system determines who receives what.

When designing one-sided matching methods, designers consider two important criteria. First, the method should adhere well to the submitted preferences, so that participants receive resources they value highly. Second, it should make reporting truthful preferences optimal, so that no participant benefits from submitting false preferences. In the mechanism design literature, these criteria are known as efficiency and strategyproofness, respectively. A well-known difficulty in one-sided matching is that these goals are often in conflict, and cannot in general be fully achieved together [2, 3]. This thesis examines this tension in the context of school choice.

In practice, school choice systems have often responded to this problem by prioritizing strategyproofness. If truthful reporting is optimal, then families do not need to game the system, the cognitive burden is lower, and the submitted preferences are a more reliable basis for assigning scarce school seats. For these reasons, strategyproofness has been an important design goal both in the school choice literature and in real-world reforms. Boston replaced its old mechanism with deferred acceptance in order to reduce the strategic burden on parents, and deferred-acceptance-based systems were also adopted in New York City for closely related reasons [4, 5]. New Orleans' OneApp similarly relied on strategyproof design principles in building a more coherent school-choice process [6].

But in one-sided matching, prioritizing strategyproofness can come at a substantial cost when it leads to inefficient allocations [5, 7]. Protecting truthful reporting is attractive, but not if it comes with allocations that are poor enough to make the system unsatisfactory in practice. The Amsterdam school choice system is a clear example of this problem. Since 2016, Amsterdam has used a single-lottery matching procedure operationalized through Random Serial Dictatorship (RSD) [1, 8]. In 2017, parents sought legal action after some students were not placed at any of the schools on their submitted preference list [9]. Once policymakers try to improve those outcomes, however, they also have to confront the possibility of incentivizing strategic behavior.

This is the core problem of this thesis. Once strategyproofness is recognized as too costly in some one-sided matching settings, we need to ask: how can we design good matching algorithms when there is scope for strategic behavior? This question, in turn, requires a shift in focus. If some relaxation of strict strategyproofness is needed in order to get better matches, then we need to understand how misreporting works,

how strategic behavior evolves, how much damage it causes, and which participants are most affected[1].

This is not merely a computational problem, but also a problem about how institutions distribute cognitive burden, risk, and trust among participants. When a mechanism rewards strategic reporting, some participants are much better placed than others to recognize the opportunity, act on it, and absorb the risk of failure: turning a formal tradeoff into a distributional one. Formal guarantees, even when they hold in theory, can mean something quite different in practice once they are mediated by institutional rules, public communication, and the lived experience of families navigating the system. For that reason, the efficiency–strategyproofness tradeoff in one-sided matching should not be treated only as a formal impossibility result, but as a practical *design problem* for real public systems.

The central argument of this thesis is that one-sided matching systems need to be studied and designed under the possibility of strategic behavior, rather than under the assumption that strict strategyproofness will always be a sufficient answer. The chapters of this thesis approach that problem from several angles: by studying how strategic behavior affects alternative matching methods, how policy rules can reintroduce strategic behavior, whether real participants actually respond strategically in practice, whether computational methods can help explore better tradeoffs, and how the Amsterdam school choice case offers broader lessons about AI, interdisciplinarity, and civic engagement in the design of public allocation systems.

1.1 AMSTERDAM SCHOOL CHOICE

Since 2015, Amsterdam has used a citywide open-choice secondary school system: students may apply to any school regardless of residence, with no neighborhood priorities [10–12]. This was introduced to promote equal access and parental choice. The result is a market with sharply skewed demand — a small number of schools attract far more first-choice applications than they can admit, while others have spare capacity [12].

Since 2016, the city has used a single-lottery procedure operationalized through Random Serial Dictatorship (RSD), administered by the Association of School Boards (OSVO) [8, 13]. Because schools do not rank students, the standard Deferred Acceptance algorithm reduces to RSD in this setting [1, 3]. Students submit a ranked list of schools (currently a minimum of 12 for the pre-university and general tracks, fewer for vocational), and are assigned to their highest-ranked school with remaining capacity in a randomly drawn priority order [8, 14].

Despite the mechanism's strategyproofness, each year a small but visible group of students were placed at schools they had ranked very low — in some cases as far down as their 18th choice [9]. This led to parental complaints, media coverage, and legal challenges [15]. In 2019, in response to sustained dissatisfaction, OSVO introduced a *placement guarantee*: students who submit a list of at least the required length and remain unassigned after the main round are guaranteed a seat at one of the schools on their list, implemented through a proportional capacity expansion of about 4% [8]. The required list length itself has shifted in response to feasibility pressures —

1 In case of school choice, this affected groups are primarily students and their parents.

reduced from 12 to 9 in 2025, then restored to 12 in 2026 when the shorter list was found to compromise the guarantee [16, 17].

Despite its unique circumstances, the Amsterdam school choice is not a marginal case, but rather contains some exemplary points. It is one of the few European cities with a fully open citywide school choice system, with publicly documented mechanism changes, an active parental advocacy ecosystem [18], and a school-boards association that has shown willingness to modify the system in response to pressure. This makes it an unusually legible setting in which to study the interplay between formal mechanism properties, institutional policy modifications, and behavioral response — the three components around which this thesis is organized.

1.2 RESEARCH QUESTIONS

The broader goal of this thesis is to understand how one-sided matching systems should be evaluated and designed when attempts to improve allocation quality can also create new incentives for strategic behavior. This leads to the following overarching research question:

> ***How should one-sided matching systems be evaluated and designed when attempts to improve allocations incentivize strategic misreporting of preferences?***

This question is addressed through five more specific research questions, each taken up in a dedicated chapter.

A natural place to begin is with the most direct version of the problem. If strategyproof mechanisms such as Random Serial Dictatorship produce inefficient allocations, then one possible response is to consider more welfare-oriented alternatives. The concern is whether such mechanisms become vulnerable to strategic manipulation, and who bears that vulnerability.

Research Question 1: *If a more welfare-oriented matching rule is used instead of a strategyproof one, what strategies become available, how severe is the resulting risk, and to whom?*

Chapter 2 takes up this question by studying rank-minimizing matching as an alternative to RSD. The chapter derives a manipulation strategy that requires only first-choice popularity information, and shows that the resulting risk is both easier to exploit and more distributionally uneven than prior "not obviously manipulable" results suggest.

Strategic misreporting is not only a concern when policymakers explicitly adopt a more manipulable mechanism. Even when the base mechanism is strategyproof, policy rules layered around it can alter the incentives participants face. This raises a second question: if a matching method is strategyproof in isolation, can it still become strategically vulnerable once it is embedded in a broader institutional setting?

Research Question 2: *If the base matching method is strategyproof, can institutional modifications create strategic vulnerabilities?*

Chapter 3 takes up this question by analyzing the placement guarantee rule used in Amsterdam. The chapter identifies a profitable strategy for students with few sincere preferences, and shows that the guarantee opens a *promise gap*: a mismatch between what the safeguard publicly assures and what it can sustain once participants adapt to it. The resulting strategic adoption can render the guarantee infeasible at the system level, and distributes its benefits unequally across groups.

The first two questions identify two sources of strategic incentive: the matching method itself, and the policy rules layered around it. Showing that an incentive exists formally is not the same as showing that participants will use it. A manipulation may be available in theory while remaining behaviorally unrealistic, too risky, or too opaque to recognize. This motivates a shift from formal possibility to behavioral response.

Research Question 3: *If strategic incentives exist formally, how do real participants respond to them?*

Chapter 4 takes up this question through a survey of 140 Amsterdam parents during the 2025 application cycle, modeled using Security/Potential Aspiration theory. The chapter finds that transparency about strategic risk acts as a deterrent rather than an enabler of manipulation, and that parents in the more competitive academic tracks are more risk-averse rather than more strategic, reversing a common assumption in the school-choice literature.

The first three questions describe where strategic incentives arise in classical matching systems and when participants are likely to act on them. A further question is whether this landscape is fixed, or whether the design space itself can be widened. Rather than choosing among a small set of classical matching methods, can learning-based methods help identify alternative points in the tradeoff between efficiency and strategic risk?

Research Question 4: *Can learning-based methods explore alternative tradeoffs between efficiency and strategyproofness?*

Chapter 5 addresses this question by formulating one-sided matching as a sequential decision problem and training Deep Q-Networks and Generative Flow Networks. The chapter tests the learned policies against four misreporting strategies of varying behavioral plausibility, and finds that robustness to misreporting is multi-dimensional, that individual incentives do not predict population-level outcomes, and that learned policies can reach positions in the efficiency-robustness landscape unavailable to classical mechanisms.

The first four chapters bring formal, institutional, behavioral, and computational perspectives to one problem. Their findings together suggest that allocation systems shape outcomes, incentives, burdens, and trust at once, which raises a broader question about how this kind of research should be done.

Research Question 5: *What do these findings imply for responsible AI research for civic purposes?*

Chapter 6 takes up this question by reflecting on the Amsterdam school choice case and proposing a framework for civic AI research organized around three dimensions: science and society, fundamental and applied, and quantitative and qualitative. Drawing on this case, the chapter argues that civic engagement is what makes interdisciplinarity practically possible, rather than the other way around.

Chapter 7 concludes the thesis by summarizing the main findings, outlining directions for future research, and ending with a personal reflection.

1.3 THESIS CONTRIBUTIONS

Together, the chapters of this thesis make contributions across theoretical, algorithmic, empirical, and policy dimensions.

THEORETICAL AND ALGORITHMIC CONTRIBUTIONS

- We derive a best-response strategy and corresponding Nash equilibrium for RM-HAL under a simplified one-step cost function (Chapter 2).
- We formalize the Amsterdam placement guarantee as an ex-post capacity expansion on top of Random Serial Dictatorship, and characterize a best-response top-k truthful strategy within a restricted class of list completions (Chapter 3).
- We introduce two learned sequential matching methods for one-sided matching with ordinal preferences: a Generative Flow Network (GFlowNet) policy and a Deep Q-Network (DQN) policy (Chapter 5).

EMPIRICAL CONTRIBUTIONS

- We show on synthetic and Amsterdam school choice data that RM-HAL produces more rank-efficient matches than RSD in 14 out of 16 tested strategic scenarios (Chapter 2).
- We provide simulation evidence on synthetic and Amsterdam-based preference data that the Amsterdam placement guarantee becomes infeasible at the system level when strategic adaptation spreads, and that unequal strategic capacity produces unequal outcomes across groups (Chapter 3).
- Through a behavioral experiment with 140 Amsterdam parents, we find that transparency about strategic risk increases strategic behavior but does not displace truthful reporting as the most common choice (Chapter 4).
- We develop a framework for measuring the robustness of matching methods to preference misreporting through structured misreporting strategies and a population-level adoption experiment, and find that learned methods are notably more stable than classical methods under population-level strategic adoption (Chapter 5).

POLICY CONTRIBUTIONS

- We argue that policy modifications to strategyproof matching methods should be tested against plausible strategic responses before being communicated as public assurances (Chapter 3).
- We challenge the assumption that transparency in school choice inherently encourages manipulation, and argue that communicating strategic risk can serve as a policy lever for fairness in algorithmic allocation systems (Chapter 4).
- We argue that responsible civic AI research requires fluid movement between theory and application, science and society, and quantitative and qualitative inquiry, with civic engagement as a driver of interdisciplinarity (Chapter 6).

CODE CONTRIBUTIONS

- We release the Python implementation for applying the best-response strategy defined in Chapter 2, available at:
`http://github.com/m-tasnim/strategic-rank-minimization`.
- We release the Python implementation for the RSD with placement guarantee simulations reported in Chapter 3, available at:
`http://github.com/JosephTrevorrow/agent-based-policy-interventions`.
- We release the Python implementation for the SP/A model used in Chapter 4, available at:
`http://github.com/m-tasnim/school-choice-survey`.
- We release the Python implementation for the DQN and GFlowNets implementation of matching along with the strategy evaluations reported in Chapter 5, available at:
`http://github.com/m-tasnim/gfn-matcher`.

1.4 LIST OF PUBLICATIONS

PUBLICATIONS FORMING THE CONTENT OF THIS THESIS:

Chapter 2: M. Tasnim, Y. Weesie, S. Ghebreab & M. Baak. *Strategic manipulation of preferences in the rank minimization mechanism.* Autonomous Agents and Multi-Agent Systems **38**, 44 (2024)

M. Tasnim	All aspects
Y. Weesie	Support with formalization
S. Ghebreab	Supervision and insight
M. Baak	Conceptualization, supervision, insight and writing

Chapter 3: M. Tasnim, J. Trevorrow, S. Ghebreab & N. Ajmeri. *The Promise Gap in Algorithmic School Choice: How Placement Guarantees Fail under Strategic Adaptation.* Under Submission (2026)

M. Tasnim	All aspects
J. Trevorrow	Simulation framework and writing
S. Ghebreab	Supervision and insight
N. Ajmeri	Supervision and insight

Chapter 4: M. Tasnim, P. Verhagen, T. Blanke, E. Acar & S. Ghebreab. *Modeling Strategic Risk in School Choice: A Case for Transparent Design.* in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* **8** (2025), 2470

M. Tasnim	All aspects
P. Verhagen	Survey support and writing
T. Blanke	Supervision and insight
E. Acar	Supervision and insight
S. Ghebreab	Supervision and insight

Chapter 5: M. Tasnim, S. Ghebreab & E. Acar. *Sequential Learning for One-Sided Matching under Preference Misreporting.* Under Submission (2026)

M. Tasnim	All aspects
S. Ghebreab	Supervision and insight
E. Acar	Conceptualization, supervision and insight

Chapter 6: M. Tasnim & S. Ghebreab. *AI, interdisciplinarity and civic engagement: insights from Amsterdam school choice.* Cambridge Forum on AI: Culture and Society **1**, e5 (2025)

M. Tasnim	All aspects
S. Ghebreab	Conceptualization, supervision, insight and writing

FURTHER PUBLICATIONS:

1. **M. Tasnim**, D. Michailidis, S. Ghebreab, F. P. Santos & E. Acar. *Mitigating School Segregation through Targeted School Relocation*. Workshop on Citizen-Centric Multiagent Systems (C-MAS), 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS) (2024).
2. D. Michailidis, **M. Tasnim**, S. Ghebreab & F. P. Santos. *Tackling school segregation with transportation network interventions: an agent-based modelling approach*. Autonomous Agents and Multi-Agent Systems **38**, 22 (2024).
3. B. Fabris & **M. Tasnim**. *Fairness inside and out: A situated approach to algorithmic allocation in complex sociotechnical systems*. AIPEX Contributions, AI Policy Lab (2026).
4. T. Portegies, L. Willems, M. Harbers, G. Sileno, R. van Dierendonck, **M. Tasnim**, L. Willemsen & S. Ghebreab. *AI From the Margins (AIM): Rethinking Participatory AI Design Through the Lived Experience of Minoritized Communities*. Under Submission (2026).

2

STRATEGIC REPORTING IN RANK MINIMIZATION

Based on
Tasnim, M., Weesie, Y., Ghebreab, S., Baak, M.
Strategic manipulation of preferences in the rank minimization mechanism.
Autonomous Agents and Multi-Agent Systems **38**, 44 (2024) [24].

Allocation mechanisms used in practice often face a tension between efficiency and strategyproofness. Random Serial Dictatorship is strategyproof but produces inefficient rank distributions, with a long tail of unfavorable allocations. Rank-minimizing mechanisms have been proposed as a welfare-oriented alternative, and recent work has argued that although they are manipulable in principle, the risk is mild: such mechanisms are *not obviously manipulable*, and empirical studies on uniform markets have found little to gain from misreporting. If this view is correct, rank minimization would be an attractive replacement for RSD. If it is not, then moving away from strategyproofness exposes participants to risks that need to be characterized before such a mechanism is adopted. This motivates the first research question of this thesis:

Research Question 1: *If a more welfare-oriented matching rule is used instead of a strategyproof one, what strategies become available, how severe is the resulting risk, and to whom?*

We address this by studying rank minimization implemented through the Hungarian algorithm (RM-HAL). We derive a best-response strategy for an agent aiming to secure one of their top-n preferred items under a simplified cost function, and extend it to a first-order heuristic strategy for the linear cost function that minimizes average rank. The heuristic requires only knowledge of which items are popular as first choices, not the full preference profile assumed in prior impossibility arguments. Using both synthetic markets and real preference data from secondary school choice in Amsterdam, we then study the empirical impact of this strategy across variations in item popularity, the fraction of strategic agents, and the value of n. We find that RM-HAL continues to produce more rank-efficient matches than RSD across most strategic scenarios, but that the strategy is easy to apply and creates a clear individual incentive to misreport. This also creates the risk of a distributional gap between agents who can act on it and those who cannot: in experiments where half of all agents strategize and the rest report truthfully, strategic agents consistently receive better matches than truthful ones across nearly all markets studied.

2.1 INTRODUCTION

Matching agents to items given agents' preferences is an essential problem with solutions existing across economics, mathematics and computer science. Its applications include admitting students to schools, allocating social housing, and matching refugees to hosts [25–27]. Deferred Acceptance with Single Tie-Breaking (DA-STB) is the most well-known matching algorithm, providing stable, envy-free and Pareto-optimal matchings for two-sided preferences [28]. When preferences are one-sided, and a random lottery is performed to simulate preferences from the other side, DA-STB reduces to Random Serial Dictatorship (RSD), and matches are no longer efficient [1, 3]. This inefficiency is reflected in the rank distributions [29] commonly reported by institutions that apply matching mechanisms [5, 30]. In particular, it is observed that RSD results in fewer agents being matched to their top preferred items, with the distribution ending in a long tail of unfavorable rankings, as seen in an example from Amsterdam school choice data in Figure 2.1.

Recent works have argued in favour of matching mechanisms that reduce the maximum rank obtained by agents.

In particular, the rank-minimizing (RM) mechanism [7, 29, 31, 32] has been proposed, which minimizes the average rank received by all agents.

Assuming n agents and n items, on average an agent receives matches that are ranked 2nd or better from RM [7], while RSD provides matches that are ranked $\log(n)$ on average, growing larger with increasing market size [33, 34].

Despite such gains in efficiency, the drawback of implementing RM in the real world is that it is not strategyproof, and agents can receive better matches by misreporting their preferences [31, 32]. Strategyproofness is considered desirable in a matching mechanism as it helps collect truthful preferences of agents and avoids inequality in outcome between truthful and strategic agents. This is a very strict property, however, as with one-sided preferences it is impossible for a mechanism to provide more efficient matches compared to RSD without being vulnerable to manipulation [35].

Troyan & Morrill introduce the notion of not obviously manipulable (NOM) mechanisms, where it is not obvious to the agent how to misreport preferences to gain beneficial matches [36]. Troyan shows that although RM is manipulable, it is not an *obviously* manipulable mechanism [32], as no single strategy ensures beneficial gains over being truthful without complete knowledge of all other agents' preferences; suggesting that the shortcoming of non-strategyproofness in RM may not be so severe. Ortega & Klein show through an empirical study that when agents are strategic using i.i.d. preferences, they do not stand to gain significantly better allocations [7]. However, this study assumes that agents misreport preferences uniformly, and is based on a simplified market with a uniform distribution of preferences over items.

In this work, we argue that agents can be strategic despite not having complete information of others' preferences, and strategies can follow specific patterns. Agents adopt heuristic strategies under uncertain conditions, as demonstrated by several works in strategic voting [37, 38].

Furthermore, markets with a skewed distribution of popularity over items can cause competition among agents, encouraging them to misreport preferences.

This work implements rank minimization using the well-known Hungarian Algorithm (RM-HAL), where costs are considered to be linear over rank. In order to

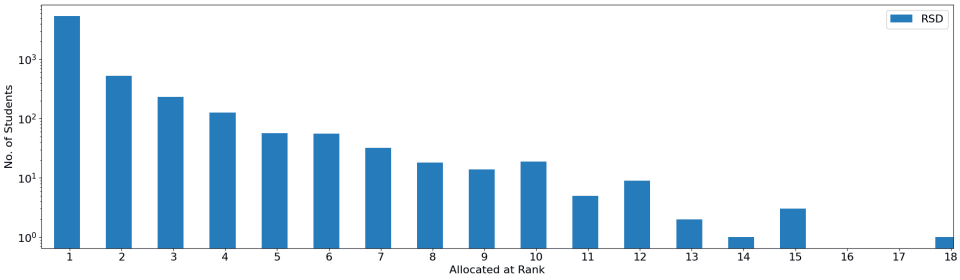


FIGURE 2.1: Rank distribution of RSD matches using Amsterdam school choice data and the original lotteries over students. The randomness of the lottery results in a long tail in the distribution, visualized using a logarithmic y-axis.

better understand the risk of adopting the RM-HAL mechanism, we need to study the impact of heuristic strategies a rational agent may apply, and how this impacts different markets. We ask the following research questions:

1. **RQ1.** Assuming an agent is aware of the top ranked preferences of others, is there a best response strategy for the agent to apply against the RM-HAL mechanism?

2. **RQ2.** What heuristic strategies can an agent apply when they are unable to find a best response to be matched to their top-n items?

3. **RQ3.** For markets with varying degrees of demand over items, what is the impact on matching efficiency when agents apply heuristic strategies?

Towards answering the research questions, the following contributions are made.

- We consider a scenario where matching is performed by RM-HAL and agents find only their top-n preferred items acceptable. We derive a best response strategy for agents in this scenario using a simple one-step cost function.
- We extend this best response strategy to a first-order heuristic strategy for a linear cost function that minimizes the average rank.
- Using synthetically generated data and real data from Amsterdam school choice, we study the impact of strategic preferences on the matching performance of RM-HAL for different market configurations. Particularly we observe how many and which agents do better and worse under different fractions of strategizers, top-n preferred items and market configurations.

The choice of RM-HAL over RSD relies on the principle that the efficiency loss caused by random priorities outweighs the efficiency loss caused by the strategic manipulation of a rank-minimizing mechanism.

The study on synthetic data shows that in markets where agents have an evenly distributed demand for items, RM-HAL provides better agent welfare regardless of the number of strategic agents. The study on Amsterdam school data finds that for cases where the number of strategic agents is limited, or where agents are risk-averse in the strategies they adopt, RM-HAL still provides better welfare than RSD. The performance for RM-HAL only becomes worse when the heuristic strategy is applied by a large fraction of agents all aiming to be matched to the top 1 or 2 preferred items.

While RM-HAL is not obviously manipulable, using the first-order heuristic strategy we find it can still be manipulated easily and effectively; in a manner that does not require complete information of all agents' preferences, only the items-list ordered by first-choice popularity. However, given the rank efficiency of RM-HAL under most configurations of strategic scenarios, we concur with [7] that RM is a mechanism worthy of further investigation for practical market design.

2.2 RELATED WORK

So far relatively little has been written in the literature about rank-minimizing mechanisms. An earlier work by Aksoy *et al.* [31] applies the Hungarian algorithm to provide matches by minimizing the total cost of assignment, where the cost is defined by transforming ordinal preference lists to cardinal utilities.

Featherstone [29] is the first to formalize the notion of *rank efficiency* and propose a mechanism that provides rank efficient matches, also showing that rank efficiency is incompatible with strategyproofness. Troyan [32] shows that RM is a *not obviously manipulable* mechanism, as agents cannot find a manipulation that will always result in being matched to a better rank as opposed to remaining truthful. Most recently Ortega & Klein [7] have shown that imposing strategyproofness in matching mechanisms can negatively impact efficiency, particularly when compared to rank efficient matches from RM.

Traditionally, random serial dictatorship (RSD) is favored for one-sided matching tasks due to desirable properties like strategyproofness, ex-post efficiency and equal treatment of equals [1, 35]. Strategyproofness is a desirable property when designing matching mechanisms as truthful preference submission is then the optimal strategy for participants [39–41]. Pathak [42] investigates the motivation behind strategyproofness as a design choice. This study highlights the inequality between truthful and strategic agents introduced by non-strategyproof mechanisms, where even a small fraction of participants manipulating their preferences can result in significant loss for the truthful ones [42]. However, an example from the city of Chicago shows that a midstream change from a highly manipulable mechanism to a less manipulable mechanism failed to induce a majority of students to change their reported preferences. Since only a small fraction (11%) of applicants responded to the change in incentive in the manipulable mechanism, the authors question if manipulability should be a factor in the design of matching mechanisms [43].

The impact of strategic preferences on Probabilistic Serial and Boston mechanism is empirically studied in [44–46]. These studies show that non-strategyproof mechanisms create an inequality in outcome between students who strategize versus those who are truthful [45], however, also that truth-telling can be a weakly dominant strategy

for these mechanisms with an increasing supply of school capacity [44]. Furthermore, even when full information about the mechanism and the preferences of others are known, many students still choose not to strategize [46].

An experimental study by Pais & Pintér [46] suggests a positive relationship between strategic manipulation of preferences and the level of information offered to the agent about the allocation mechanism. That is, the more information agents have access to about the preferences of others, the more likely they are to lie about their preferences in order to achieve a better outcome – even for strategyproof mechanisms. Similar results are shown by Calsamiglia, Haeringer & Klijn [47], where 34.7–44.4% of agents remain truthful for the Boston mechanism in different settings. In summary, it is seen that large fractions of agents are willing to strategize, albeit not everyone.

A well-known concept in one-sided matching problems is rank-maximality [48, 49]. A rank-maximal matching results in the maximum number of students being matched to their first choice school, and subject to this condition, the maximum number of students are matched to their second choice school, and so on. While rank-maximal matchings maximizes the number of students matched to their top preferred schools, this is different from minimizing the average rank of all matches, as shown in [29, 31]. The latter also focuses on minimizing the number of students being matched to less preferred schools.

Toward exploring preference manipulations in RM, an empirical study in [7] analyzes the impact of strategic preferences on the matching performance of the RM mechanism, measured by average rank. Two manipulations are considered; students rank last on their submitted preferences: (i) the school they would have been assigned to without manipulation, or (ii) their true first preference. The empirical results using data from schools in Budapest show that these manipulations do not significantly affect the average rank of RM matches. However, the manipulations studied use two allocation rounds, and are only applied to students who have initially received unfavorable allocations – something impossible in practice as submitted preferences cannot be updated. The authors also comment that the scope of manipulations individuals might employ against the RM mechanism is an open research question.

No other work studies possible preference manipulations in RM. Our work addresses this research gap by implementing RM using the Hungarian algorithm (RM-HAL) and proposing strategies agents may apply by analyzing its matching procedure.

2.3 RANK MINIMIZATION USING HUNGARIAN ALGORITHM (RM-HAL)

The one-sided matching problem for multi-unit items is defined as follows. Let S and M be sets denoting agents and items respectively. Individual agents and items are indexed by i and j and each multi-unit item has capacity to admit k_j agents. Let the total capacity of all items be $t = \sum_{\forall j \in M} k_j$.

A matching $\mathcal{A} : S \to M$ is an assignment of all agents to the available items. For a valid allocation $\mathcal{A}$, every agent is assigned to only one item, and each item cannot be assigned more agents than its capacity.

In one-sided matching, only agent preferences over items are considered. Each agent ranks available items in a personal preference list, $p_i = \{p_{i,j}\}$, where each item appears at most once and $p_{i,j} \in M$ is the jth preference of agent i. The ordering is strict, with $p_{i,j} \succ p_{i,j+1}$, meaning there is no indifference in preferences.

In this work, we consider incomplete preference lists with a fixed length l, i.e., $|p_i| = l$ where $l < |M|$. As preferences are incomplete, our problem allows for an agent to not be assigned to any item. To model this, a sink item $m_{sink} \in M$ with capacity $k_{sink} = |S|$ is added at the end of every preference list as the lth item. Assignment to m_{sink} represents not being matched to any item. Preference lists with a variable length are not considered in this work.

Let $r_{ij} \in \mathbb{Z}^+$ be the linearly increasing rank at which item j appears in the preference list of agent i. The objective of the rank-minimizing (RM) mechanism is to find a set of matchings $(i, j) \subset S \times M$ such that it minimizes the average rank at which all agents are matched:

$$\mathcal{A} = \underset{ij}{\operatorname{argmin}} \sum_{\substack{i \in S \\ j \in M}} r_{ij} \tag{2.1}$$

The RM mechanism is implemented using the Hungarian algorithm [50–53], referred to as RM-HAL throughout this work. The Hungarian method considers a cost corresponding to possible matches between agents and items and finds a matching such that the total cost is minimized.

Let $C = (c_{ij})$ be a square cost matrix of size $t \times t$, with $c_{ij} = r_{ij}$. Items not in the preference list of an agent get assigned a constant large cost, larger than the cost of the sink item. The Hungarian algorithm accepts square matrices only, and therefore when $t > |S|$, the cost matrix C is made square by adding rows of zeros. [1]

As $c_{ij} = r_{ij}$, applying the Hungarian algorithm over C provides a matching that minimizes the average rank.

The Hungarian algorithm is well-known for its simplicity, availability and reproducibility. Further reasons behind choosing this algorithm for rank minimization are discussed in Appendix A.1. The steps of the Hungarian algorithm are discussed and exemplified below, such that insights can be derived on how to submit preferences strategically against this algorithm in the context of matching. These insights are further applied in Sec. 2.4.

2.3.1 *Algorithm Description*

The Hungarian algorithm is summarized in Algorithm 1, which calls two sub-routines Algorithm 2 and 3 respectively. Matching $\mathcal{A}$ is created through iteratively finding a set of *independent zeros* equal to the number of agents in the cost matrix C.

If C is a matrix with real entries, a set of independent zeros is given by a set of ordered pairs such that none of the zeros in the set are in the same row or column. A method for finding the maximal set of independent zeros is given in Algorithm 2. The set of independent zeros is obtained by iterating over rows in C, sorted from the fewest to the highest number of zeros per row and column (lines 3–7).

Algorithm 1 contains a cost reduction step (lines 8–15) where the minimum uncovered element is subtracted from all uncovered elements (lines 9–12). This step is repeated until the number of independent zeros and agents are equal, at which point all agents can be matched.

1 These represent fake agents who can be matched to any item without cost. They have no impact on the matching in practice.

Algorithm 1: Hungarian Algorithm (Kuhn-Munkres)

procedure HUNGARIAN(C) ▷ Input: Cost matrix C of size $t \times t$
  for every row i and column j in C **do**
    $C_i = C_i - \texttt{min}(C_i)$ ▷ Subtract row and column minimum
    $C_j = C_j - \texttt{min}(C_j)$
  end for
  $\mathbf{Z} \leftarrow \texttt{MaxZ}(C)$ ▷ Get maximal independent zeros
  $\mathbf{cx}, \mathbf{cy} \leftarrow \texttt{MinCover(C, Z)}$ ▷ Get index of covered rows and columns
  while $\texttt{len}(\mathbf{cx}) + \texttt{len}(\mathbf{cy}) < n$ **do**
    minVal $\leftarrow \texttt{min}(C_{ij})$ such that $i \notin$ cx and $j \notin$ cy
    for every row $i \notin \mathbf{cx}$ and column $j \notin \mathbf{cy}$ **do**
      $C_{ij} = C_{ij} -$ **minVal**
    end for ▷ Subtract minimum value from uncovered elements
    $\mathbf{Z} \leftarrow \texttt{MaxZ}(C)$
    $\mathbf{cx}, \mathbf{cy} \leftarrow \texttt{MinCover(C, Z)}$
  end while
  Output: Set Z containing index (i, j) of assignments
end procedure

Algorithm 2: Find Maximal Set of Independent Zeros

procedure MAXZ(C) ▷ Input: Cost matrix C of size $t \times t$
  $Z \leftarrow \{\}$
  while $0 \in C_{i,j} \forall (i, j) \notin Z$ **do**
    $x \leftarrow$ row such that C_x has the fewest marked 0 elements
    $y \leftarrow$ column such that $C_{x,y} = 0$ and C_y has fewest marked 0 elements
    Add (x, y) to set Z ▷ Mark $C_{x,y}$ as an independent zero
  end while
  Output: Set Z containing independent zeros
end procedure

Theorem 10.1 in [53] states that the maximum number of independent zeros in C is equal to the minimum number of lines through rows and columns that cover all zeros of C. The Hungarian algorithm employs Algorithm 3 to cover all zeros in C using the minimum number of horizontal or vertical lines. For a strategic agent trying to find the best order in which to state preferences, the placement of these horizontal and vertical lines is crucial to the final assignment. This is further discussed in Section 2.3.3.

In setting the horizontal and vertical lines that cover all zeros, the set of uncovered rows is the complement of all covered rows, and consists of all agents who are competing for a limited set of spots at possibly multiple items. Their objective is to achieve a spot among the limited set of independent zeros among the uncovered rows. This is illustrated in Section 2.3.2 using an example.

Algorithm 3: Cover all zeros in matrix with minimum lines

```
1: procedure MinCover(C, Z)                    ▷ Input: Cost matrix C and Set Z
2:     coveredRows ← {}                        ▷ Rows covered with horizontal lines
3:     coveredCols ← {}                        ▷ Columns covered with vertical lines
4:     while C_{i,j} = 0 and i ∉ coveredRows do
5:         if 0 ∈ C_i and i ∈ Z then           ▷ Row i contains independent zero
6:             coveredRows ← i                 ▷ Cover row i
7:         end if
8:         if C_{i,j} = 0 and i ∉ coveredRows then   ▷ 0 exists in uncovered row
9:             coveredCols ← j                 ▷ Cover column j
10:        end if
11:        for (i, j) ∈ Z do                   ▷ For all independent zeros
12:            if j ∈ coveredCols then
13:                Remove i from coveredRows
14:            end if
15:        end for
16:    end while
17:    Output: coveredRows, coveredCols
18: end procedure
```

2.3.2 *School Match Example*

The effect of strategic preferences in RM-HAL is illustrated using a school matching example. Consider three schools A, B, C, and seven students $S_1 \dots S_7$. The preferences of the students over the schools are as follows:

$$\begin{aligned} S_1 &: B \succ ? \succ ? \\ S_2 &: A \succ B \succ C \\ S_3 &: A \succ C \succ B \\ S_4 &: A \succ B \succ C \\ S_5 &: B \succ C \succ A \\ S_6 &: B \succ A \succ C \\ S_7 &: C \succ B \succ A \end{aligned}$$

Here the school capacities are $|A|=2$, $|B|=2$ and $|C|=3$.

Figure 2.2a shows the cost matrix for all students. Costs correspond to the integer rank r_{ij} at which student i prefers school j, as defined in Section 2.3. Student S_1's ranks for schools A and C are unknown and are represented by r_a and r_c respectively. The goal of student S_1 is to rank A and C in such a way as to ensure placement at the preferred school B. The sink item is excluded for brevity. [2]

In Figure 2.2a, the row and column minimum values have already been subtracted according to Algorithm 1 lines 2–4. Note that each school has duplicate columns,

2 Including the sink item and corresponding fake agents will not alter the allocation shown in this example.

representing each unit of its capacity. As a result, the zeros of the cost matrix have a relatively simple block structure.

Schools A and B are popular schools, both having a first-choice demand larger than its capacity. In Figure 2.2b, a minimal number of four vertical lines and one horizontal line covers all zeros of the matrix. Not all students can be matched yet, so a cost reduction step needs to be applied. The minimum value of 1 gets subtracted from the uncovered elements in the columns C_1, C_2, C_3.

This results in the matrix shown in Figure 2.2c, where all zeros can now be covered by seven lines: the same five lines and two new horizontal lines through the new zeros created in C_1, C_2, C_3. Now all seven students can get matched, as seen in matrix 4, and the minimum cost value has been reached.

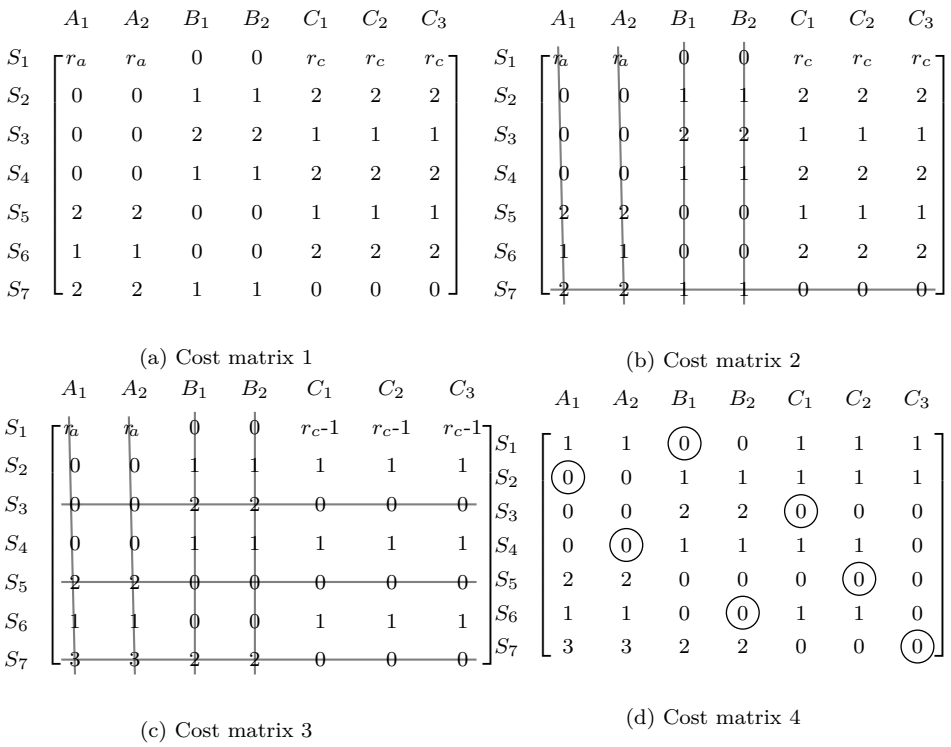


FIGURE 2.2: Application of the Hungarian algorithm to an example cost matrix. The preferences of student S_1 are chosen to ensure placement at school B.

For student S_1, their best-response strategy needs to ensure that no additional zeros are created in its row, besides those already at B_1 and B_2, thereby guaranteeing their match to school B. The unknown cost values r_a at A_1 and A_2 do not get reduced in the cost reduction step; they are covered with vertical lines as the first-choice demand for school A is larger than its capacity. The unknown cost value r_c at C_1, C_2 and C_3 gets reduced by one. It needs to have the value of 2 in order to remain non-zero after the cost reduction step. This leaves the value of 1 for r_a, resulting in the preference

list for student S_1: $B \succ A \succ C$. This ordering ensures that additional zeros are not created in the row of student S_1.

In Figure 2.2b the popular schools A and B receive vertical lines, and therefore do not participate in the cost reduction step. The unpopular school C is assigned a high cost value by student S_1, so it would take a higher number of cost reduction steps before it becomes available for being matched to S_1. In the example, S_3 and S_5 get matched to C after one cost reduction step, as it was their second choice school.

Before the cost reduction step, each student has the possibility of being matched to their first-choice school, as that choice always gets assigned zero cost. After the cost reduction step, new zeros are created and students S_3 and S_5 get matched to schools ranked lower in their preference lists. This does not happen in the row of student S_1, keeping only the first-choice school as a potential match.

2.3.3 *Insights on Strategy*

There are two main insights from the Hungarian algorithm and the minimal covering algorithm related to strategy. These two insights are applied in Sec. 2.4 to form strategies to manipulate RM-HAL.

Firstly, when trying to ensure allocation to top-n preferred items there should be as many zeros as possible from other agents in all columns but the columns of the first n preferences, such that those columns (items) become covered during the minimal covering subroutine. If covered, these elements cannot become zero in the cost reduction step of the Hungarian algorithm and the agent thus avoids the risk of being assigned to these items in the (final) assignment. This insight is used in the proof of the best response strategy of Sec. 2.4.1.

Secondly, when an item outside the top-n preferences does not get covered by vertical lines, it is imperative that the cost for being assigned to that item is higher than that of other uncovered items, such that it minimizes the chance of becoming zero through iterations of the cost reduction step. Each iteration lowers the cost, so the higher the initial cost is, the better. This insight used in the heuristic strategy of Sec. 2.4.2.

2.4 STRATEGIES FOR RM-HAL

Strategies to manipulate RM-HAL are considered for a scenario where being matched to only the top-n ($n \in \mathbb{Z}^+$) preferred items is deemed acceptable by an agent.

First, a best response strategy is derived for matches using a simplified one-step cost function, where possible assignments either have a high or a low cost. This finding is then extended to a first-order heuristic strategy for RM-HAL where the cost of each possible assignment increases linearly with the agent's rank of items.

Let allocation using the Hungarian algorithm be a normal-form game (S, P, u), where S represents all agents undergoing the matching process, P represents the set of all possible ordered preference lists agents can submit, and u represents the set of all payoff functions from a Hungarian assignment such that $u_i : P \to \mathbb{R}$ for any agent i. Note that P contains permutations of all subsets of items M.

Then $q = (q_1, q_2, \ldots, q_n)$ denotes a pure strategy profile containing preference submissions from all agents such that $q_i \in P$.

For a single agent i, let $q_{-i} = (q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n)$ contain the strategy of all other agents except i. Therefore, $q = (q_i, q_{-i})$.

Assuming q_{-i} is known, the best response strategy for agent i corresponds to a strategy $q_i{}^*$ such that $u_i(q_i{}^*, q_-i) \geq u_i(q_i, q_{-i})$ for all $q_i \in P$.

In the scenario where agent i only cares about being allocated to items among their top-n preferences, and is indifferent among those n items, their individual payoff function becomes:

$$u_i(n) = \begin{cases} 1 & \text{if } \mathcal{A}(i) \in \{p_{i,1}, \ldots, p_{i,n}\}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

That is, agent i receives a positive payoff only when they are allocated to their top-n preferred items. Note that this payoff function is based on how an individual agent values their matched item; this is distinct from the cost function in Eqn. 2.1 used during RM-HAL. The goal of the strategic agent i is to ensure the costs corresponding to their top-n preferences, encoded in row C_i of cost matrix C, are a part of the set of independent zeros created at the final step of the Hungarian algorithm. A trivial way to ensure this is to state top-n preferences truthfully so that these items are assigned the lowest initial costs.

Each iteration of the cost reduction step of the Hungarian algorithm reduces cost values and potentially creates new zeros in C_i, particularly in columns that do not yet have vertical lines covering them. Thereby each iteration increases the chance that agent i is assigned to a item lower on their preferences list. The goal of the best response strategy is to prevent this from happening.

Two scenarios are considered: matching agents to their top-n preferred items using either a one-step cost function, or the linear cost function of Eqn. 2.1. The first scenario permits the derivation of a best-response strategy and proof of the corresponding Nash equilibrium. This best-response strategy serves as input and inspiration for a heuristic response strategy for a top-n allocation using the linear cost function.

2.4.1 *Best Response Strategy for One-Step Cost Function*

A simple one-step cost function is considered to study the best response strategy for being matched to the top preferred item. The one-step cost function is the inverse of the payoff function $u_i(n)$ described in Eqn. 2.2. In this case, the corresponding single-step cost function assigns zero cost to all agents' top-n preferred items, and a constant cost to other submitted items, including the sink item. Unpreferred items get assigned a constant, large penalty value.

By construction, this function results in at most one iteration of the reduction step of the Hungarian algorithm.

The game played is assumed to be one of complete information up to top-n preferences, i.e., the strategizing agent i knows all top-n preferred items in q_{-i}. Every preference list has a fixed length l, with $l > n$. Each preference list is padded with a sink item, which is assigned a cost value of $l+1$. The sink item has a large capacity, i.e.,

enough units to be matched to all agents who cannot be matched to their submitted preferences.

The popularity of an item, $f_{j,n}$, is defined as:

$$f_{j,n} = \sum_i \sum_n [p_{i,n} = j] - k_j \, .$$

The n-rank popularity for item j, $f_{j,n}$, is denoted by the difference between the total number of agents picking j within their top-n ranked choices and the capacity of j. Items with $f_{j,n} > 0$ are considered popular, those with $f_{j,n} \leq 0$ are less popular. Let N be the set of all popular items.

This leads to the following theorem and corollary:

Theorem. *Under complete information of q_{-i}, with cost function $u_i(n)$ and $|N| \geq l$, the best response strategy for agent i to be matched to their top-n preferred items with RM-HAL is $(p_{i,1}, \dots p_{i,n}, m_{i,1}, ..., m_{i,l-n})$, where item $m_{i,j}$ has $f_{j,n} > 0$ and $m_{i,j} \neq p_{i,n}$.*

The proof consists of three steps, related to the minimal covering subroutine of the Hungarian algorithm (Algorithm 3):

Proof. To maximize the probability of being matched to top-n preferred items, an independent zero in the first n items of preference is desired, and none outside.

1. The column of a popular item ($f_{j,n} > 0$) holds more zeros than the number of columns of that item, as there are more applicants than the item capacity. Therefore, the minimal covering algorithm covers these items with vertical lines. By the same argument, unpopular items ($f_{j,n} \leq 0$) get covered with horizontal lines, as do rows with only zeros.

2. Being covered by vertical lines, the costs for popular items do not become zero in the reduction step of the Hungarian algorithm, which excludes the possibility of being matched to them.

3. When all other items are covered by vertical lines, an agent may still be allocated to the sink item. In this case, the sink item becomes zero through the first cost reduction step.

□

Note that the same outcome is obtained when all agents submit only their top-n true preferences, i.e., $|p_i| = n$, where all agents are allocated to either one of their top-n choices or the sink item. An example application of the best response strategy can be found in Appendix A.2.

Corollary. *A Nash Equilibrium is $\big((p_{1,1}, \dots p_{1,n}, m_{1,1}, ..., m_{1,l-n}), \dots, (p_{|S|,1}, \dots p_{|S|,n}, m_{|S|,1}, ..., m_{|S|,l-n})\big)$, where item $m_{i,j}$ has $f_{j,n} > 0$ and $m_{i,j} \neq p_{i,n}$.*

Proof. Assume that an agent plays any strategy different from the best response.

1. By definition of the one-step cost function, the cost of top-n preferred items is zero, so it is optimal to report top-n items truthfully.

2. Any agents with different strategies than the best response would include items with $f_{j,n} \leq 0$ after the top-n preferred items. These items are not covered by vertical lines in Algorithm 3, thereby creating zeros in items outside the top-n preferences. This decreases the odds of being allocated to the top-n preferred items, making the strategy sub-optimal in terms of a positive payoff.

□

In summary, for the simplified scenario of a single-step cost function, with complete information on all top-n preferences, it is a Nash equilibrium for agents to strategize with only popular items ($f_{j,n} > 0$) beyond their top-n preferences.

2.4.2 *First-order Heuristic Strategy for Linear Cost Function*

Extending the best response strategy of Sec. 2.4.1 to the linear cost function of Eqn. 2.1 is non-trivial. The linear cost function typically requires multiple iterations of the cost reduction step to match all agents. At each iteration, the agent aims to prevent zeros in the cost matrix corresponding to the items outside of their top-n preferences. The number of iterations this takes is dependent on the reported preferences of other agents. In essence, such a strategy requires complete information on all preferences. An example illustrating this is provided towards the end of this section.

Instead, inspired by the best response strategy of Sec. 2.4.1, several heuristics are proposed. The proposed strategy is referred to as the *first-order heuristic* strategy. Complete information up to top-1 preference is assumed. Note that for the linear cost function, only the top-1 preferred item has an initial zero cost, making it optimal to report this item truthfully.

Heuristic. *First-order heuristic strategy:*

1. *Based on the Theorem: beyond top-n, pick items with $f_{j,1} > 0$.*

2. *Sort the items in descending order of popularity. The larger the popularity, the larger the excess of agents needing to be matched elsewhere. Generally, this leads to more cost reduction steps before costs beyond the top-n items become zero.*

3. *Less popular items may be added to the preference list if these are close to being full, e.g., when some items are popular choices as safety items and are ranked lower by many on the preference list. The assumption is that these items may be fully assigned during the cost reduction steps, thus also getting covered by vertical lines.*

The first heuristic prevents zero cost elements being created during the first iteration of the reduction step, as with the single-step cost function, and aims to do so at the

next steps as well. The second heuristic ensures that, if somehow popular items do receive zero cost, their initial cost values are as high as possible. The third heuristic extends the preference list and thus increases the number of Hungarian algorithm iterations, e.g., in order to compete with other strategizing agents.

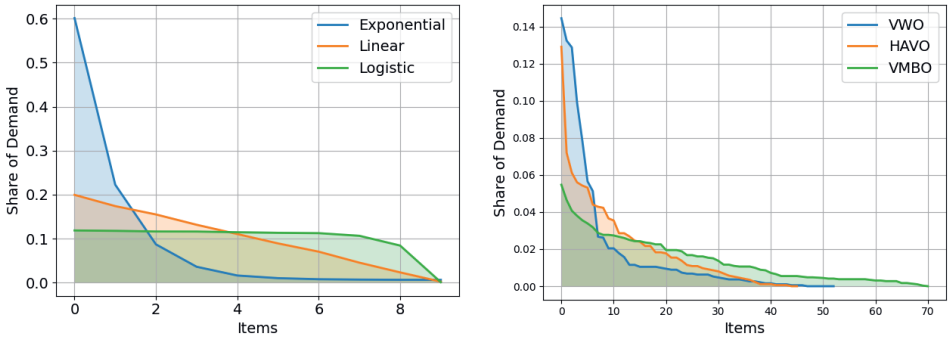
The first-order heuristic strategy is not a best response. To illustrate this, an edge-case example is presented where heuristic 1 fails in the cost reduction step of the Hungarian algorithm.

Let there be three items A, B, and C such that $f_{A,1} > f_{B,1} > 1 > f_{C,1}$ and C has a large enough capacity to admit everyone. As A is the most popular item, for any agent I with B as their true first preference, the first-order heuristic strategy recommends reporting $B \succ A \succ C$. Now imagine a scenario where all agents who pick A as their first choice, have C as their second choice, reporting $A \succ C \succ B$. As C has a large capacity, all these agents are allocated to either A or C after the first iteration of the cost reduction step. At this point the vertical lines passing through A become horizontal, meaning agent I now has a chance to be assigned to A instead of B, making their strategy non-optimal.

In practice, this may happen when item choices are very correlated. The edge-case example shows that picking a popular item (beyond top-n) which is guaranteed to have vertical lines throughout multiple iterations of the Hungarian algorithm requires detailed knowledge of the correlations between item choices. Although this is possible with complete preference information, the best set of items to pick beyond top-n becomes a non-trivial exercise due to these correlations.[3]

2.5 DATASETS

Both synthetic and real data are analyzed in this study, each having three different markets representing different demand for items by agents.



(a) Synthetic data.

(b) Amsterdam data.

FIGURE 2.3: Markets with varying demand over items. Y-axis shows the share of agents who ranked each item as their top preferred item.

3 The derivation of a best-response strategy for RM-HAL is left as future work.

2.5.1 *Synthetic Data*

Three markets are generated with 2000 agents and 10 multi-unit items each with enough units for every agent. Each item has a capacity of 200 units. Each market has a non-uniform demand distribution over items, as shown in Figure 2.3a. The *logistic* market represents a near-uniform market where most items are equally demanded by the agents. The *linear* market presents a scenario where demand for items increases linearly. The *exponential* market reflects a situation where a few popular items are demanded by a large fraction of agents.

2.5.2 *Amsterdam Data*

The real-world preference data used in this work is sourced from the Association of School Boards in Amsterdam Secondary Education (OSVO), which has implemented RSD as the central school choice mechanism since 2016. The dataset contains anonymized preferences of 7.5k agents over 60+ schools from the year 2017.[4] More details on the Amsterdam dataset can be found in Appendix A.4.

The Dutch education system has three levels for secondary education: pre-university education (VWO), senior general secondary education (HAVO) and pre-vocational secondary education (VMBO). The three education levels represent different markets in the context of school choice, with varying levels of demand, as shown in Figure 2.3b. A minimum of 12 entries per preference list is recommended for VWO and HAVO students and 4-6 for VMBO students, although no minimum or maximum length is enforced in practice. For example, some students fill in just one or two preferred schools. Although there is enough capacity for all students in the region, there are multiple popular schools that cannot accept all applicants.

A survey performed by OSVO [14] has measured the percentages of parents or students who find allocation to the top-n school ranks acceptable, for increasing values of n. For example, 63% of parents find allocation to the top-3 ranks acceptable, but this level drops rapidly after that, down to 30% for the top-5. The study records responses from both students and their parents; in this work we use the average of both sets (see Appendix A.3). From the survey, one can deduce the fraction of people who consider allocation to a certain rank n unacceptable, but who accept allocation up to rank $n-1$. For ranks 2-9 the corresponding fractions are $18.5, 22.5, 23, 9, 14, 5, 2, 1.5\%$. These numbers are used in the strategy studies in Sec. 2.6.1.

2.6 EXPERIMENTAL RESULTS

Four experiments are conducted on both Amsterdam and synthetic datasets to assess the impact of the first-order heuristic strategy on the matching performance of RM-HAL. The synthetic dataset simulates different market conditions, including uneven distributions of agents' demand for items. In the first three experiments, the first-order heuristic strategy is applied for various fractions of agents and values of top-n. These settings simulate different scenarios of strategic manipulation, ranging from moderate to aggressive. The matches from RM-HAL are compared against those resulting from

4 The data can be obtained from OSVO upon request.

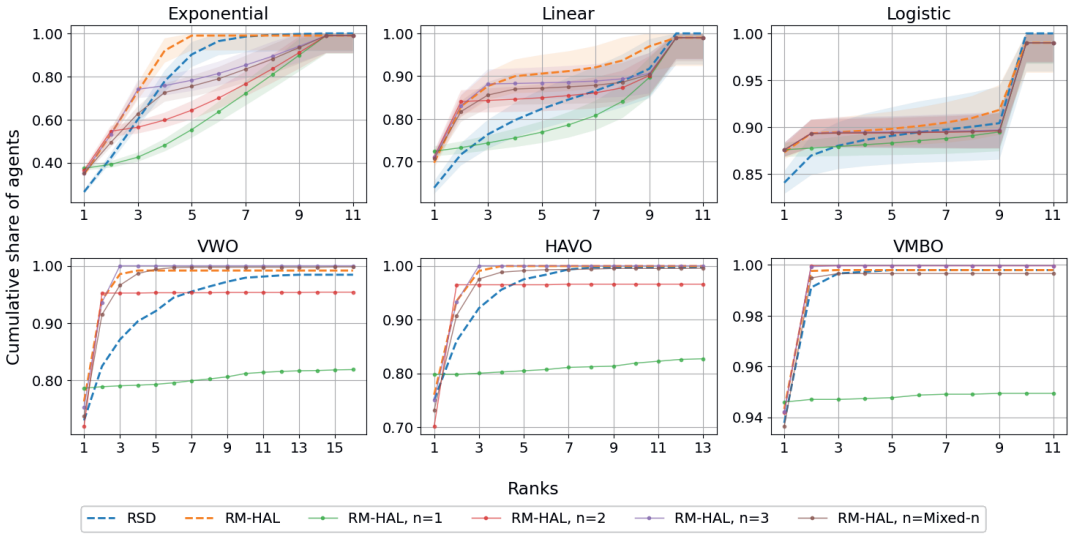


FIGURE 2.4: Experiment 1. Cumulative share of matched agents over ranks in RSD and RM-HAL with strategic behavior over different top-n preferred items. The legend shows matches using different values of top-n, for synthetic (top) and Amsterdam (bottom) markets. The fraction of strategic agents is 100%. Lines not reaching 1.0 indicate agents allocated to the sink item or a item not on the truthful preference list. The shaded uncertainty bands reflect 1-standard deviation variations observed in the results when varying the simulated sample and lottery.

RSD using truthful preferences, as RSD does not incentivize agents to manipulate their preferences. A fourth experiment is performed to measure the impact of strategic preferences on truthful and non-truthful agent groups.

2.6.1 *Setup*

RSD is run multiple times, always using truthful preferences, and the resulting average match is considered as the baseline. For RM-HAL, matching performance is always evaluated w.r.t. truthful preferences of agents, even when agents submit false preferences. Preferences from the datasets in Section 2.5.2 (collected for RSD) are considered truthful and strategic preferences are simulated using the heuristic from Section 2.4. Each preference list is also padded with the sink item, with a capacity equal to the total number of agents, which keeps track of agents not matched to any other item.[5] Agents not matched to any item are tracked for both RM-HAL and RSD.

While simulating strategic preferences the assumption is made that all agents are truthful about their top-n item choices, for various values of $n \geq 1$, as listing them first maximizes the probability to be matched to any of them.

5 This is particularly important for matching in practice using RM-HAL when preference lists do not have a fixed length, as having a short preference list can allow an agent to have an advantage over others. This advantage is removed by the sink item.

Market	Method	Top-1 (%)	Top-3 (%)	Not Placed (%)	Max Rank	Avg Rank
Exponential	RSD	26.36	60.04	0	10	3.09
	RM-HAL	34.57	73.09	1	11	**2.53**
	RM-HAL, n=1	37.36	42.5	1	11	4.72
	RM-HAL, n=3	35.27	74.03	1	11	3.35
Linear	RSD	63.86	76.37	0	10	2.74
	RM-HAL	69.71	87.68	1	11	**2.06**
	RM-HAL, n=1	72.41	74.36	1	11	2.95
	RM-HAL, n=3	70.98	88.17	1	11	2.25
Logistic	RSD	84.05	88	0	10	2.04
	RM-HAL	87.3	89.45	1	11	**1.92**
	RM-HAL, n=1	87.57	87.92	1	11	2.05
	RM-HAL, n=3	87.55	89.45	1	11	1.97

TABLE 2.1: Experiment 1. Matching performance of RM-HAL with truthful and strategic preference submissions in the synthetically generated matching markets. The best average rank achieved for each market is emphasized. The fraction of strategic agents is 100%.

2.6.2 *Impact of strategic preferences on matching performance*

Three experiments are conducted testing different strategic settings. In Experiment 1, n is varied between $[1, 3]$ and a mixed-n scenario is also considered where groups of agents use different values of n. In the mixed-n setting, values of $n \in [1, 3]$ are picked uniformly for synthetic markets, while for Amsterdam values of n are taken from a survey [14]. In Experiment 2, the fraction of strategic agents, denoted by f, is varied between $[25, 50, 75, 100]\%$. In both experiments the parameters n and f are varied only one at a time, and when not varied, default values are set to mixed-n and 100% respectively. In Experiment 3, the values of n and f are varied simultaneously.

Experiment 1: Matching results are shown in Figure 2.4 in the form of cumulative rank distributions, which show for each rank the average share of agents assigned to that rank or better. The top and bottom figures show the synthetic and Amsterdam markets respectively. For synthetic markets averages are computed from 50 distinct generated datasets. The resulting average values are visualized for each synthetic market, including their associated uncertainties (± 1 s.d.). For Amsterdam markets

Market	Method	Top-1 (%)	Top-3 (%)	Not Placed (%)	Placed Elsewhere (%)	Max Rank	Avg Rank
VWO	RSD	73.52	88.85	1.41	0	21	1.87
	RM-HAL	76.56	98.59	0.84	0	5	**1.3**
	RM-HAL, n=1	78.65	78.75	0	17.84	27	3.51
	RM-HAL, n=3	75.2	100	0	0	3	1.31
HAVO	RSD	74.47	90.98	0.46	0	14	1.62
	RM-HAL	75.73	99.2	0.06	0	4	**1.31**
	RM-HAL, n=1	79.78	80.41	0	17.93	22	2.98
	RM-HAL, n=3	74.76	100	0	0	3	1.32
VMBO	RSD	93.68	99.86	0.1	0	5	1.07
	RM-HAL	94.33	99.9	0.1	0	3	**1.06**
	RM-HAL, n=1	94.6	94.74	0	5.13	14	1.33
	RM-HAL, n=3	94.26	99.97	0	0.03	3	**1.06**

TABLE 2.2: Experiment 1. Matching performance of RM-HAL with truthful and strategic preference submissions in the Amsterdam school choice markets. The best average rank achieved for each market is emphasized. The fraction of strategic agents is 100%. "Not Placed" refers to being matched to the sink item, and "Placed elsewhere" to an item on the submitted preference list that is not on the truthful list.

the results shown are the average over the single dataset with 50 different lottery assignments for RSD.

In Figure 2.4, for both synthetic and Amsterdam data it is observed that RM-HAL with truthful preferences has the best matching performance for nearly all markets (see orange dashed lines). This is not unexpected, however particularly for markets with uneven distribution of demand over items (e.g., exponential, linear, VWO and HAVO) RM-HAL with truthful preferences performs significantly better than RSD (blue dashed line), with differences in cumulative share as large as 10% points. In markets where demand is more evenly distributed (e.g., logistic and VMBO) RM-HAL still performs better than RSD, although the difference in matching performance is less visible.

The gain in performance is negatively impacted when agents begin to misreport their preferences. In Figure 2.4 the performance of RM-HAL declines when agents apply the heuristic strategy for decreasing values of n. For example, in all markets matching performance of RM-HAL is worse than RSD when all agents apply the

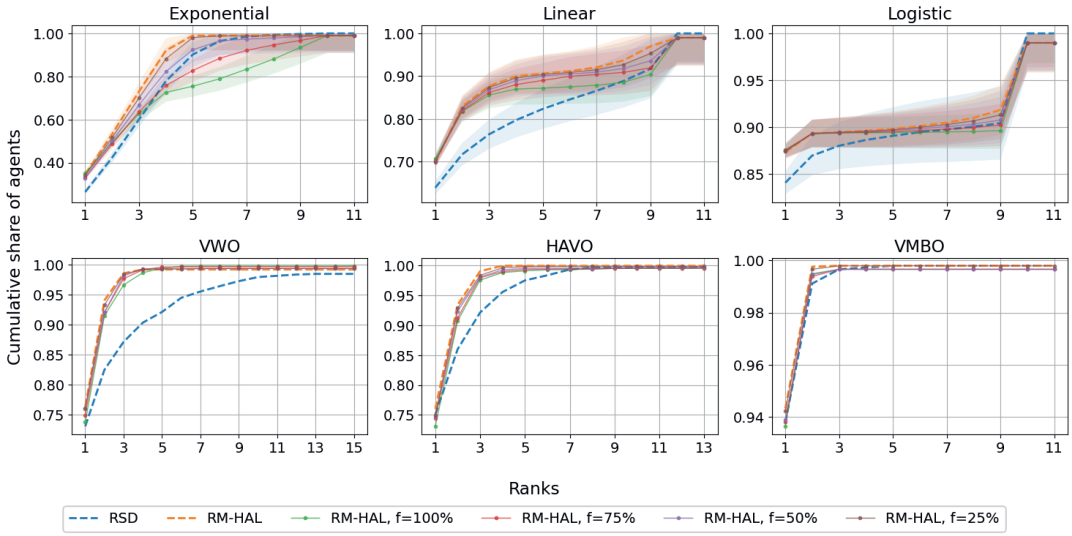


FIGURE 2.5: Experiment 2. Cumulative share of matched agents over ranks in RSD and RM-HAL with strategic behavior over increasing fractions (f) of agents. The legend shows matches using different fractions, for synthetic (top) and Amsterdam (bottom) markets. The top-n value for strategizing agents is mixed.

first-order heuristic strategy at $n = 1$ (represented by the green lines). For some strategic agents this strategy is successful, resulting in a greater share of matched agents at rank 1 than RSD. However, for all other agents this strategy fails, and they are either matched to the sink item or to a item they do not truthfully prefer.

This impact is also seen in Tables 2.1 and 2.2, which summarize the average rank of each match, along with the share of agents matched to a particular rank, for the synthetic and Amsterdam markets respectively. For example, Table 2.1 shows more agents matched to their top-1 item compared to RSD for each strategic scenario. In Table 2.2 the VWO market demonstrates in particular that the first-order heuristic strategy is risky at $n = 1$, with a much higher average rank than RSD (3.51 versus 1.87).

In Figure 2.4, for $n \geq 2$ the matching performance of RM-HAL compared to RSD is market dependent (represented by red and purple solid lines). In the synthetic markets with linear and logistic demands, allocation results are better than RSD. Exponential demand results in worse matches than RSD at ranks greater than two. However, notably for the VWO and HAVO markets from the Amsterdam dataset, $n > 2$ results in better matches than RSD, again with differences of cumulative share of more than 10% at rank two.

For the Amsterdam school data in Table 2.2, note that the maximum rank for RSD is a high as 21 (for VWO). Using RM-HAL with truthful preferences, the maximum rank gets reduced to 5. With RM-HAL, n=3, the maximum rank even gets reduced to 3. This happens because in the Amsterdam markets the items within the top-3

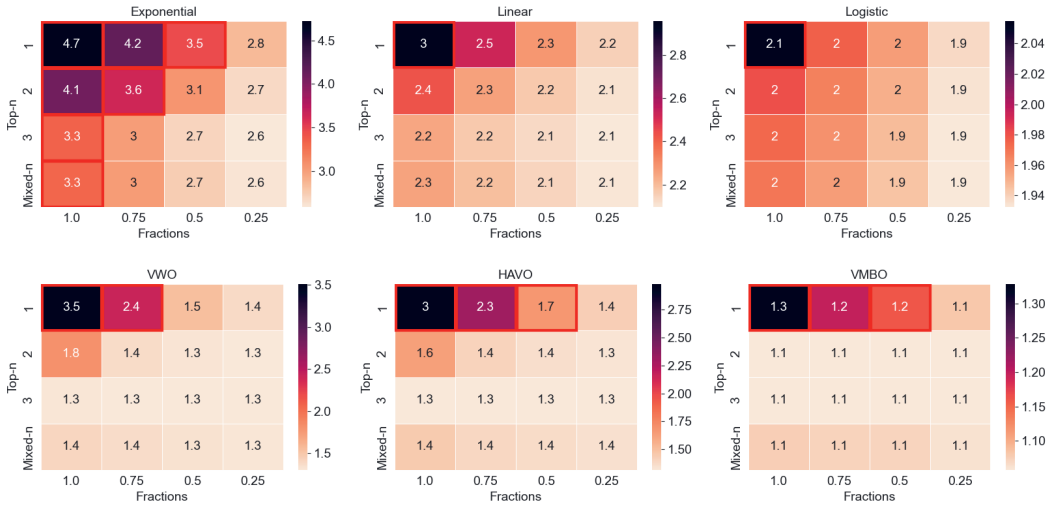


FIGURE 2.6: Experiment 3. Average rank of RM-HAL matchings when agents apply the first-order heuristic strategy, for varying values of both n and f. Scenarios where RM-HAL has a worse average rank than RSD are highlighted in red.

preference of all agents have enough capacity to accommodate everyone.

Experiment 2: The value of f is varied between $[25, 50, 75, 100]\%$, with n set to mixed-n, in order to observe the impact of increasing shares of strategic agents. The corresponding matching results, summarized in Figure 2.5, show that overall matching performance declines when more agents submit strategic preferences. E.g., in the exponential market, matching performance is worse than RSD for $f > 50\%$. However in all other markets, the decline in matching performance is not significant. This can be attributed to the lack of competition in markets such as logistic and VMBO. However in competitive markets like VWO and HAVO, RM-HAL is again able to make more efficient matches as the agents use mixed values of n, and these top-n items collectively have enough capacity to match all agents.

Experiment 3: The average rank is studied for varying values of both n and f, as summarized in Figure 2.6. Scenarios where the average rank from RM-HAL matches is worse than RSD with truthful preferences are outlined in red. For most markets, the matching performance of RM-HAL can be worse than RSD when: a) agents only find top-1 or top-2 items acceptable and b) 50% of agent or more are acting strategic. However, in all other scenarios RM-HAL provides matches with a better average rank. The exponential market is clearly the most extreme scenario, where the competition over items is the highest; even in that market 9 out of 16 scenarios have better average rank from RM-HAL than RSD. Notably in markets using real data from Amsterdam, only 2-3 out of 16 scenarios have worse average ranks than RSD, where all strategizing agents want to be matched their top item.

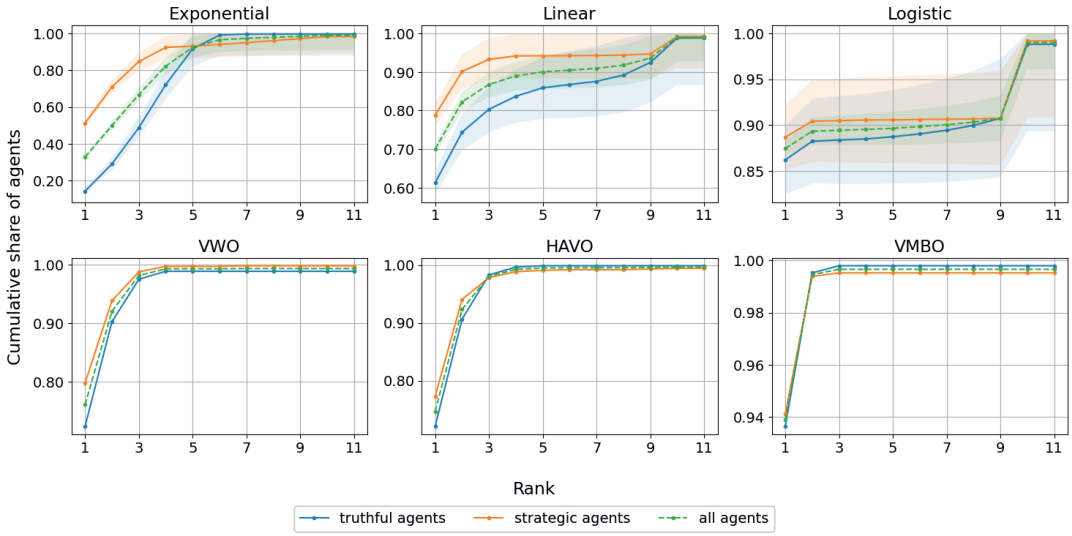


FIGURE 2.7: Experiment 4. Cumulative share of matched agents over ranks in RM-HAL for truthful agents vs. strategic agents applying the first-order heuristic strategy with a mixed top-n preferred items.

This shows that RM-HAL with strategic preferences can provide more rank-efficient matches than RSD with truthful preferences, except for cases when agents display the most extreme strategic behavior.

2.6.3 *Impact of strategic preferences on agent groups*

Experiment 4: Matching performance is observed for half of all agents reporting truthful preferences, and the rest applying the first-order heuristic strategy for n =mixed-n. The results, summarized in Figure 2.7, show that for most markets strategic agents (orange lines) are provided better matches than truthful ones (blue lines).

For instance in the market with exponential demand over items, about 90% of all strategic agents receive better matches until rank 5, but for higher ranks the strategy fails. For the HAVO market this tipping point is observed at rank 3. However, in the VWO market strategic agents consistently receive better matches than truthful agents over all ranks. In the VMBO market, where all items are in near even demand, very little differences are seen between strategic and truthful agents.

These results imply that regardless of demand, agents have an incentive to be strategic in their preference reporting, especially if they know other agents will report truthfully. This finding also holds also for $n = [1, 3]$, as shown in Figure A.2 of Appendix A.5.

2.7 DISCUSSION

Rank minimization directly optimizes the rank distribution commonly reported as default performance metric when applying matching mechanisms. While the mechanism is not strategyproof, Troyan has shown that rank minimization is not obviously manipulable. In addition, according to Ortega & Klein the concern of manipulation is mitigated because the potential gains from misrepresentation are small.

Our results present a more realistic view of the risks in implementing the rank minimization mechanism in practice. Our empirical studies on simulated and real school choice data show that the first-order heuristic strategy for top-n allocation is easy to apply and effective. While not obviously manipulable, our findings show improved top-n allocation probability and imply an incentive for agents to be strategic in their preference reporting, contradicting Ortega & Klein. However, for most strategic scenarios tested RH-HAL still outperforms RSD, confirming it is an appealing mechanism for assignment problems.

With respect to the research questions posed in Section 5.1, our findings are as follows:

RQ1. A best response strategy for RM-HAL has been derived for top-n allocation using a simple one-step cost function.

RQ2. The first-order heuristic strategy has been derived with the assumption of complete information: an agent is aware of the rank 1 demand over all items. The strategy makes sense intuitively. The popular items placed after the first n choices form a high-cost "wall", as these items are already matched to other agents. The first-order heuristic strategy further sorts the popular items in descending popularity, aiming to make it very costly for RM-HAL to not assign agents to their top-n items.

RQ3. Simulations studying the impact of the first-order heuristic strategy on matching efficiency shows that average rank can be negatively affected when large fractions of agents apply the strategy with a goal to be matched to their top-1 or top-2 items, particularly in markets with uneven demand over items.

The matching performance of RM-HAL is observed to be significantly better than RSD when truthful preferences are considered, particularly in markets with uneven demand. In those markets, with more competition over a few items, agents have an incentive to submit strategic preferences to maximize their chances of being matched to top-n preferred items. However, applying the first-order heuristic strategy comes at a risk of not being matched to any item or to an item not truthfully preferred, especially when the acceptable rank of matched item n is limited. This is a natural consequence of the way RM-HAL minimizes the sum of all ranks. The probability of agents applying this strategy therefore depends on individual perception of risk and reward.

In the simulation studies reported in this work, RM-HAL performs better than the baseline of RSD for most strategic scenarios. In use cases where rank-efficient matches are important and agents are not able to be strategic (e.g., they do not have enough information about the preferences of others), RM-HAL is a better option than RSD. However, there is an inherent inequality in matching performance between strategic and truthful agents. Truthful agents are especially affected when more than half of the agents submit strategic preferences.

In real world matching scenarios, this inequality will be reflected between sophisticated agents who submit preferences strategically, and agents who report preferences truthfully due to lack of knowledge or resources to compute the heuristic strategy.

2.8 LIMITATIONS

The best response strategy has been derived for a constrained, simplified scenario: assuming complete information about top preferred items, a one-step cost function, a fixed preference list length, as many popular items as that length, and a version of RM-HAL where a sink item gets added to each preference list. These requirements are artificial and strict, and in practice will often be missing. The best response strategy serves as a stepping stone to the first-order heuristic strategy.

In practice, the first-order heuristic strategy requires an up-to-date ordering of the most popular first-choice items, information that is not always available to agents. However, this ordering may be approximated from historical data, especially in scenarios where matching is performed repeatedly (*e.g.,* school choice). Once known, this information may be passed around, for example, through online forums [54]. As such, in practice the first-order heuristic strategy – or a close approximation thereof – could become available to a wide range of agents.

The heuristic strategy is not necessarily the best response strategy for any top-n allocation. The derivation of the best response strategy valid for any value of top-n is non-trivial due to correlations between preferences, and is left as possible future work.

The simulations explore variations of the first-order heuristic strategy, from taking less to more risk of being matched to a truthfully preferred item. There are multiple untested scenarios, e.g., agents may not want to risk being matched to an item they do not truthfully prefer. Such moderate risk strategies can be followed up in a future work. Also not yet explored are irrational strategies, such as not ranking top preferred items truthfully. In a follow-up work, experiments are conducted with real individuals as in [46, 47] to gain insights into behaviors that occur in practical settings.

Matching performance has been evaluated with respect to truthful preferences. This will not be possible in practice, as only reported preferences are available.

Strategyproofness is an important property for matching as it levels the playing field: agents who lack the information to be strategic against the matching mechanism are disadvantaged compared to those who do not. As such, when aggressive strategies are applied in the rank minimization mechanism it could disproportionately hurt the former group; for example in the school choice context, this could affect students coming from a lower socioeconomic background.

2.9 CONCLUSION

This work explores the trade-off between efficiency and strategyproofness in the context of applying the rank maximization mechanism to allocation problems, and aims to promote discussion on the nature of strategic manipulations and their impact. In particular the inefficiency of RSD matches motivates a more rank efficient mechanism.

Under a simplified, constrained scenario, the best response strategy is derived for an agent to maximize the probability of being assigned to their top preferred item

using RM-HAL. The use of the best response strategy by all agents results in a Nash equilibrium. Based on this finding, a first-order heuristic strategy is derived for maximizing the probability of being matched to top-n preferred items when minimizing the average rank. This strategy is then tested on simulated data and real school choice data from Amsterdam, for several scenarios ranging from moderate to worst-case, and performance results have been compared against RSD.

While RM-HAL is not obviously manipulable, the first-order heuristic strategy is found to be effective and easy to apply. Our findings show improved match probabilities for the top-n preferred items in both simulated and real-world datasets. Although the strategy is not devoid of risks, these results suggest an incentive for agents to employ strategic preference reporting. In summary, contrasting with [7, 32, 36] our findings present a more realistic view of the risks in implementing rank minimization in practice.

While the incentive for agents to strategize is market dependent and should be further studied, this work shows that the rank maximization mechanism provides more rank efficient matches than RSD in moderate strategic situations. For most strategic scenarios tested, performance gains in rank are significant compared to RSD. In the simulated market with the highest skew in popularity, the average rank worsens only when more than half of all students apply the first-order heuristic strategy to protect their top-1 or top-2 choice. On Amsterdam school data this happens when half of the students or more protect their top-1 choice. Notably, on Amsterdam school data, when 100% of all agents protect their top-3 preferences using the first-order heuristic strategy, all students are allocated within the top-3 ranks, down from a maximum rank of 21 for RSD.

3

STRATEGIC REPORTING IN PLACEMENT GUARANTEE

Based on
Tasnim, M., Trevorrow, J., Ghebreab, S., Ajmeri, N.
The Promise Gap in Algorithmic School Choice: How Placement Guarantees Fail under Strategic Adaptation.
Manuscript under submission.

The previous chapter showed that moving from a strategyproof mechanism to a more welfare-oriented one creates exploitable incentives. A natural next question is whether such incentives only arise when policymakers actively choose a manipulable mechanism, or whether they can also emerge inside systems built around a strategyproof core. In practice, mechanisms are rarely deployed in textbook form: local authorities introduce tie-breaking rules, priority categories, and policy fixes intended to address fairness, transparency, or capacity concerns. Each modification is typically justified on its own terms, but the implemented system, rather than the underlying algorithm, is what participants actually face. This motivates our second research question:

Research Question 2: *If the base matching method is strategyproof, can institutional modifications create strategic vulnerabilities?*

We address this through the Amsterdam secondary school choice system, in which Random Serial Dictatorship is paired with a placement guarantee: students who submit a sufficiently long preference list and remain unassigned in the main round are promised placement at a school on their list through a proportional capacity expansion. We characterize a simple reporting strategy for students with fewer sincere preferences than the required list length, in which they report their sincere preferences first and fill the remaining positions with heavily oversubscribed schools, and show that this strategy is a best response within a restricted class of list completions. Using simulations on anonymized Amsterdam preference data and synthetic profiles, we find that as strategic adaptation spreads, the number of guarantee-eligible students can exceed the additional seats the policy creates, rendering the guarantee infeasible at the system level. We further show that the guarantee distributes its benefits unequally: groups with more sincere preferences, or better information about which schools are oversubscribed, are better positioned to exploit the rule. We use the Amsterdam case to articulate a broader *promise gap* in public algorithmic systems: a mismatch between what an implemented system publicly assures and what it can sustain once participants adapt to it.

3.1 INTRODUCTION

School choice systems are widely used to assign students to secondary schools through centralized procedures. A well-known method is the Deferred Acceptance (DA) algorithm, which produces stable allocations and incentivizes students to report preferences truthfully [28, 39, 40, 55]. Random Serial Dictatorship (RSD) is another widely used mechanism when schools have no preferences over students; under suitable conditions, it also incentivizes truthfulness and is ex-post efficient with respect to reported preferences [1]. In practice, however, local authorities rarely implement these mechanisms in textbook form. They introduce random tie-breaking, priority categories, and policy fixes motivated by local constraints and political objectives [56, 57]. Such modifications are intended to adapt the system to local conditions, but can also alter its efficiency and incentives [4, 58]. The relevant object of evaluation is therefore the implemented system, not the textbook algorithm alone.

This chapter studies what happens when a practical modification is introduced to a well-understood allocation algorithm in order to address a real institutional problem. Such modifications are often introduced as safeguards: they are meant to make the system more protective, better aligned with local policy goals, or easier to justify publicly. Yet, when modifications are made to an allocation mechanism, policymakers cannot have confidence that the mechanism retains desirable properties, or that the mechanism is reflective of the policy aims it sets out to achieve.

We use the term *promise gap* to refer to the mismatch between the assurance attached to an implemented algorithmic system and what that system can sustain once participants adapt their behavior to it. The Amsterdam secondary school system provides a concrete case of the promise gap in action. Secondary school places are allocated using RSD combined with a *placement guarantee*. Students submit a ranked list of at least l schools (currently $l = 12$ in Amsterdam). They are first assigned via RSD with fixed capacities and a random order over students. If a student who listed at least l schools remains unassigned after this main round, the system promises a place at one of the schools on their list, implemented by increasing capacities at those schools by a small fraction (currently around 4%) [59]. Public communication emphasises that families can list schools truthfully and that the guarantee protects students from non-placement [59]. At the same time, many families in Amsterdam have fewer sincere preferences than the l required by the matching algorithm; [1] the remaining positions must therefore be filled with schools they know less well. This creates a practical problem and an avenue for strategically misreporting preferences that is not addressed by the public message of straightforward truth-telling. Families who do not have enough sincere preferences must still decide how to complete the required list, which creates scope for strategic responses to the guarantee rule itself [60].

The chapter addresses three questions. First, how does the placement guarantee change incentives for students who do not have enough sincere preferences to fill the required list? Second, under what conditions does strategic adaptation make the guarantee infeasible at the system level? Third, when access to strategic knowledge is unequal, does the same rule generate unequal outcomes across groups?

To answer these questions, we formalize the Amsterdam system and study a strategy in which students with sincere preferences over at most $k < l$ schools report these k

1 Based on a report on school preferences in Amsterdam [14].

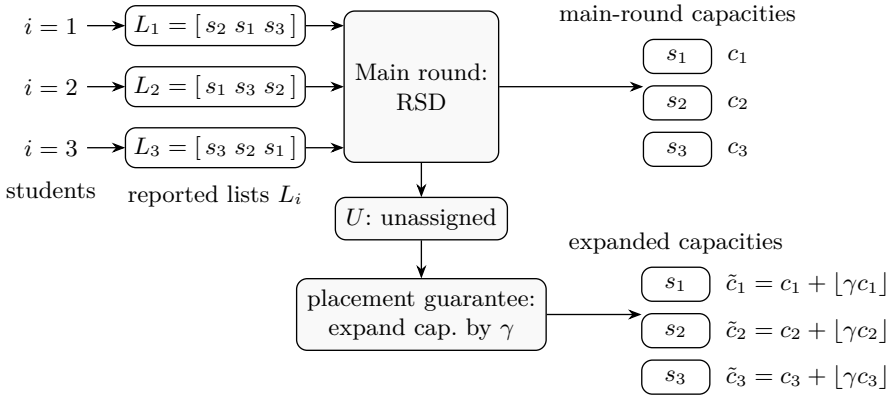


FIGURE 3.1: Conceptual model of school placement under RSD with a placement guarantee. Students submit reported lists L_i, are first assigned via RSD with capacities c_s, and unassigned students enter the placement guarantee stage with expanded capacities $\tilde{c}_s$.

schools truthfully and fill the remaining $l - k$ positions with highly demanded schools that are known to be heavily oversubscribed. Under mild assumptions, we show that this strategy is a best response within a restricted class of list completions that preserve the sincere top k schools. We then study the population-level consequences of this behavior in an agent-based simulation framework using both synthetic preference profiles and anonymized Amsterdam data. At the system level, we examine how feasibility varies with the number of sincere preferences, k, and the fraction of strategic agents, α. At the group level, we examine how unequal strategic capacity produces unequal outcomes when groups differ either in how many schools they can rank sincerely or in how accurately they identify popular schools.

Our contributions are threefold:

- We formalize how the Amsterdam placement guarantee changes incentives and identify a simple profitable reporting strategy for students with fewer than l sincere preferences.
- We evaluate the system-level consequences of this strategy using simulations on synthetic and Amsterdam-based preference data, and identify regimes in which the placement guarantee becomes infeasible because the number of guarantee-eligible students exceeds the additional seats created.
- We show that unequal strategic capacity, captured through unequal k and unequal information about school popularity, can generate unequal outcomes across groups.

This work uses the Amsterdam case to illustrate a promise gap in public algorithmic systems: a gap between what institutions publicly assure and what modified allocation systems can sustain once participants adapt to them.

3.2 AMSTERDAM SCHOOL CHOICE AND THE PLACEMENT GUARANTEE

Amsterdam secondary school choice is a city-wide allocation system in which students may rank schools across the city rather than being assigned through strict neighborhood catchments. The implemented procedure combines lottery-based priorities with local policy adjustments such as the placement guarantee. Students submit a ranked list, are considered in a random order in the main round, and are assigned to the highest-ranked school on their list with remaining capacity. Students who remain unassigned after this round are eligible for the guarantee only if their submitted list satisfies the required minimum length l [59].

The guarantee stage is modeled here as an ex-post capacity expansion. Eligible students are reconsidered for placement at schools on their submitted lists after additional capacity is created. This modeling choice focuses on the feature that is central to our analysis: the guarantee creates additional seats only after the main-round allocation has produced a set of unassigned but eligible students. The relevant question is therefore whether the number of eligible students can exceed the additional capacity created by the policy.

The list-length requirement has also changed in practice. For the 2025 round, the minimum list length for pre-university level (vwo) and general secondary level (havo) was reduced from 12 to 9 schools [16]. For 2026, the requirement was restored to 12 after OSVO concluded that the 9-school list used in 2025 had affected the placement guarantee; while school boards acknowledged that a shorter list is preferable, they stated that being able to provide the guarantee was more important [17, 61]. The guarantee is thus influenced by the choice of list length because this length affects how many schools remain available when the system tries to place students who are unassigned after the main round. This motivates treating l not as a neutral technical parameter, but as an institutional design choice that can change both behavior and feasibility.

Two features of the Amsterdam setting are central to the analysis that follows. First, many families may have sincere preferences over fewer schools than the required list length. Second, some schools are widely known to be heavily oversubscribed. These features together make the tail of the submitted list strategically important: it can be filled either with acceptable but less-preferred schools, or with highly demanded schools that are unlikely to admit the student in the main round. This motivates the formal model developed in Section 3.4.

3.3 RELATED WORK

This chapter relates to work on school choice mechanism design, behavioral responses in school choice, simulation-based studies of information inequality, and algorithmic systems in public institutions.

A large literature studies the formal properties of school choice mechanisms, including Deferred Acceptance and Random Serial Dictatorship [1, 28, 40, 55]. Practical implementations of these algorithms may include local policy adjustments [56, 57], which can materially change efficiency and incentives [4, 58]. Our chapter studies modified school choice algorithms directly: specifically, a placement guarantee added to an RSD-like allocation procedure.

A second strand of work shows that participant behavior is central to how school choice systems perform in practice. Fack et al. [62] show that even under Deferred Acceptance, submitted rank-order lists may diverge from sincere preferences when students anticipate their admission chances and omit schools they regard as infeasible. Other work shows that beliefs, information, and strategic sophistication shape behavior across allocation mechanisms [45–47, 63], and that responses to changed strategic incentives vary across participants and institutional settings [43]. We build on this literature by studying a policy modification that is intended as a safeguard, but that creates a profitable reporting strategy once behavioral response is taken into account.

The chapter also relates to simulation-based work on school choice under limited or unequal information. Díaz et al. [64] develop an agent-based model of school choice with information asymmetries and show that better information can improve outcomes without eliminating inequality. Prior work also shows that families differ in their access to information and in how they interpret school choice options [65, 66]. Our contribution differs in that we focus on how a protective rule can both become infeasible at the system level and distribute advantages unevenly when strategic capacity differs across groups.

Finally, the chapter connects this school-choice case to broader work on algorithmic systems in public administration. Recent work stresses that such systems should be evaluated both in terms of formal design and in relation to public values, institutional context, transparency, and deployment conditions [67, 68]. The Amsterdam placement guarantee is therefore not only a technical addition to an allocation rule, but also a public assurance whose credibility depends on behavioral assumptions. We use this case to show how a promise gap can emerge when an implemented algorithmic system is not tested against strategic response.

3.4 SCHOOL CHOICE MODEL

We consider a finite set N of n students and a set S of m schools. Each school $s \in S$ has base capacity $c_s \in \mathbb{N}$, and each student $i \in N$ has a strict true preference order $\succ_i$ over S. Students submit a reported list L_i, an ordered list of distinct schools in S. There is a minimum list length $l \geq 1$: a student is eligible for the placement guarantee if and only if $|L_i| \geq l$.

In the main round, allocation is performed by Random Serial Dictatorship (RSD) on reported preferences. A permutation π of N is drawn uniformly at random. Starting from capacities c_s, students are considered in the order of π; when student i is called, they are assigned to the highest-ranked school in L_i with remaining capacity. If no school in L_i has remaining capacity, student i remains unassigned. Let $a_i \in S \cup \{\varnothing\}$ denote the main-round outcome for i, and define $U = \{i \in N : a_i = \varnothing \text{ and } |L_i| \geq l\}$ as the set of unassigned students who qualify for the placement guarantee.

The placement guarantee is modeled as a proportional capacity expansion at rate $\gamma \geq 0$ for each school: $\tilde{c}_s = c_s + \lfloor \gamma c_s \rfloor, \qquad s \in S$. Here, γ is the policy parameter governing how much additional capacity is created at each school relative to its base capacity. We define the total additional capacity induced by the guarantee as $C^{\text{extra}} = \sum_{s \in S} \lfloor \gamma c_s \rfloor$. In the guarantee stage, seats allocated in the main round remain fixed, and only students in U are reconsidered. Students in U are reconsidered in the same relative order in which they appeared in the original permutation π. Under the

expanded capacities $\tilde{c}_s$, each such student is assigned to the highest-ranked school on their submitted list with remaining capacity, if such a school exists. We use C^{extra} as the total number of additional seats created by the guarantee. If the number of guarantee-eligible students exceeds this quantity, the guarantee cannot be satisfied for all eligible students.

We distinguish between truthful and strategic preference reporting. A truthful student reports a list L_i that coincides with the top $|L_i|$ positions of $\succ_i$. We assume that each student has sincere preferences over at most k schools, with $1 \leq k \leq l$, and assigns zero utility to all other schools. A fraction $\alpha \in [0,1]$ of students are strategic, the remaining fraction $1-\alpha$ are truthful, and we use (k, l, α, γ) as the main parameters of interest. For the analytical results, we assume that there exists a set $P \subseteq S$ of popular schools with $|P| \geq l-k$ that attract a large share of demand and have limited capacity, and that it is common knowledge which schools belong to P. We focus on a simple strategic behavior in which students list their top k schools truthfully and use popular schools from P to fill the remaining positions up to l. The following section formalizes this strategy and provides a best-response argument.

3.5 STRATEGIC REPORTING UNDER PLACEMENT GUARANTEE

To explain students' behavior, we focus here on the incentives of a single student i, taking the behavior of all other students as fixed. Student i has sincere preferences over at most $k < l$ schools, $r_1 \succ_i r_2 \succ_i \cdots \succ_i r_k$, and is indifferent between all other schools except for their effect on eligibility for the placement guarantee. Let $u_i(s)$ denote the utility of i for school s, with $u_i(r_j) > 0$ for $1 \leq j \leq k$ and $u_i(s) = 0$ for any school that i considers unacceptable.

Let $P \subseteq S$ be the set of popular schools introduced in Section 3.4, with $|P| \geq l-k$. The *top-k truthful* strategy for student i is the reported list $L_i^{\text{TP}} = [r_1, \ldots, r_k, p_1, \ldots, p_{l-k}]$, where $(p_1, \ldots, p_{l-k})$ is any ordering of $l-k$ distinct schools from P. Thus i reports the sincere top k schools and uses popular schools to fill the remaining positions up to the guarantee threshold l.

We restrict attention to reported lists that agree with the sincere ranking in the top k positions and differ only in the tail. Formally, let $\mathcal{L}_i$ be the set of lists $L_i = [r_1, \ldots, r_k, t_{k+1}, \ldots, t_l]$, where $t_{k+1}, \ldots, t_l$ are distinct schools in $S \setminus \{r_1, \ldots, r_k\}$. By construction, any school in positions $k+1, \ldots, l$ yields zero direct utility for i.

To isolate the incentive created by the tail of the list, we impose a stylized assumption on the guarantee stage.

Assumption 1. *For student i, the expected utility from the placement guarantee depends only on the top k positions of L_i and is the same for all lists $L_i \in \mathcal{L}_i$. Moreover, schools in P are sufficiently oversubscribed that the probability of i being assigned to any $p \in P$ in the main round is negligible.*

For a reported list $L_i = [s_1, \ldots, s_l]$, let $p_j(L_i)$ denote the probability that student i is assigned to s_j in the main RSD round, and let $P_{\text{MR}}(L_i) = \sum_{j=1}^{l} p_j(L_i)$ be the probability that i is assigned to some school on the list in the main round. Let $u_i^{\text{PG}}(L_i)$ denote the expected utility from the placement guarantee stage, conditional on being

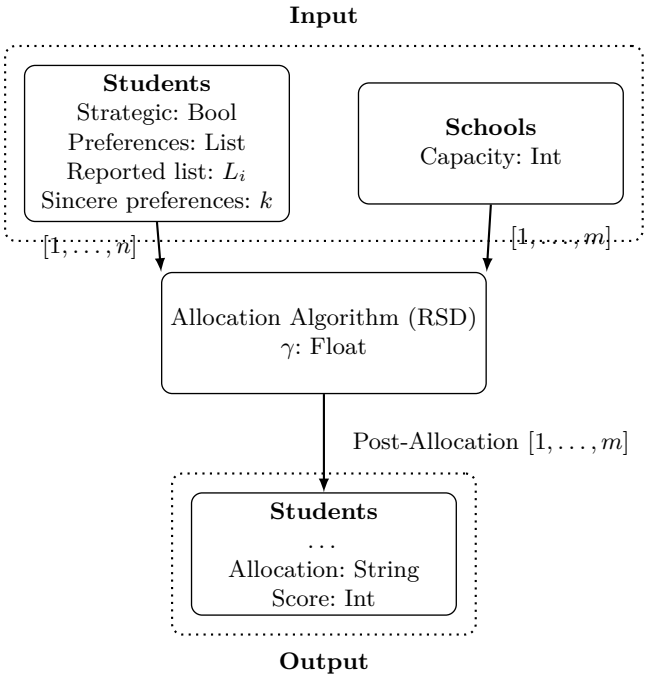


FIGURE 3.2: Pipeline of the simulation environment.

unassigned and guarantee-eligible. The expected utility of i under L_i can be written as

$$\mathbb{E}[U_i(L_i)] = \sum_{j=1}^{l} p_j(L_i)\, u_i(s_j) + \big(1 - P_{\mathrm{MR}}(L_i)\big)\, u_i^{\mathrm{PG}}(L_i).$$

For lists $L_i \in \mathcal{L}_i$, the top k positions coincide with $[r_1, \ldots, r_k]$, and all tail schools $t_{k+1}, \ldots, t_l$ yield zero direct utility. Hence the first term reduces to $\sum_{j=1}^{k} p_j(L_i)\, u_i(r_j)$. Under Assumption 1, $u_i^{\mathrm{PG}}(L_i)$ is the same for all $L_i \in \mathcal{L}_i$, so the tail affects the second term only through $P_{\mathrm{MR}}(L_i)$.

Proposition. *Consider a student i with sincere preferences over at most $k < l$ schools. Under Assumption 1, any list in $\mathcal{L}_i$ that uses only schools from P in positions $k+1, \ldots, l$ weakly dominates any list in $\mathcal{L}_i$ that uses at least one other school in the tail. In particular, the top-k truthful strategy L_i^{TP} is a best response within $\mathcal{L}_i$.*

Proof sketch. Replacing a tail school that may assign i in the main round by a popular school $p \in P$ with negligible assignment probability leaves the distribution over $\{r_1, \ldots, r_k\}$ unchanged and weakly reduces $P_{\mathrm{MR}}(L_i)$. Since $u_i^{\mathrm{PG}}(L_i)$ is unaffected within $\mathcal{L}_i$, this weakly increases the weight on the guarantee term while preserving the first term, and thus weakly increases $\mathbb{E}[U_i(L_i)]$. Iterating this replacement over all tail positions yields the claim. □

3.6 SIMULATION ENVIRONMENT AND DATA

We study the mechanism in Section 3.4 using an agent-based model (ABM) implemented in Python, and tested using both synthetic and real-world preference data from Amsterdam.

3.6.1 *Simulation Environment*

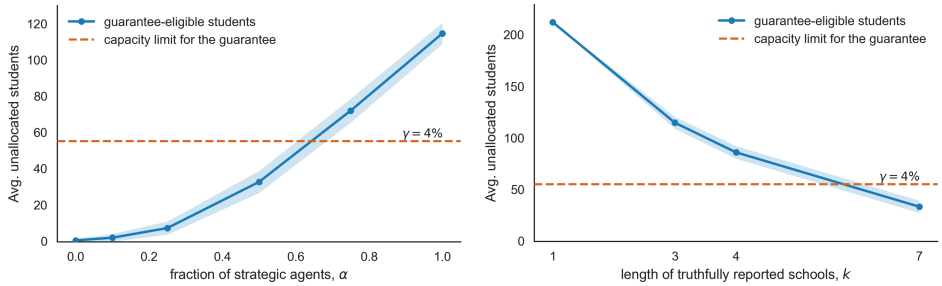
A pipeline diagram of the environment is shown in Figure 3.2. The environment contains n students and m schools with fixed base capacities. In each run of the environment, every student submits a ranked list of length l. The mechanism runs RSD with a placement guarantee as defined in Section 3.4, and each student receives an allocation and a rank-based score. Agents follow one of two reporting strategies: (i) *truthful* agents report the top l schools of their true preference order; and (ii) *strategic* agents use the top-k truthful schools, then the popular strategy from Section 3.5, reporting their sincere top k schools followed by $l - k$ popular schools. The set and ordering of popular schools are provided in the environment from the datasets. For some experiments, this ordering is perturbed using Gaussian noise governed by an error parameter P_{error}. This allows us to model differences in agents' ability to identify the schools that are useful for strategic padding. The parameter P_{error} is interpreted as the level of uncertainty in a group's perceived popularity ordering, and two groups with the same P_{error} may still receive different realized noisy popularity orderings.

For a given experiment, we fix all environment parameters and run the same allocation problem 10,000 times to average over the randomness in the RSD order. School capacities, student preferences, and the popular set remain fixed within an experiment, and we report averages over repeated runs of the matching procedure. The ABM assumes: (i) stationary agents, with true preferences fixed across repeated runs of the matching procedure and no learning or adaptation; (ii) baseline sincerity, so that truthful agents report their true preferences and strategic behavior arises only through the specified reporting rules; (iii) random order, with the RSD order drawn from Python's pseudo-random number generator and fixed seeds across experiments for comparability; and (iv) school popularity is known and derived from the relevant preference dataset.

Our main outcome measures are: (i) the number of *guarantee-eligible students*, defined as students who remain unassigned after the main RSD round while satisfying the list-length requirement; and (ii) final allocation outcomes after the guarantee stage, which we use in the group-based experiments on unequal strategic capacity.

3.6.2 *Datasets*

We instantiate the above simulation framework on two types of preference data. The first is a synthetic preference dataset designed to isolate the incentive effects of the placement guarantee. In this setting, we consider $n = 100$ students, $m = 20$ schools, and list length $l = 8$. Preferences are generated using a popularity-skewed procedure in which a small subset of schools receives higher weight and therefore attracts a larger share of rank-1 demand. This produces a recognizable set of popular schools



(a) varying α (b) varying k

FIGURE 3.3: **Feasibility of the placement guarantee under strategic adaptation in the Amsterdam-based setting.** The figure shows the average number of guarantee-eligible students after the main RSD round as the fraction of strategic agents, α, and the number of truthfully reported schools, k, vary. The dashed red line marks the total additional capacity C^{extra} created by the placement guarantee. Error bars show 95% confidence intervals over 10,000 simulation runs.

and provides a controlled environment in which strategic padding can be studied directly.

Our second dataset is a real-world preference dataset provided by the Amsterdam School Boards Association (OSVO) [13]. We use preference lists from 2017 for pre-university level students (vwo), where competition is highest for several popular schools, and retain only lists containing at least 12 schools. Because restricting the student population without adjusting capacities would distort the market, we first allocate the excluded students under RSD using their observed submitted preferences and then use the residual school capacities for our dataset. Since this data was collected before the introduction of the placement guarantee, we treat the observed preference lists as truthful reports that are not yet shaped by strategic responses to the guarantee.

The same strategies are then applied to both datasets. Truthful agents report their preferences sincerely, while strategic agents are assigned modified lists according to the reporting rules studied in the chapter. Popular schools are identified using the rank-1 demand for each school in the datasets, which we use as a tractable proxy for widely recognized school popularity.

3.7 EXPERIMENTAL DESIGN

We study two sets of experiments. (1) Feasibility experiments examine when the placement guarantee becomes infeasible at the system level. (2) Inequality experiments examine whether unequal strategic capacities produce unequal outcomes across groups. Each set of experiments is run on both the synthetic and Amsterdam-based preference data described in Section 3.6.2.

The feasibility experiments vary the two main parameters that determine how strongly agents can exploit the placement guarantee. The first is the fraction of

Experiment	Varying Parameters	Fixed Parameters		
		k	α	P_{error}
Feasibility #1	α	3	-	0
Feasibility #2	k	-	1.0	0
Unequal strategy #1	k	4	1.0	0
Unequal strategy #2	P_{error}	4	1.0	0.05

TABLE 3.1: **Summary of experimental settings.** Placement guarantee expansion rate is fixed at $\gamma = 0.04$ for all experiments. For the unequal strategic capacity experiments, only Group 2's parameters are varied, while Group 1's are fixed.

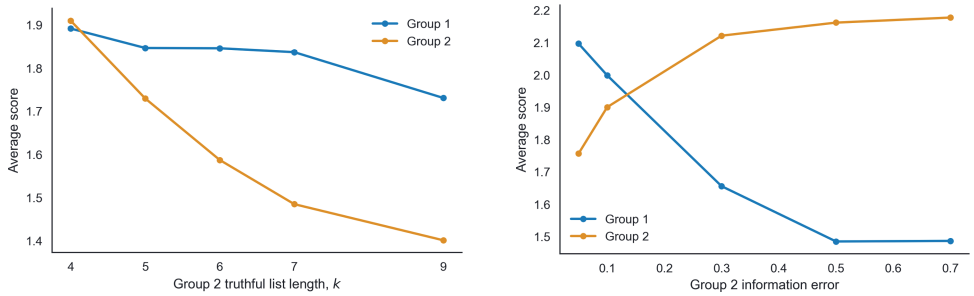
strategic agents, α, which captures the extent of strategic reporting in the population. The second is the number of sincere preferences, k, which captures how many schools an agent can rank meaningfully before the remainder of the list must be filled strategically. For each parameter setting, we compute the average number of guarantee-eligible students after the main RSD round.

These experiments identify some conditions under which the placement guarantee can no longer be satisfied.

The inequality experiments examine whether the same algorithm produces unequal outcomes when strategic agents differ in their ability to apply the reporting rule studied in this chapter. We consider two forms of unequal strategic capacity where an environment contains two strategic groups of equal size, and no truthful agents. In the first, the two groups differ in the number of schools they can rank sincerely, operationalized through different values of k. In the second, the two groups differ in how accurately they identify popular schools, operationalized through different values of P_{error} in the popularity ordering used for strategic padding, as described in Section 3.6.1. In both cases, we compare final outcomes across groups after the guarantee stage. These experiments are intended to distinguish between system-level infeasibility and distributional effects, and to test whether the placement guarantee disproportionately benefits agents with better strategic knowledge.

3.8 RESULTS

We report results corresponding to the three research questions introduced in Section 5.1. First, the formal analysis in Section 3.5 shows how the placement guarantee changes incentives for students with fewer than l sincere preferences. Second, the feasibility experiments examine when strategic adaptation makes the guarantee infeasible at the system level. Third, the inequality experiments examine whether unequal strategic capacity produces unequal outcomes across groups. In the main text, we focus on the Amsterdam-based experiments, since the chapter is motivated by the Amsterdam case. Corresponding synthetic results, which preserve the same qualitative patterns in a more controlled setting, are reported in the appendix.



(a) Different truthful-list lengths (k) across groups

(b) Different popularity-information accuracy across groups

FIGURE 3.4: **Unequal strategic capacity experiments in the Amsterdam-based setting.** Panel (a) shows average score by group as Group 2's truthful-list length, k, varies while Group 1 is held fixed. Panel (b) shows average score by group as Group 2's error in identifying popular schools varies while Group 1 is held fixed. In both panels, the two groups are of equal size and all agents are strategic. Lower average score indicates a more favorable final outcome. Error bars show 95% confidence intervals over 10,000 simulation runs, but are too small to be visible.

3.8.1 *Feasibility of the Placement Guarantee*

The feasibility experiments demonstrate that the placement guarantee becomes increasingly difficult to sustain as strategic reporting becomes more widespread and as students have fewer sincere preferences.

Figure 3.3 reports the Amsterdam-based feasibility results. Panel (a) varies the fraction of strategic agents, α, while holding $k = 3$ and $\gamma = 0.04$ fixed. The average number of guarantee-eligible students increases monotonically in α, from close to zero when all agents are truthful to well above the available additional capacity once strategic reporting becomes common. At $\alpha = 0.50$, the average number of guarantee-eligible students remains below the capacity threshold, but by $\alpha = 0.75$ it exceeds the total additional capacity created by the guarantee, and by $\alpha = 1.00$ it is more than twice as large. This indicates that the guarantee can become infeasible once strategic reporting is sufficiently widespread.

Panel (b) varies the number of sincere preferences, k, while holding $\alpha = 1$ and $\gamma = 0.04$ fixed. The average number of guarantee-eligible students decreases sharply as k increases. When students can rank only a small number of schools sincerely, a large share of the list is used for strategic padding and the guarantee is exceeded by a substantial margin. As k increases, the pressure on the guarantee falls. In the Amsterdam-based setting, the average number of guarantee-eligible students remains above the capacity threshold for $k \in \{1, 3, 4\}$, but falls below it at $k = 7$. This shows that the feasibility of the guarantee depends strongly on how many preferences students can rank meaningfully before they are incentivized to complete the list strategically.

These results show that the placement guarantee fails under exactly the conditions that make the strategic rule attractive. As strategic behavior spreads, or as students

have fewer sincere preferences, the number of guarantee-eligible students can exceed the additional capacity created by the policy. The synthetic experiments reported in Appendix B.2.1 show the same qualitative mechanism in a more stylised environment.

3.8.2 *Unequal Strategic Capacity*

The inequality experiments demonstrate that unequal strategic capacity produces unequal outcomes. In the discussion below, lower average score indicates a more favorable final outcome.

Figure 3.4 reports the Amsterdam-based inequality results. Panel (a) varies Group 2's truthful-list length while holding Group 1 fixed. As Group 2's k increases, Group 2's average score decreases substantially, indicating better final outcomes. Group 1's score also improves but more modestly. When both groups have the same truthful-list length, their outcomes are similar. As the gap in truthful-list length widens, the group that can rank more schools sincerely performs better. This suggests that agents with a larger sincere preference set are better positioned to benefit from the mechanism.

Panel (b) varies Group 2's popularity-information error, P_{error}, while holding Group 1's P_{error} fixed. As Group 2's information becomes noisier, its average score worsens steadily, while Group 1's score improves in this shared allocation setting. The difference at the equal-error baseline arises because the noisy popularity ordering is sampled separately for each group. This suggests that unequal information about which schools are useful for strategic padding can also translate directly into unequal final outcomes.

These experiments show that in addition to creating a system-level feasibility problem, the placement guarantee can produce unequal outcomes when groups differ in their capacity to use the rule strategically. Groups are better positioned under the rule when they can either rank more schools sincerely or identify oversubscribed schools more accurately. The synthetic inequality experiments reported in Appendix B.2.2 preserve the same qualitative pattern.

3.9 DISCUSSION

3.9.1 *The Promise Gap*

The central claim of this chapter is that the Amsterdam placement guarantee creates a *promise gap*: a mismatch between the protection communicated to families and what the implemented system can sustain under strategic adaptation. The guarantee is introduced as protection against non-placement, and truthful reporting is presented as safe. For students with fewer than l sincere preferences, however, the guarantee changes the role of the remaining list positions. These positions no longer only represent less-preferred schools; they can also be used to preserve eligibility for the guarantee itself. The experiments show how this promise gap becomes consequential. When strategic reporting becomes sufficiently widespread, the number of guarantee-eligible students can exceed the additional capacity created by the policy, and the guarantee cannot be satisfied for all eligible students. A safeguard introduced to protect families can therefore induce behavioral responses that weaken the very protection it was meant to provide.

This matters institutionally because public communication about the Amsterdam system has long drawn on the legitimacy of the Deferred Acceptance algorithm and its association with Nobel Prize-winning work [69, 70]. Yet the implemented system differs from textbook Deferred Acceptance: in the Amsterdam setting, lottery-based priorities make the main round closer to Random Serial Dictatorship, while the placement guarantee adds a further policy modification. The relevant object of evaluation should therefore be the implemented system as it operates in practice, rather than the theoretical mechanism to which it is publicly connected. Broadly, this shows the need for more robust evaluation of public algorithmic systems beyond the formal results reported in the academic literature: even small local modifications can have large consequences for system behavior, and should therefore be tested rigorously before being adopted as policy.

3.9.2 *Unequal Strategic Capacity and Institutional Fairness*

The results also show that the placement guarantee has a distributional dimension. The same rule can produce different outcomes when groups differ in their strategic capacity. In our experiments, groups obtain better outcomes when they can rank more schools sincerely or when they have more accurate information about which schools are useful for strategic padding. The benefits of the safeguard are therefore not distributed evenly across participants.

This raises a broader fairness concern. A rule can be formally neutral while still having unequal effects in practice. When participants differ in their knowledge, resources, or ability to navigate the system, protections designed to reduce risk may advantage those who are already better positioned to use them. In this sense, the placement guarantee can reproduce or amplify existing asymmetries in access to strategic knowledge.

3.9.3 *Transparency, Trust, and Responsible Communication*

These findings also have implications for institutional communication, especially in settings where transparency is treated as a public value but may also affect strategic behavior [68]. In Amsterdam, families are told that reporting preferences truthfully gives them the best outcome. Our results show that this message is incomplete once the placement guarantee is added. For students with fewer than l sincere preferences, truthfully ranking preferred schools does not determine how the remaining positions should be filled. The placement guarantee creates a strategic incentive that is not reflected in the public message of straightforward truth-telling.

This communication gap can become a trust gap. When families are told that the system is safe to use truthfully, but the implemented rule creates incentives to complete lists strategically, it can create an additional burden among parents. The recent changes to the list-length requirement illustrate this problem. As noted in Section 3.2, the minimum list length was changed back from 9 to 12 in 2026 as the placement guarantee was difficult to satisfy with a list length of 9. Without a clear explanation of the strategic pressures created by the guarantee, such changes can frustrate families and weaken confidence in the system.

Institutions should therefore communicate the limits of the guarantee and the behavioral conditions under which it can be sustained. Before presenting safeguards as reliable public assurances, they should test whether those safeguards remain feasible under plausible strategic responses. Where strategic adaptation can undermine feasibility, this risk should be communicated explicitly rather than hidden behind a general message of truthfulness.

3.10 LIMITATIONS

This chapter has several limitations. First, the strategic model is deliberately simplified. We identify one profitable reporting strategy and study its consequences, but do not characterise the full set of best responses or equilibria under the placement guarantee, as our goal was to study a condition behaviorally plausible for Amsterdam. The results should therefore be read as a stress test of the guarantee and not as a calibrated prediction of behavior in a specific admission year.

Second, the simulations assume stationary agents. Strategic behavior is imposed through reporting rules rather than emerging through learning, observation, or imitation. In practice, families may adapt over time as information circulates through schools, parent networks, media reports, or previous experience with the system.

Third, the Amsterdam-based experiments are based on preprocessing choices that simplify the real institutional process. The retained population, residual-capacity adjustment, and use of pre-guarantee preference data are reasonable for isolating the incentive effects of the guarantee, but they do not reconstruct the Amsterdam system in full.

Fourth, strategic capacity is operationalized through two proxies: the number of schools agents can rank sincerely and the accuracy of their information about popular schools. These proxies capture important differences across groups, but families may also differ in risk attitudes and familiarity with the school choice process.

Future work could examine richer strategic behavior under the placement guarantee, including broader best-response and equilibrium analyses. It should also model learning and information diffusion explicitly, so that strategic behavior emerges over time rather than being fixed in advance. A further direction is to study alternative guarantee designs that are more robust to strategic response and easier to justify as public assurances.

3.11 CONCLUSION

This chapter studied Amsterdam secondary school choice as a case of a modified public algorithmic system. We examined three consequences of the placement guarantee: the strategic incentive it creates for students with fewer than a fixed number of sincere preferences, the conditions under which this behavior makes the guarantee infeasible, and the unequal outcomes that arise when groups differ in strategic capacity. We showed that students with fewer than a fixed number of sincere preferences can benefit from reporting their top schools truthfully and filling the remaining positions with highly demanded schools. Using simulations on Amsterdam-based and synthetic preference data, we then showed that as this behavior spreads, the number of guarantee-

eligible students can exceed the additional seats created by the policy. The guarantee can therefore become infeasible under the strategic responses it helps create. The broader lesson is that policy fixes can create promise gaps. A modification introduced to make a public algorithmic system more protective can change incentives in ways that undermine the assurance it was meant to provide. For public-sector algorithmic systems, responsible evaluation cannot stop at the formal properties of the original algorithm, but must also examine how implemented modifications interact with behavior, and whether the promises used to justify those modifications can still be delivered in practice.

4

STRATEGIC REPORTING IN PRACTICE

Based on
Tasnim, M., Verhagen, P., Blanke, T., Acar, E., Ghebreab, S.
Modeling strategic risk in school choice: A case for transparent design.
Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society **8**(3), 2470–2479 (2025) [71].

The previous two chapters identified two sources of strategic incentive in one-sided matching: the mechanism itself, and the policy modifications layered around it. In both cases, the analysis showed that a profitable misreporting strategy exists and that, in principle, participants could use it to obtain better outcomes. Showing that an incentive exists, however, is not the same as showing that participants will act on it. A strategy may be available in theory while remaining too risky, too opaque, or too cognitively demanding to be adopted in practice, and the policy implications of formal manipulability depend on which of these is true. This shift from formal possibility to behavioral response motivates the third research question of this thesis:

Research Question 3: *If strategic incentives exist formally, how do real participants respond to them?*

We address this through a behavioral experiment embedded in a survey of 140 parents participating in Amsterdam's 2025 secondary school choice process. Each parent selects among truthful, mildly strategic, and aggressively strategic preference rankings across three scenarios that vary how much information about admission probabilities is disclosed. We then estimate a discrete-choice model grounded in Security-Potential/Aspiration (SP/A) theory, which represents each decision as a trade-off between aspirational goals and the risk of poor outcomes, and compare its fit against expected utility and prospect theory benchmarks. We find that disclosing probability information does increase strategic behavior, but that truthful reporting remains the most common choice under every condition, and that aggressive strategies are rarely adopted even when they appear favorable. Two findings challenge common assumptions in the school choice literature: transparency about strategic risk acts as a deterrent rather than an enabler of manipulation, and parents in the more competitive academic tracks are more risk-averse rather than more strategic. These results suggest that the trade-off between efficiency and manipulability may be less sharp in practice than formal analyses imply, and that risk communication can serve as a policy lever for fairness in algorithmic allocation systems.

4.1 INTRODUCTION

School choice mechanisms co-determine access to educational opportunity and, in turn, long-term social mobility [66]. In school choice, assignment algorithms often operate as one-sided matching mechanisms, where students rank schools but schools do not rank students [5]. Policymakers and researchers have long debated the role of transparency in these systems: while making the rules of the system visible can improve trust and accountability, it is also feared to invite "gaming" that undermines fairness. This tension is especially acute in matching algorithms that are not *strategyproof*, where families can potentially benefit from misreporting their preferred schools [5]. For example, strategyproof algorithms like Random Serial Dictatorship (RSD) encourage truth-telling, but lead to inefficient outcomes, leaving many students without access to top-ranked schools [1]. In contrast, more efficient matching methods are manipulable and incentivize families to misrepresent their preferences [3, 29, 31].

This balance between efficiency and equity has shaped school assignment policies in many cities. The prevailing assumption is that parents with a high socioeconomic status (SES) are more likely to identify and exploit strategic advantages, while lower-SES families may struggle to do so [4, 65, 66]. To avoid reinforcing inequality, many systems retain strategyproof mechanisms despite reduced efficiency, or limit transparency when using non-strategyproof mechanism. Despite the existence of this assumption, little is known in the literature about how families actually respond to incentives under manipulable systems. Previous work has shown that strategic behavior increases when mechanisms are manipulable, particularly when assignment data or algorithmic rules are made transparent [46, 47]. However, these findings are largely derived from experimental studies in lab settings and may not capture how real-world parents behave, particularly in a high-stakes context like school choice, where parents can likely be more risk-averse [72].

Amsterdam offers a relevant case. In 2016, the city adopted a lottery-based mechanism that prioritized strategyproofness but led to dissatisfaction due to inefficiency, with many students matched to schools ranked low on their preference list [19, 73]. Following complaints and legal appeals, a 2019 policy change introduced a placement guarantee, which also rendered the system manipulable. The change was not widely communicated, and strategic knowledge on its manipulability spread informally. To mitigate this, the school boards reduced transparency by no longer publishing anonymized preference data or the lowest lottery number at which a student secured a place at each school, which had previously helped parents estimate admission probabilities [74].

This shift also reflects a common policy assumption derived from the literature: that reducing information transparency deters manipulation by increasing uncertainty. Yet opacity has its own drawbacks, including reduced accountability and trust in systems [68]. Moreover, in many matching settings, it is difficult to determine an optimal strategy, even when providing high levels of information transparency [19, 32]. This raises a key question: rather than incentivizing manipulation, can transparency about the probabilities of success of strategies serve as a deterrent by making the risks of manipulation more explicit?

To investigate this, we conduct a behavioral experiment embedded in an online survey of 141 parents whose children are about to participate in Amsterdam's lottery-

based school choice system. This population was chosen intentionally as real-world stakes may shape how parents perceive and respond to strategic risk, in contrast to typical lab settings, with perceived manipulability being different than actual manipulability. Participants were asked to complete a simplified school matching task under a manipulable algorithm, choosing between being truthful, mildly strategic, and aggressive manipulation. They repeated the task under three sequential conditions: (1) no probability information, (2) full information with high risk, and (3) full information with reduced risk. This experimental setup is designed to allow us to isolate the effects of transparency and strategic risk.

Our findings show that transparency does increase the rate of strategic behavior, but that most parents opt for mild deviations. Aggressive manipulation tactics are not adopted universally, even when risk is low. Truthful reporting remains the most commonly picked strategy in all three scenarios, and among those who deviate, mild strategies are far more common than aggressive ones. To explain this, we model parents' decisions using the Security-Potential/Aspiration (SP/A) framework from behavioral decision theory. In this model, each strategy is evaluated based on its security (avoiding bad outcomes) and potential (achieving one's aspiration), moderated by a parent-specific risk preference. We estimate this model using a multinomial logit and recover interpretable heterogeneity in behavior. The fitted risk preferences are bimodally distributed, and show that parents with children in a more competitive school type tend to be risk-averse, which runs counter to the common assumption that competition fuels strategic behavior.

Overall, this chapter makes three contributions. First, we provide empirical evidence on how parents respond to strategic incentives and transparency under a manipulable mechanism. Second, we introduce SP/A theory as a behavioral framework for understanding strategy use in school choice, showing that aspiration-driven behavior aligns with observed patterns. Third, we offer a policy insight: transparency, rather than uniformly enabling manipulation, may reduce harmful strategies by clarifying the risks under manipulable matching methods. Our findings suggest that sharing accurate information about strategy success and failure could discourage harmful manipulation without sacrificing institutional trust. This aligns with recent simulation-based work showing that only a small fraction of strategic reports lead to better outcomes in Amsterdam, implying that explicit transparency could act as a behavioral deterrent [19]. Our findings suggest that transparency can support both equity and efficiency in the design of school choice algorithms.

4.2 RELATED WORK

4.2.1 *Algorithmic Transparency*

With the growing prevalence of algorithmic school allocation, there has been increased scrutiny of how such systems align with public values and requirements that come with the use of algorithms in the public sector.

Transparency in particular is emphasized in key regulatory frameworks such as the EU AI Act [75] and the Digital Services Act [76]. The latter, for example, mandates in-scope service providers to publish an annual transparency report with respect to content moderation, and disclose algorithmic operations that may impact the

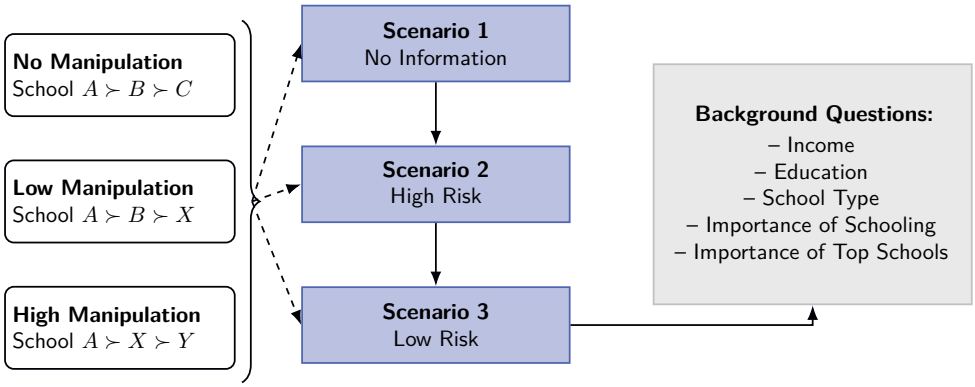


FIGURE 4.1: Survey structure and experimental design. Respondents select from one of three preference lists in each of three scenarios varying in the availability of information and strategic risk. Background questions follow the final scenario.

public domain. However, while transparency is increasingly seen as a normative requirement for public sector algorithms, some scholars have argued that great algorithmic transparency carries the risk of strategic manipulation and gaming [77]. Moreover, there can be legal constraints both from a privacy requirement and an intellectual property requirement that prevent transparency. Crucially for this study, policy choices can also be important variables in determining the level of transparency, including preventing negative media exposure, equity concerns, or fear of strategic manipulation.

This latter argument is central in the Amsterdam school allocation context, with the choice of Random Serial Dictatorship (RSD), orienting around strategic manipulation and the associated fairness concerns. Particularly, this is based on the assumption that greater transparency regarding admission probabilities might incentivize parents to engage in strategic manipulation, exacerbating social inequalities, as parents with a higher socioeconomic background would be more likely to engage in such manipulation [4, 65, 66]. This positions school allocation as not merely a technical problem, but a politically embedded problem that involves political and normative tradeoffs, including between efficiency and fairness, and between transparency and manipulability. At the same time, algorithmic transparency has been shown to have positive effect; prior works suggest it can strengthen trust in algorithmic decision making [67], support actionable recourse [78] and even improving predictive power of algorithms[77]. Our study examines whether, within the context of school allocation in Amsterdam, greater transparency on algorithmic outcomes leads to greater (and potentially undesirable) strategic behavior, and whether these effects vary based on parent's aspirational levels. In particular, we focus our investigation on the behavior of the higher SES parents, as the literature suggests that these are most prone to engage in strategic behavior.

4.2.2 *Strategic Behavior in School Choice*

Centralized school choice systems use algorithmic mechanisms to allocate students to schools based on stated preferences and school capacities. The Deferred Acceptance (DA) algorithm [40, 55] is widely used for its strategyproofness and stability guarantees. However, DA does not necessarily maximize welfare, leading to dissatisfaction among students assigned to lower-ranked choices [4]. Alternative mechanisms such as Rank Minimization (RM) [7, 29] have been proposed to improve efficiency by minimizing the number of students receiving low-ranked school assignments, albeit at the cost of introducing incentives for strategic manipulation [32].

Empirical studies have documented that when mechanisms are manipulable, strategic behavior becomes widespread. Abdulkadiroglu et al. [4] show manipulation under the original Boston mechanism, while Calsamiglia et al. [47] report similar findings in Barcelona. Pathak and Sönmez [45] highlight that non-strategyproofness can disproportionately disadvantage less-informed families. Although these studies establish that manipulation occurs, they offer limited insight into *who* chooses to strategize and *how* individual differences affect strategic behavior.

Socioeconomic factors have been shown to influence school choice behaviors and strategic sophistication. Hastings and Weinstein [66] demonstrate that access to better information alters school selections across income groups, and He [72] finds that higher-income families exhibit greater strategic behavior in school choice settings.

Beyond observational data, experimental and survey-based methods have been used to investigate strategic behavior in assignment systems. Bronfman *et al.* [79] conduct laboratory experiments to study strategic preference reporting under different mechanism designs, while Rees-Jones [80] demonstrates that even with clear strategic incentives, participants often fail to optimally manipulate their preferences.

4.2.3 *Behavioral Theories on Decision Making*

Decision-making under uncertainty has long been studied in behavioral economics and psychology. The Security-Potential/ Aspiration (SP/A) theory, developed by Lopes and Oden [81], models individuals as balancing two competing motivations when facing risky choices: 1) Security, prioritizing the minimization of potential losses, and 2) Potential, focusing on maximizing the best possible outcomes. Aspiration levels serve as reference points in this trade-off, with individuals holding higher aspirations more willing to take calculated risks to achieve desired goals. SP/A theory explains probability distortions not through perceptual biases, as in Prospect Theory [82], but through the emotional dynamics of hope and fear. It emerged in response to empirical violations of the independence axiom in Expected Utility Theory [83], such as those demonstrated in the Allais and Ellsberg paradoxes [84]. SP/A theory has been widely applied in financial decision-making and consumer behavior but has not been extensively explored in the context of school choice. Although expected utility theory models choice as maximizing expected value, it cannot fully explain the reluctance towards risking highly undesirable outcomes. Meanwhile, prospect theory adds probability weighting and loss aversion, but evaluates outcomes along a single utility scale. SP/A theory, by contrast, conceptually separates the drive to secure a satisfactory outcome from the aspiration to achieve the best possible one,

capturing the tendency for people to reject high-probability gains when they risk highly undesirable outcomes.

Our work contributes to this literature by designing a behavioral survey based on realistic Amsterdam school choice conditions, varying transparency of admission probabilities to observe how parents respond to strategic opportunities, and model this behavior using SP/A theory.

4.3 METHODS

We conducted a quantitative behavioral survey (n = 140) to investigate how parental aspirations and socioeconomic background influence strategic behavior under a manipulable school choice mechanism.

We ground our survey design in the SP/A theory [81], as introduced in Section 5.2. This theory models decision-making under risk as a balance between potential, security, and aspiration. In the context of school choice, potential refers to the possibility of achieving a highly desirable outcome, such as being matched with a top-ranked school, while security reflects the desire to avoid the worst-case outcome, namely not being placed at any school at all. Aspiration represents the level of outcome parents hope to achieve, and the gap between this aspiration and their expected outcome influences their willingness to take risks.

4.3.1 *Survey Design*

Participants completed a hypothetical school choice task based on a manipulable matching mechanism. Each respondent was assigned a fixed true preference ranking over three fictional schools, $A \succ B \succ C$, and informed that it might be advantageous to misrepresent their preferences—e.g., by including oversubscribed schools such as X and Y that they did not genuinely prefer.

In each of three scenarios, participants were asked to select one preference ranking from three pre-constructed lists. The first list reflected truthful reporting ($A \succ B \succ C$). The second list, representing a *mild strategy*, included $A \succ B \succ X$, using the oversubscribed school X to improve the chance of placement in the top two choices. The third list, representing an *aggressive strategy*, prioritized two highly competitive schools ($A \succ X \succ Y$) to increase the likelihood of receiving the top choice through a riskier manipulation. The design of the strategies follows from the best-response analysis in [19], where it is shown that listing popular schools after one's top preferred schools can increase the likelihood of being assigned to those top schools. The mild strategy implements this by being truthful about the top two choices and placing one popular school later in the list, applying moderate pressure to secure a placement within the top two. The aggressive strategy is truthful only about the top choice and places two popular schools afterwards, exerting greater pressure to obtain the top-ranked school at the cost of a higher risk of non-placement.

The three scenarios varied in the degree of information available. In the *baseline* scenario, no admission probabilities were shown. In the second scenario, participants were shown realistic probabilities for each list, where strategic options carried substantial risk and offered limited benefit. In the third scenario, the probability of admission to

the first-choice school on the aggressive list was artificially boosted to 90%, simulating a low-risk, high-reward condition. Probabilities in Scenarios 2 and 3 were generated using historical Amsterdam placement data under a manipulable rank-minimizing method similar to [7, 19]. Following each scenario, participants could optionally provide open-ended explanations for their choices. Additional demographic data were collected, including household income, education level, school track preference (e.g., pre-university vs. vocational), and perceptions of school quality and the importance of elite school access.

4.3.2 *Hypothesis*

We test the following hypotheses in our experiment:

H1—STRATEGIC BEHAVIOR: When assignment probabilities are disclosed and the perceived risk of manipulation is low, parents are more likely to behave strategically.

H2—STRATEGY TYPE: Even under low-risk, most parents will avoid aggressive strategies and instead favor mild deviations.

H3—RISK PROFILE AND COMPETITION: Parents applying to more competitive school tracks will exhibit greater risk aversion, regardless of socioeconomic background.

H1 reflects how the disclosure of admission probabilities affects parents' perception of security, while H2 suggests that the perceived loss of security associated with aggressive strategies continues to dominate most parents' choices. H3 builds on the assumption that competitive school tracks elevate the perceived cost of failure. As a result, parents may prioritize security over potential when competition makes negative outcomes (such as being unplaced) more probable and consequential.

4.3.3 *Sampling*

Participants were reached through the Onderwijs Consumenten Organisatie (OCO), a non-profit organization supporting parents navigating the Amsterdam school system. Eligibility criteria included being a parent or guardian with a child eligible for Amsterdam's secondary school lottery. Participants were offered the opportunity to enter a prize lottery upon survey completion. The study was conducted under standards consistent with exemption from formal IRB review, given its minimal risk profile. The questionnaire for the survey is disclosed in Appendix C.1.

Demographic information collected included parental income, education level, school aspiration attitudes, and prior experience with the school lottery process. The demographic of our sample is reported in Table 4.1.

4.3.4 *Analysis*

To evaluate the outlined hypotheses, we conducted a series of descriptive statistical tests. For H1, we compared participants' strategy choices across the three scenarios to assess whether transparency and reduced risk increase the incidence of strategic behavior. We applied McNemar's test [85] to paired responses from Scenario 2 and 3 to

Variable	Category	Count (%)
Education	PhD	11 (7.9%)
	Master's degree	62 (44.3%)
	Bachelor's degree	32 (22.9%)
	Associate's degree	16 (11.4%)
	High School	18 (12.9%)
	Primary school	1 (0.7%)
Income	≥ €100,000	49 (35.0%)
	€50,000–€100,000	44 (31.4%)
	€25,000–€50,000	18 (12.9%)
Track	Pre-University	43 (30.7%)
	General/Pre-University	29 (20.7%)
	General	21 (15.0%)
	Vocational	12 (8.6%)
	Practical education	31 (22.1%)

TABLE 4.1: Demographic characteristics of the sample ($n = 140$). Counts are rounded to the nearest whole number.

assess whether the observed decline in truthful reporting was statistically significant. For H2, we analyzed the distribution of non-truthful choices, comparing the prevalence of mild versus aggressive strategies across scenarios. For H3, we disaggregated strategy use by parent characteristics, particularly school track preference and aspiration level, and assessed whether behavioral differences across groups were consistent with increased risk aversion in competitive contexts. All comparisons were visualized and supported by chi-square tests of independence to identify statistically meaningful group differences.

4.3.4.1 *Modeling Strategic Risk with SP/A Theory*

We estimated a discrete choice model grounded in the Security-Potential/Aspiration (SP/A) theory of decision-making under uncertainty. SP/A theory posits that individuals evaluate risky options based on two competing motivations: the *potential* to achieve an aspirational goal and the *security* of avoiding disappointing outcomes. Rather than assuming distortions in probability perception, as in Prospect Theory, SP/A models decision-making as an affective trade-off between hope for gain and fear of loss.

Formally, we define the utility of strategy s for parent i as:

$$U_{i,s} = \alpha_i \cdot \text{Potential}_s + (1 - \alpha_i) \cdot \text{Security}_s,$$

where $\alpha_i \in [0, 1]$ is a latent parameter capturing individual risk preference. Higher values of α_i indicate a stronger orientation toward potential (risk-seeking), whereas lower values reflect greater emphasis on security (risk aversion).

In our implementation, *potential* was defined as the probability of being placed at the top-ranked school in a given list, while *security* was defined as the probability of avoiding non-placement (i.e., being matched to any school). These probabilities were drawn directly from the scenario-specific admission estimates shown to participants during the survey.

To capture heterogeneity across individuals, we modeled α_i as a logistic function of standardized parent-level features:

$$\alpha_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot \text{Education}_i + \beta_2 \cdot \text{Income}_i + \ldots)}}$$

We estimated a multinomial logit model over the three strategic choices (truthful, mild, aggressive) in each scenario, with utilities computed according to the SP/A formulation above. A single model was fit jointly across all three scenarios, using the scenario-specific potential and security values for each strategy. Each participant contributed one observation per scenario; standard errors were clustered at the respondent level to account for within-subject correlation. Although SP/A theory can be extended with non-linear specifications such as aspiration cutoffs or weighted utility distortions, we employed a linear formulation to prioritize model interpretability and ensure stable parameter estimation given our small sample size of 140.

For comparison, we implemented two benchmark models: an Expected Utility (EU) model [83] with linear probability-weighted outcomes, and a Prospect Theory (PT) model [82] using a single-parameter Prelec weighting function. All models were estimated via maximum likelihood using default convergence settings. Model performance was evaluated using log-likelihood, Akaike Information Criterion (AIC), and predictive accuracy on held-out data.

4.4 RESULTS

4.4.1 *Descriptive Patterns of Strategic Behavior*

Participants' strategy choices varied systematically across the three experimental conditions, as shown in Figure 4.2. In the baseline scenario, where no information about admission probabilities was provided, 80.1% of participants submitted the truthful list. This share declined to 57.4% under the high-risk transparency condition and

Scenarios	**Truthful (%)**	**Mild (%)**	**Aggressive (%)**
Baseline	80.14	14.18	5.67
High Risk	57.45	38.30	4.26
Low Risk	40.43	20.57	39.01

TABLE 4.2: Strategy choices across the three experimental scenarios.

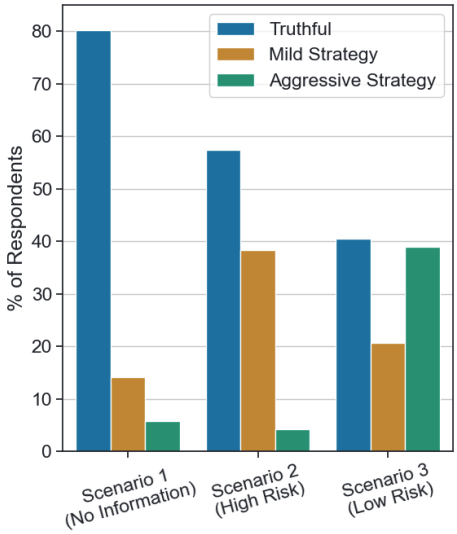


FIGURE 4.2: Strategic behavior by scenario. Each bar shows the percentage of participants choosing Truthful, Mild Strategy, or Aggressive Strategy across Baseline, High Risk, and Low Risk scenarios.

further to 40.4% when the aggressive strategy was presented as favorable. McNemar's test confirms that the reduction in truth-telling between the High Risk and Low Risk conditions is statistically significant ($\chi^2 = 14.69$, $p = 0.00013$). These results support our first hypothesis (H1), indicating that transparency combined with low perceived risk increases the incidence of strategic behavior.

We also examine the specific types of strategic behavior that participants adopted. As shown in Table 4.2, the share of participants choosing the aggressive strategy (List 3) increased sharply from 4.3% in the High Risk condition to 39.0% in the Low Risk condition. This shift suggests that participants were responsive to the underlying probability structure and that some were willing to pursue high-reward strategies when the risk of non-assignment was reduced.

However, even when the aggressive strategy offered a 90% chance of securing the top choice, many participants (∼60%) continued to avoid it. Among strategic respondents, a sizable proportion favored the mild deviation (List 2), which carried less risk of being misplaced. In Scenario 3, 20.6% of all participants selected the mildly strategic option, despite the more favorable probabilities associated with aggressive behavior. This sustained preference for lower-risk strategies provides support for our second hypothesis (H2): while transparency shifts behavior away from truth-telling, risk aversion continues to moderate the extent of strategic deviation.

4.4.2 *Risk Preference Estimation and Model Fit*

The SP/A-based discrete choice model provided a substantially better fit to the data than a constant-risk baseline. A likelihood ratio test comparing the full model to the

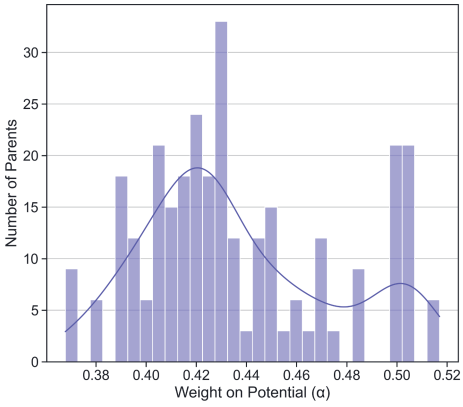
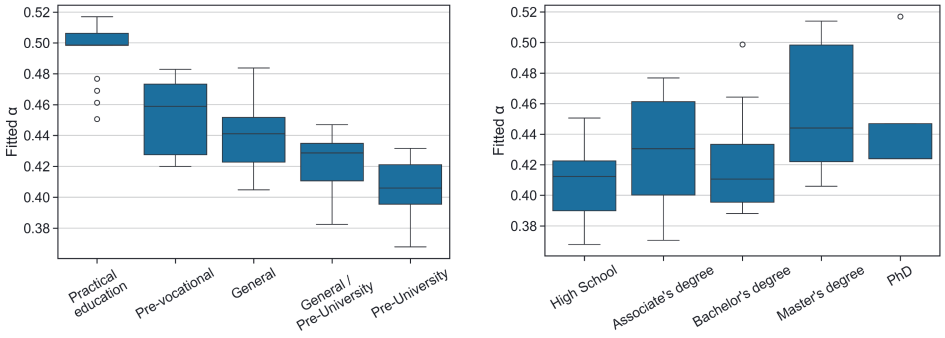


FIGURE 4.3: Distribution of fitted risk preference parameters α_i across respondents. Higher α indicates greater emphasis on potential vs. security.



(a) School Type

(b) Parent Education

FIGURE 4.4: Distribution of fitted risk preference parameters α_i across demographic variables of respondents.

baseline yielded $\chi^2 = 258.55$ with $p < 10^{-50}$. The model also achieved a pseudo-R^2 of 0.276, and outperformed the baseline on model selection criteria (AIC: 689.1 vs. 937.6; BIC: 713.4 vs. 941.7). While none of the individual covariates reached conventional levels of statistical significance ($p > 0.05$), the overall model captures meaningful variation in strategy behavior, suggesting that the combination of features helps explain latent risk preferences.

Figure 4.3 shows the estimated distribution of individual risk parameters α_i. The distribution is bimodal, with one group of participants placing substantial weight on security (e.g., $\alpha_i < 0.45$), and another group moderately prioritizing potential outcomes (e.g., $\alpha_i > 0.45$). Very few respondents exhibit strongly risk-seeking behavior ($\alpha_i > 0.6$), even under favorable conditions, consistent with the observed rarity of aggressive strategic choices.

Covariate	**Estimate** (β)	**Effect on** α_i
Intercept	0.058	Baseline
Education	0.088	↑ Risk-seeking
Income	–0.023	No effect
School Type	–0.094	↓ Risk-seeking
Importance: Schooling	–0.020	Negligible
Importance: Top School	–0.00006	Negligible

TABLE 4.3: Estimated coefficients for covariates explaining heterogeneity in α_i.

Model	**Log-Likelihood**	**AIC**	**BIC**	**Pseudo-R^2**
SP/A	338.55	665.09	647.40	1.73
EU	−464.71	931.43	934.37	0.00
PT	−464.71	933.43	939.32	≈ 0.00

TABLE 4.4: Model comparison using log-likelihood, AIC, BIC, and pseudo-R^2 for SP/A, Expected Utility (EU) and Prospect Theory (PT) models.

We next examined which parent characteristics best explain variation in α_i. As shown in Table 4.3 and Figure 4.4a, school track preference emerged as a key predictor: parents whose children were applying to competitive programs tended to exhibit greater risk aversion, possibly reflecting heightened sensitivity to the consequences of poor placement outcomes. Education was weakly positively associated with risk-seeking, while income and attitudinal factors, such as the importance placed on elite schooling, showed no significant effects once other variables were accounted for.

These findings provide partial support for our third hypothesis (H3). While parents do become more strategic under low-risk conditions, their choices appear moderated by how they perceive and manage risk, not by education or income.

There are several plausible explanations for the lack of statistically significant individual predictors despite strong overall model fit. Firstly, the sample size ($n = 140$) may limit statistical power to detect modest effects. While correlations among covariates such as income, education, and aspiration are moderate, multicollinearity does not appear to be a primary concern (see Appendix C.5). Alternatively, the influence of these features may be context-dependent or mediated by unobserved factors, requiring further study to disentangle their role in shaping risk preferences in school choice settings.

We also compared the SP/A model against two baselines: Expected Utility (EU) and Prospect Theory (PT). Table 4.4 reports model fit statistics across specifications. The SP/A model achieves the highest log-likelihood (338.55), outperforming both EU and PT, which converge to −464.71. AIC, BIC, and pseudo-R^2 metrics likewise favor the SP/A model, indicating better in-sample fit.

4.5 DISCUSSION

Our research centers on examining the behavior of parents in school choice mechanisms that are not strategyproof, and under varying levels of risk. We observe that a considerable share of respondents chose to report their preferences truthfully, even when given information that made the risks and potential gains of strategic behavior transparent.

However, the statistically significant increase in strategic behavior from Scenario 2 to Scenario 3 suggests that parents are sensitive to changes in the strategic landscape, particularly when the cost of deviation (i.e. risk) is lowered. Indeed, the preference for List 2 over the more aggressive List 3, even under favorable conditions, indicates that risk aversion plays a central role in shaping behavior and exceeds the importance of reward. These patterns align with the qualitative predictions of SP/A theory: participants tend to adopt secure strategies when uncertainty is high and only transition toward potential-maximizing strategies when the environment becomes less risky.

While strategic behavior does increase when probabilities are disclosed, its uptake appears to be governed not merely by access to information but also by individuals' tolerance for risk which, crucially, is not uniform. High-SES parents, often presumed to be the primary beneficiaries of manipulable systems, did not disproportionately adopt aggressive strategies in our study. Instead, parents at the most competitive school levels showed higher levels of risk aversion. We do not find strong evidence that income or education systematically explains variation in strategic behavior.

These findings should be interpreted in light of key design constraints. In order to reach parents currently participating in the school choice process, we recruited participants through informational seminars for the ongoing placement year. To avoid compounding their cognitive burden, the survey was designed to be short and completable within five minutes. Many participants were already navigating a complex matching system and facing emotionally charged decisions; the idea of deliberately crafting a strategic submission added a further layer of stress that may have shaped both engagement and responses.

This difficulty in forming and executing strategies under uncertainty underscores the need for a behavioral model that goes beyond rational choice. The SP/A model provides a compelling explanation for this. Rather than treating parents as either fully rational or systematically biased, SP/A theory captures a more affective mode of decision-making, where aspirations are weighed against perceived threats. The model's better fit relative to Expected Utility and Prospect Theory supports the claim that this balancing act governs how families engage with school choice mechanisms. Importantly, the model reveals a bimodal distribution of risk preferences, suggesting that population-level behavior cannot be adequately understood through average tendencies alone.

From a policy perspective, these findings challenge the notion that limiting transparency is a necessary safeguard against strategic behavior. Rather than concealing information, school choice platforms might better support fairness by disclosing probabilities alongside tools that communicate the risks of manipulation in intuitive and context-sensitive ways. Designing for transparency need not mean encouraging exploitation: our findings suggest parents are highly risk-averse, and only shift to aggressive strategies when the perceived chance of success is exceedingly high—in our

case, 90%. Prior simulation studies in Amsterdam suggest that strategic reporting improves outcomes for only 1–2% of students [19], making this level of confidence in strategy success extremely unrealistic. This gap between perceived and actual risk highlights the opportunity to build interactive tools that display the probabilities of admission under different ranking strategies, explicitly including uncertainty and known error bounds. Such interventions may deter manipulation not by limiting information, but by making risk legible.

These findings should be interpreted in light of several limitations. Our results are based on responses to hypothetical scenarios, and real-world implementation involving more students and schools could involve additional dynamics. Nonetheless, the consistent tendency of parents to minimize the risk of non-placement suggests that these patterns may generalize in live settings where the consequences directly affect their children. The simplified three-school setup allowed us to isolate strategic risk preferences, but actual school choice involves more complex preferences, repeated decisions, heterogeneous stakes, and social learning. Future work should test similar probability disclosures in full-scale choice environments, ideally in collaboration with school districts, to assess how these patterns translate when participants have diverse preferences. The small sample size of low-SES respondents also constrains our ability to draw strong subgroup conclusions. Furthermore, because our sample is drawn from Amsterdam, results may reflect sociocultural norms specific to the Netherlands. School choice policies, however, are often adapted internationally, suggesting similar dynamics could emerge elsewhere. Future work should extend this study to multiple cities and policy settings to examine how institutional context, cultural factors, and mechanism design interact in shaping strategic behavior.

Overall, this work underscores the value of integrating behavioral theory into the evaluation of algorithmic systems, particularly in high-stakes domains like education where institutional trust and individual aspiration intersect. From a policy perspective, our findings suggest that the trade-off between high efficiency and manipulability of an algorithm may not be as sharp as previously thought, and that transparency, rather than fueling strategic behavior, may in fact mitigate it. This contributes to the broader debate on transparency in human–AI interactions, showing that its effects depend on context, the nature of the information disclosed, and user incentives. Furthermore, given the nexus between educational opportunity and equality, what merits further investigation is not only the use of AI in education, but also its application in the broader educational ecosystem, including school choice.

4.6 CONCLUSION

In this study, we investigated how probability transparency in school choice mechanisms influences parental strategic behavior. Using a survey of 140 parents in Amsterdam, we examined whether aspiration levels and socioeconomic background affect the tendency to manipulate school preferences under a non-strategyproof allocation method. While transparency increases strategic behavior, a large fraction of parents still choose truth-telling. Among those who deviate, most opt for mild, low-risk strategies rather than aggressive manipulation. To interpret these patterns, we modeled parental decision-making using the Security-Potential/Aspiration (SP/A) framework, which captures the affective trade-off between avoiding bad outcomes

and striving for aspirational goals. The SP/A model not only outperforms Expected Utility and Prospect Theory baselines, but also reveals a bimodal distribution of risk preferences—underscoring that population-level behavior cannot be explained by average tendencies alone. These findings have direct implications for the design of algorithmic school choice mechanisms, particularly when considering non-strategyproof methods that allow gaming. Rather than restricting transparency, policymakers should focus on communicating the risks of strategic behavior clearly and intuitively, enabling parents to make informed decisions. Additionally, school choice models should account for heterogeneous behavioral responses, as a uniform transparency policy may not work equally well for all families. Future work could explore behavioral modeling techniques such as agent-based simulations or reinforcement learning to understand better how parents interact with different allocation methods over time. Further validation through controlled experiments and comparative studies of non-strategyproof mechanisms would enhance our understanding of strategic behavior in realistic settings.

5

SEQUENTIAL LEARNING FOR ONE-SIDED MATCHING

Based on
Tasnim, M., Ghebreab, S., Acar, E.
Sequential Learning for One-Sided Matching under Preference Misreporting.
Manuscript under submission.

The previous chapters take the design space of one-sided matching as given: each chapter analyzes a specific matching method, or a specific modification of one, and asks how it behaves under strategic reporting. Across these chapters, the same trade-off shows up repeatedly: strategyproof methods are inefficient, efficient methods are manipulable, and the choice between them is often treated as binary. The classical matching methods each take one side of this binary choice, but the space between them is largely unexplored. We ask whether this design space is genuinely as constrained as the classical results suggest, or whether learning-based methods can reach intermediate positions that classical methods cannot. This shift from analyzing existing methods to exploring the design space itself motivates the fourth research question of this thesis:

Research Question 4: *Can learning-based methods explore alternative tradeoffs between efficiency and strategyproofness?*

We address this by formulating one-sided matching as a sequential decision problem on a bipartite graph and training two learned policies: a Generative Flow Network that samples a distribution over high-reward matchings, and a Deep Q-Network that constructs a single high-value matching per profile. Because learned policies admit no analytical strategyproofness results and exhaustive incentive analysis is computationally intractable, we develop a framework for empirical robustness evaluation built around four misreporting strategies that vary in behavioral plausibility and search strength, complemented by a population-level adoption experiment that measures outcomes for strategic and truthful agents separately as the strategy spreads. Applied to Random Serial Dictatorship, Probabilistic Serial, Rank Minimization, and the two learned methods, the framework yields three findings that complicate the classical picture: robustness to preference misreporting is multi-dimensional, with a method that resists one type of misreport potentially exposed to another; learned methods remain stable under population-level strategic adoption where classical methods deteriorate for both strategic and truthful agents; and truthful efficiency is not a reliable indicator of robustness, with methods at similar efficiency levels differing substantially in which misreports they are vulnerable to.

5.1 INTRODUCTION

In one-sided matching, agents are assigned to indivisible items based on the agents' strict ordinal rankings over those items. The problem has been studied extensively in economics, social choice theory, and multi-agent systems, and arises in a wide range of practical allocation settings such as school choice, public housing allocation, and course assignments [1, 3]. A matching method for this problem should ideally assign agents to items they prefer highly, while also giving them little reason to misrepresent their preferences. These two goals are particularly difficult to reconcile in one-sided matching, where no method can achieve both efficiency and strategyproofness simultaneously [2]. Methods with strong incentive guarantees therefore tend to sacrifice efficiency, and methods that improve efficiency open up opportunities for agents to misreport their preferences in their favor [35, 86].

The three standard one-sided matching methods illustrate this tension cleanly. Random Serial Dictatorship (RSD) is strategyproof but can produce inefficient allocations [1, 7]. Probabilistic Serial (PS) and Rank Minimization (RM) are ordinally efficient and rank efficient respectively, but neither is strategyproof [3, 19]. Each method occupies a position in the efficiency-strategyproofness trade-off, and a natural question is whether one can construct matching methods that occupy different positions in this trade-off, beyond the set of available methods.

This chapter approaches this question from a machine learning perspective. One-sided matching is a natural setting for sequential decision-making: agents can be matched one at a time, partial matchings define meaningful intermediate states, and the set of feasible next assignments shrinks deterministically as the matching is built up. Despite this fit, learning-based methods for one-sided matching with ordinal preferences remain largely unexplored. We address this gap by formulating one-sided matching as a sequential decision problem on a bipartite graph, in which a policy builds a matching step by step. We then introduce two learned matching methods for this setting, one based on Generative Flow Networks (GFlowNet) [87] and one based on Deep Q-Networks (DQN) [88]. GFlowNet is well-suited to one-sided matching because it learns a distribution over high-reward objects rather than converging on a single optimum. This matches a structural property of matching that single-optimum methods do not exploit: many distinct matchings can achieve similar efficiency while differing in their distributional and strategic properties. DQN serves as a sequential-learning baseline that constructs a single high-value matching per profile, allowing us to isolate the contribution of distribution-aware sampling.

Introducing learned methods, however, comes with the methodological difficulty of evaluating strategyproofness. Classical matching methods come with analytical strategyproofness results: RSD is provably strategyproof, RM and PS are provably not. Learned matching policies admit no such guarantees, so their performance under preference misreporting has to be assessed empirically. The standard empirical approach of incentive-ratio analysis [89] enumerates over all possible misreports for every agent in every preference profile, and becomes factorially explosive as the problem grows. We therefore propose a framework that tests the robustness of matching methods to preference misreporting through a small set of structured misreporting strategies that can be evaluated efficiently. We test four strategies, chosen to span both behavioral plausibility and search strength:

(i) A popularity-based heuristic motivated by competitive school-choice settings [19], in which an agent keeps their top truthful choices and fills the rest of the report by item popularity.

(ii) A random adjacent swap, which probes sensitivity to arbitrary small perturbations of the truthful report.

(iii) The best adjacent swap, which identifies the most beneficial single local deviation [90].

(iv) The best permutation of the top-h ranks, which identifies the most beneficial structured reordering of an agent's most preferred items.

For each method and strategy, we report the fraction of agents who can gain by misreporting and the average gain, as a method might give many agents small gains or a few agents large ones. We evaluate one deviating agent at a time, while everyone else reports truthfully. In reality, though, a misreporting strategy that is profitable for a lone agent may have very different effects when many agents adopt it at once. To capture this, we complement the individual experiments with a population-level adoption experiment. We vary the fraction of agents who adopt each strategy and measure outcomes separately for strategic adopters, for agents who continue to report truthfully, and for the overall population. Although the framework was designed to evaluate learned methods, it applies also to classical methods, which lets us compare both kinds of matching methods on their performance under misreporting. We apply it to RSD, PS, RM, GFlowNet and DQN methods on synthetic preference profiles with $n = 10$ averaged across five dataset seeds.

We report three main findings. First, robustness to preference misreporting is multi-dimensional. A method that resists one class of misreport may be vulnerable to another: the GFlowNet policy, for example, is robust to the popularity heuristic but vulnerable to the best adjacent swap. Second, learned methods provide greater stability than classical methods under population-level strategic adoption. As the popularity heuristic spreads through the population, GFlowNet and DQN remain notably more stable than RSD, PS, and RM, and this stability holds for strategic adopters and truthful agents alike. Third, efficiency under truthful reporting is not a reliable indicator of robustness to misreporting. RM achieves the best truthful efficiency but is the most exposed to the popularity heuristic, and methods with similar rank performance can differ substantially in their vulnerability to specific misreporting strategies.

The contributions of this chapter are therefore threefold:

- We formulate one-sided matching as a sequential decision problem on a bipartite graph and introduce two learned matching methods for this setting, one based on GFlowNet and one on DQN.

- We develop a framework for measuring the robustness of matching methods to preference misreporting. The framework combines four misreporting strategies and a population-level adoption experiment. It is computationally tractable, applies to both classical and learned methods and reveals structural differences between methods that the binary notion of strategyproofness cannot capture.

- We apply the framework to RSD, PS, RM, GFlowNet and DQN methods, and find that the learned methods occupy distinct positions in the efficiency-robustness landscape, most notably providing greater stability under population-level strategic adoption than classical methods.

5.2 RELATED WORK

Preference Misreporting. A growing body of work studies strategic behavior in matching through restricted or behaviorally motivated misreports rather than through full strategyproofness. The existence of some profitable misreport does not by itself indicate whether agents are likely to find it or benefit from it substantially. Shirakawa [91] argues that full strategyproofness may be too demanding when participants are limited in their ability to identify complex profitable misreports, and studies immunity to simple misreports in school choice. Mennle et al. [90] study local manipulation strategies in assignment methods including PS, and show experimentally that such strategies capture much of how human agents approach the misreporting problem. Nesterov et al. [92] compare non-strategyproof school-choice methods through a robustness criterion rather than treating strategyproofness as a binary property. A complementary line studies worst-case incentive bounds via the *incentive ratio* introduced by Chen et al. [89], which quantifies the worst-case ratio between an agent's utility under their best misreport and under truthful reporting. Wang et al. [93] apply this concept to PS and show that the ratio is bounded. Our framework follows the behavioral and restricted-misreport perspective rather than the worst-case one. The popularity-based heuristic is closest in spirit to Tasnim et al. [19], in which agents preserve a small set of top choices and use information about item popularity to construct the rest of the report. We extend this idea to a comparison of classical and learned matching methods under several restricted misreport classes.

Machine Learning Methods for Matching. Machine learning has been applied to several matching problems. A first line of work studies learning-based methods for two-sided matching. Ravindranath et al. [94] use deep learning to characterize the trade-off between stability and strategyproofness in two-sided markets, and Maruo et al. [95] learn a strategyproof two-sided matching method from example matchings. A second line uses multi-agent reinforcement learning to find stable matchings in two-sided settings, either through decentralized agent interaction [96] or through a planner interacting with two-sided agents in a Markov matching market [97, 98]. A third line studies online bipartite matching, in which vertices arrive over time and assignments must be made under uncertainty; Alomrani et al. [99] train reinforcement learning policies for edge-weighted and submodular online bipartite matching. A fourth line uses reinforcement learning for combinatorial assignment problems such as the quadratic assignment problem [100], in which solutions are constructed sequentially but the inputs are pairwise costs rather than ordinal preferences. Related work in learning-augmented mechanism design has examined whether neural models can learn high-welfare rules while respecting incentive constraints in voting and social-choice problems [101].

Each of these settings is structurally different from one-sided matching as we study it: two-sided matching introduces preferences on both sides and is organized around

stability rather than efficiency; online matching considers arrivals over time and does not assume full preference profiles; QAP optimizes cardinal pairwise costs rather than ordinal preferences. Learning-based methods for static, one-sided matching with ordinal preferences have not, to our knowledge, been systematically studied. A common concern across these works is that learned policies do not inherit the formal incentive guarantees of classical methods, which motivates the empirical robustness framework we develop.

5.3 PRELIMINARIES

We formally define the one-sided matching problem, the efficiency metric and the classical matching methods used as baselines. We also summarize the GFlowNet and DQN frameworks used in our learned methods. The learned sequential policies are described in Section 5.4 as a part of our main contributions.

5.3.1 *One-sided matching.*

Let $A = \{a_1, a_2, \ldots, a_n\}$ be a set of n agents and let $I = \{i_1, i_2, \ldots, i_n\}$ be a set of n indivisible items. Each agent $a \in A$ has a strict and complete preference ordering R_a over the items in I. A preference profile is denoted by $R = (R_a)_{a \in A}$. A deterministic matching is a set of pairs $\mu \subseteq A \times I$ such that each agent is matched to exactly one item and each item is matched to exactly one agent. We write $\mu(a)$ for the item assigned to agent a.

A matching mechanism maps a reported preference profile to either a deterministic matching or a distribution over matchings. The latter case applies whenever the matching method induces a distribution over outcomes given fixed reports. We denote a mechanism by M, and write $M(R)$ for the matching, or distribution over matchings, induced by the reported profile R.

For an agent a and item i, let $r_a(i) \in \{1, 2, \ldots, n\}$ denote the position of item i in agent a's (true) ranking, where smaller values indicate more preferred items. We also associate a utility weight $w_{a,i} = n - r_a(i) + 1$ with each agent-item pair, which converts ranks into utility values such that the top-ranked item has weight n and the lowest-ranked item has weight 1. For a deterministic matching μ, the average rank (AR) is defined as follows:

$$AR(\mu, R) = \frac{1}{n} \sum_{a \in A} r_a(\mu(a)). \tag{5.1}$$

For stochastic methods, we evaluate expected average rank:

$$\mathbb{E}_{\mu \sim M(R)}[AR(\mu; R)]. \tag{5.2}$$

In the experiments, this expectation is estimated by Monte Carlo rollouts when the matching method is stochastic. We report normalized average rank, AR/n, so that allocation quality is comparable across market sizes. Lower values indicate that agents are assigned, on average, to more preferred items.

5.3.2 *Classical One-Sided Matching Methods*

Random Serial Dictatorship (RSD). Random Serial Dictatorship assigns items according to a randomly drawn priority ordering over agents [1]. Agents are considered one by one according to this order, and each agent receives their most preferred item among those still available. RSD is strategyproof, since an agent cannot improve their outcome by misreporting their preferences when their priority position is fixed.

Probabilistic Serial (PS). Probabilistic Serial produces a *fractional assignment* $p = (p_{a,i})_{a \in A, i \in I}$, where $p_{a,i} \in [0, 1]$ is the probability that agent a is assigned item i, with $\sum_i p_{a,i} = 1$ for each agent and $\sum_a p_{a,i} = 1$ for each item [3]. The fractional assignment is computed using the simultaneous-eating algorithm: each agent is treated as continuously consuming probability mass from their most preferred available item at unit rate, all agents eating in parallel. When an item's total probability mass is fully consumed, the agents currently eating it move on to their next most preferred available item, and so on, until every agent has accumulated total probability 1. PS is ordinally efficient as a fractional assignment, but it is not strategyproof. Given the fractional assignment p, the expected rank of agent a is $\sum_i p_{a,i} \cdot r_a(i)$. In our experiments, we report PS performance using this expected rank directly, and obtain deterministic matchings when needed via a Birkhoff–von Neumann decomposition of the fractional assignment.

Rank Minimization (RM). Rank Minimization selects a matching that minimizes the total rank of the assigned items:

$$\mu^{RM} \in \arg\min_{\mu} \sum_{a \in A} r_a(\mu(a)). \tag{5.3}$$

We implement RM using the Hungarian algorithm for the assignment problem [102]. RM directly optimizes the efficiency metric used in our experiments, and therefore serves as a strong efficiency-oriented baseline. At the same time, because RM optimizes the reported preference profile, it can create opportunities for strategic misreporting [19].

5.3.3 *Generative Flow Networks*

Generative Flow Networks (GFlowNet) are generative models that learn to sample structured terminal objects with probability proportional to a non-negative reward [87]. The generative process is modeled as a directed acyclic graph (DAG) over states. A unique initial state s_0 has no incoming edges, and terminal states x have no outgoing edges and correspond to the objects of interest. A trajectory $\tau = (s_0 \to s_1 \to \cdots \to s_T = x)$ is a sequence of states starting at s_0 and ending at a terminal state x. A *forward policy* $P_F(s' \mid s; \theta)$ assigns a probability to each transition out of a non-terminal state and induces a distribution over trajectories, and a *backward policy* $P_B(s \mid s'; \theta)$ assigns a probability to each transition into a non-initial state. A non-negative reward function $R(x)$ is defined on the terminal states, and a state flow function $F(s; \theta)$ assigns a non-negative flow value to each non-terminal state.

A GFlowNet is trained so that the forward policy samples terminal states with probability proportional to their reward:

$$P_F(x;\theta) \propto R(x), \tag{5.4}$$

where $P_F(x;\theta)$ denotes the total probability that the forward policy generates a trajectory ending at terminal state x. In contrast to reward-maximizing reinforcement learning, which converges toward a single high-reward terminal state, a GFlowNet that satisfies Equation 5.4 samples a distribution over high-reward terminal states.

A GFlowNet can be trained to satisfy Equation 5.4 through a local consistency condition between the forward policy, the backward policy, and the state flow function at every transition. When intermediate rewards are available along a trajectory, the *forward-looking* detailed balance objective [103] uses these intermediate signals to improve credit assignment. This objective extends R from terminal states to all states: at terminal states the per-state reward coincides with the terminal reward in Equation 5.4, while at non-terminal states it serves as an intermediate signal that shapes credit assignment along the trajectory. For a transition $s \to s'$, the forward-looking detailed balance loss is

$$\begin{aligned}\ell_{\mathrm{FL}}(s \to s';\theta) = \big(&\log R(s) + \log F(s;\theta) + \log P_F(s' \mid s;\theta) \\ &- \log R(s') - \log F(s';\theta) - \log P_B(s \mid s';\theta)\big)^2,\end{aligned} \tag{5.5}$$

At terminal states, F is fixed to one, since the full reward is captured in R at that state. At a global minimizer of this objective, the forward policy satisfies Equation 5.4 exactly [103].

5.3.4 *Deep Q-Networks*

Deep Q-Networks (DQN) are a reinforcement learning method that learns an action-value function $Q(s,a)$ representing the expected return from taking action a in state s and following the learned policy thereafter [88]. The Q-function is parameterized by a neural network with weights θ and trained to satisfy the Bellman optimality equation

$$Q(s,a;\theta) \approx r(s,a) + \gamma \max_{a'} Q(s',a';\theta^-), \tag{5.6}$$

where s' is the state reached after taking action a in state s, $r(s,a)$ is the immediate reward, $\gamma \in [0,1]$ is a discount factor, and θ^- denotes a periodically updated target-network copy of the parameters. Training proceeds by drawing transitions from an experience replay buffer and minimizing the squared residual of Equation 5.6. At evaluation time, actions can be selected greedily as $\arg\max_a Q(s,a;\theta)$ or stochastically by sampling from a softmax distribution over Q-values.

5.4 MATCHING AS A SEQUENTIAL DECISION PROBLEM

One-sided matching admits a natural sequential formulation. A matching can be constructed one edge at a time, with each step selecting an agent-item pair and removing the conflicting alternatives. Each intermediate stage of this process is itself

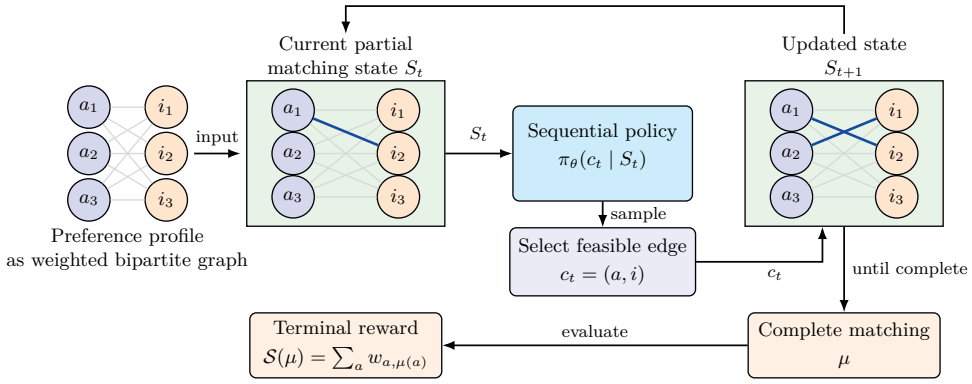


FIGURE 5.1: **One-sided matching as a sequential decision problem.** A preference profile is represented as a weighted bipartite graph between agents and items. At each step, a learned policy observes the current partial matching state and selects a feasible agent–item edge. The environment updates the state by adding the selected match and removing conflicting edges. A terminal score based on rank utility is assigned once a complete matching is formed. Both the GFlowNet and DQN policies use this sequential formulation.

a well-defined object: a partial matching together with the set of feasible assignments that remain. Casting the problem in this way lets us train policies that build matchings step by step, conditioning each assignment on the choices made so far. We formalize this view as a Markov decision process (MDP) on the bipartite graph of agents and items, and use it to instantiate two learned matching policies: a Generative Flow Network policy and a Deep Q-Network policy.

5.4.1 *Matching MDP*

We formulate one-sided matching as a sequential decision problem on the bipartite graph $G = (A \cup I, E)$, where agents and items form the two node sets and edges encode candidate assignments. A policy constructs a matching by selecting feasible edges one at a time. Two edges (a, i) and (a', i') *conflict* if they share an endpoint, that is, if $a = a'$ or $i = i'$. After an edge is selected, all edges that conflict with it are removed from the set of feasible candidates. This process continues until a complete one-to-one matching is formed. Given n agents and n items, the matching MDP is defined as follows:

States. A state s represents a partial matching through an $n \times n$ matrix over agent-item edges. The entry at (a, i) holds the utility weight $w_{a,i}$ if the edge is either still feasible or has been selected, and 0 if the edge has been removed. The initial state has all entries equal to the corresponding $w_{a,i}$. After an edge (a, i) is selected, every other entry in row a and column i is set to zero, leaving (a, i) as the unique nonzero entry in both its row and its column. Selected edges are therefore identified structurally: an edge (a, i) has been selected if and only if its entry is nonzero and it is the only nonzero entry in both row a and column i. A feasible action at state s is any edge

(a, i) whose entry is nonzero but is not yet a row-and-column singleton; we denote the set of feasible actions by $\mathcal{A}(s)$. A terminal state is reached when every agent has been matched to exactly one item and every item to exactly one agent, that is, when every nonzero entry is a singleton in its row and column.

Actions. At each step, a policy selects a single action $c = (a, i) \in \mathcal{A}(s)$. After the action is taken, the entries at all other positions in row a and column i are set to zero, enforcing the one-to-one matching constraint.

Terminal score. The terminal score of a complete matching μ is the sum of the utility weights of the selected edges:

$$\mathcal{S}(\mu) = \sum_{a \in A} w_{a,\mu(a)}. \tag{5.7}$$

Since utility weights are derived from ranks, maximizing $\mathcal{S}(\mu)$ is equivalent to minimizing the sum of assigned ranks. The learned policies are trained to favor high-utility, low-rank matchings. The terminal score depends only on efficiency; strategic behavior is evaluated separately, after training.

5.4.2 *GFlowNet matching policy*

We instantiate the GFlowNet framework from Section 5.3.3 on the matching MDP. The terminal objects are complete one-to-one matchings, and the reward on a terminal matching μ is the terminal score $\mathcal{S}(\mu)$ defined in Equation 5.7. A rollout starts from the initial graph state and constructs a matching by repeatedly sampling feasible edges under the learned forward policy. A trajectory is a sequence

$$\tau = (s_0 \to s_1 \to \cdots \to s_T), \tag{5.8}$$

where s_0 is the initial state and s_T is a terminal state corresponding to a complete matching μ. The forward policy $P_F(s_{t+1} \mid s_t; \theta)$ assigns probabilities to the feasible actions in $\mathcal{A}(s_t)$ at each step. After training, sampling from this policy produces a distribution over complete matchings, biased toward those with higher terminal scores.

We parameterize the GFlowNet policy with a graph neural network (GNN) over the bipartite matching graph. A GNN is a natural choice here: the matching MDP is defined on a bipartite graph, and the policy must respect the permutation invariance of agents and items, which GNNs handle by construction through message passing on the graph structure. Specifically, we use a Graph Isomorphism Network (GIN) [104], which is among the most expressive message-passing GNN architectures and has been used effectively for graph combinatorial optimization with GFlowNets [105].

Each edge (a, i) is associated with a single scalar feature, the current state entry at (a, i), which encodes both the utility weight (when the edge is still active) and the removal status (when the edge has been removed). The network outputs a score for each feasible edge, which is converted into a probability distribution over $\mathcal{A}(s)$ via a masked softmax. We train the GFlowNet using the forward-looking detailed balance objective defined in Equation 5.5. The reward-scaling parameter β controls how strongly the GFlowNet biases sampling toward high-scoring matchings: larger

values concentrate probability mass on near-optimal matchings, while smaller values give more uniform sampling. We use $\beta = 500$ in the main experiments based on preliminary validation runs. Additional details on the GFlowNet architecture and hyperparameters are given in Appendix D.1.

5.4.3 *DQN matching policy*

We instantiate the DQN framework from Section 5.3.4 on the matching MDP. DQN learns a value function over partial matching states and constructs a single high-value matching per profile, in contrast to GFlowNet, which samples across multiple high-reward matchings. Comparing the two within the same matching MDP lets us study whether the diversity-aware sampling of GFlowNet confers advantages that single-optimum sequential learning does not.

The DQN baseline uses the same matching MDP as the GFlowNet policy, but with two differences in how the reward signal is handled during training. First, we use intermediate per-step rewards: at each transition, the reward is the utility weight of the selected edge, normalized by the market size n for training stability:

$$r(s, c) = w_{a,i}/n, \quad \text{where } c = (a, i). \tag{5.9}$$

Summed undiscounted over a full trajectory, this gives $\mathcal{S}(\mu)/n$ where $\mathcal{S}(\mu)$ is the terminal score defined in Equation 5.7. Second, we apply a discount factor $\gamma < 1$ to the Bellman target during training. This helps stabilize Q-value targets when the magnitude of the cumulative return varies across episodes.

The Q-network is parameterized by a GNN over the bipartite matching graph. Agent and item nodes are first embedded by a linear node encoder, then refined through two graph convolution layers that aggregate information from neighboring nodes via message passing. For each feasible edge (a, i), the network combines the embeddings of agent a and item i with two scalar features: the current state entry at (a, i) and the original utility weight $w_{a,i}$, and passes the concatenation through a small MLP to produce a scalar Q-value.

At training time, actions are selected via ϵ-greedy exploration with ϵ decayed from 1.0 to 0.05. At evaluation time, we report results under the deterministic policy $\arg\max_{c \in \mathcal{A}(s)} Q(s, c)$. Further implementation details and hyperparameters for the DQN baseline are provided in Appendix D.2.

5.5 EVALUATING ROBUSTNESS TO PREFERENCE MISREPORTING

The classical mechanism-design tool for evaluating strategic vulnerability, known as the incentive ratio [89], requires enumerating possible misreports for every agent in every profile and is computationally intractable for matching markets. We instead measure robustness empirically through a small set of structured misreporting strategies. Each strategy modifies an agent's truthful report in a different way; we then compare the agent's outcome under the modified report with their outcome under truthful reporting.

5.5.1 *Measuring strategic impact*

Recall from Section 5.3.1 that each agent a has a utility weight $w_{a,i}$ associated with each item i, derived from their true ranking R_a. A report is truthful if it ranks items in decreasing order of utility: $r_a(i) < r_a(j) \iff w_{a,i} > w_{a,j}$. A misreport is any reported ordering R'_a that differs from the truthful ordering R_a. When evaluating a misreport, we always compute utility or rank with respect to the agent's true preferences, not with respect to the reported ranking. For brevity, we write $u_a(\mu) = w_{a,\mu(a)}$ for the utility agent a receives under matching μ.

A matching method is strategyproof if no agent can obtain a strictly better outcome by misreporting their preferences [86]. For a deterministic method, this means that for every agent a, true profile (R_a, R_{-a}), and possible misreport R'_a,

$$u_a(M(R_a, R_{-a})) \geq u_a(M(R'_a, R_{-a})). \tag{5.10}$$

For stochastic methods, the comparison is between expected utilities under the induced distributions over matchings: $\mathbb{E}_{\mu\sim M(R_a,R_{-a})}[u_a(\mu)] \geq \mathbb{E}_{\mu\sim M(R'_a,R_{-a})}[u_a(\mu)]$.

In our empirical evaluation, we fix a misreporting strategy σ, to be defined in the next subsection. For a profile R and an agent a, let $\sigma_a(R)$ denote the report generated for agent a by applying strategy σ to the profile R. We then compare agent a's expected utility under truthful reporting with their expected utility when only agent a changes their report to $\sigma_a(R)$, while all other agents remain truthful.

Given a profile R, an agent a, and a misreporting strategy σ, we define the non-negative misreporting gain as

$$G_a^\sigma(R) = \max\left\{0,\, \mathbb{E}_{\mu\sim M(\sigma_a(R),R_{-a})}[u_a(\mu)] - \mathbb{E}_{\mu\sim M(R_a,R_{-a})}[u_a(\mu)]\right\}. \tag{5.11}$$

For deterministic methods, the expectation is taken over the deterministic outcome. For stochastic methods, the expectation is estimated over sampled outcomes.

We report two aggregate strategic-impact measures. First, the fraction of agents for whom the misreporting strategy is profitable:

$$F^\sigma(R) = \frac{1}{n}\sum_{a\in A} \mathbf{1}\{G_a^\sigma(R) > \epsilon\}, \tag{5.12}$$

where $\epsilon > 0$ is a small numerical tolerance. Second, the average misreporting gain per agent, normalized by market size:

$$\bar{G}^\sigma(R) = \frac{1}{n^2}\sum_{a\in A} G_a^\sigma(R). \tag{5.13}$$

The first factor of $1/n$ averages the gain across agents. The second factor of $1/n$ normalizes by market size, since the utility weights $w_{a,i}$ take values in $\{1,\ldots,n\}$ and an unnormalized gain would scale with n. This double normalization makes gains comparable across markets of different sizes.

We use four misreporting strategies to instantiate σ. They vary in behavioral plausibility and search strength: the popularity heuristic models a structured, behaviorally motivated form of strategic reporting, while the adjacent-swap and top-h strategies test sensitivity to local and structured deviations.

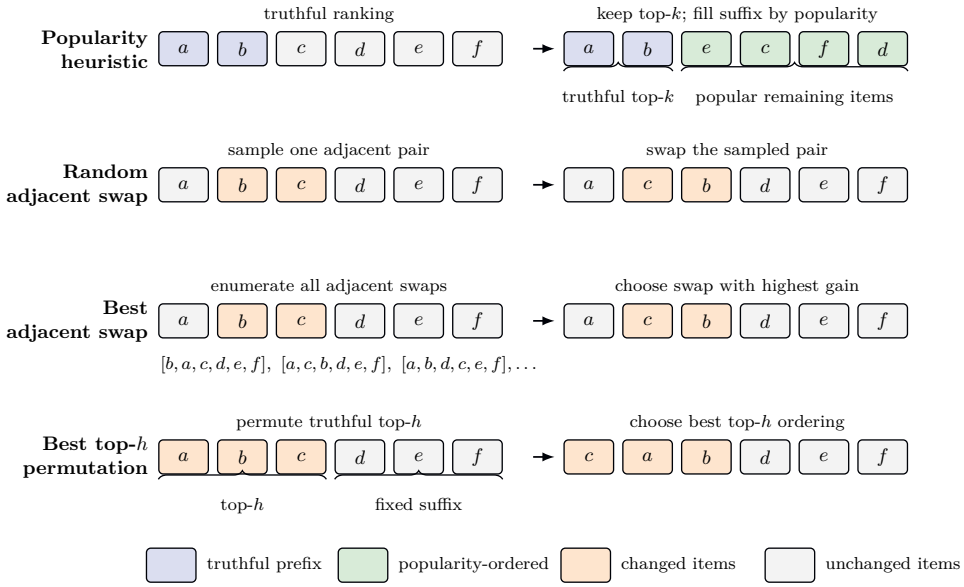


FIGURE 5.2: **Illustration of the four misreporting strategies.** Each row starts from the same truthful ranking $[a, b, c, d, e, f]$ and shows how the reported ranking is constructed under the corresponding strategy.

5.5.2 *Misreporting Strategies*

Popularity-based heuristic. The popularity heuristic captures a simple form of strategic reporting in markets where agents have a small number of genuinely preferred options and use the rest of the ranking strategically [19]. For agent a, the heuristic keeps the agent's first k truthful choices fixed and fills the remainder of the ranking with items ordered by global popularity. Item popularity is computed from the truthful profile as the number of agents who rank the item first.[1] The resulting report begins with agent a's truthful top-k items in their original order, followed by all remaining items in decreasing popularity order, with items in the top-k prefix excluded from the suffix so that the final report is a complete ranking without duplicates. This heuristic is the main behaviorally motivated misreporting strategy in the chapter and is also used in the population-level adoption experiment.

Random adjacent-swap perturbation. As a weak local-sensitivity baseline, we evaluate a random adjacent-swap perturbation. For agent a, we sample one adjacent pair of items in the truthful ranking uniformly at random and swap the two items. This measures whether a method is sensitive to arbitrary small perturbations of the reported ranking.

1 The popularity heuristic assumes the strategic agent has access to aggregate information about which items are most popular, motivated by school-choice settings.

Experiment	Main setting	Metrics
Truthful efficiency	$n = 10$, uniform preferences, 5 dataset seeds	Normalized average rank
Individual misreporting strategies	$n = 10$, uniform preferences, 5 dataset seeds	Fraction of profitable agents; normalized misreporting gain
Population adoption	Popularity heuristic only; $\rho \in \{0, 0.25, 0.5, 0.75, 1\}$; 3 repetitions per adoption rate	Overall, strategic-agent, and truthful-agent normalized average rank
Appendix robustness	Correlated-demand markets; alternative population strategy types	Same metrics as the corresponding main experiments

TABLE 5.1: Summary of the experimental regime.

Best adjacent swap. We also evaluate a local misreporting strategy based on adjacent swaps, similar to the local manipulation strategies introduced by Mennle et al. [90]. For agent a, we enumerate all rankings obtained by swapping one adjacent pair in the truthful ranking. We then evaluate the agent's expected utility under each candidate report, keeping all other agents truthful, and select the best candidate. This strategy measures vulnerability to the smallest non-trivial strategic deviation.

Best top-h prefix permutation. Lastly, we evaluate a stronger misreporting strategy based on permutations of the top of the ranking. For agent a, we enumerate all permutations of their truthful top-h items while keeping the remainder of the ranking fixed. We then select the report that gives the highest expected utility to the agent, again computing utility with respect to the agent's true preferences and keeping all other agents truthful. This strategy captures whether agents can benefit from reordering a small set of highly preferred items, without allowing arbitrary full-list misreporting.

5.6 EXPERIMENTS

We evaluate classical and learned matching methods under both truthful and misreported preferences, organized around three questions. First, how well do the methods perform when agents report truthfully? Second, how vulnerable are the methods to individual misreporting strategies? Third, how do outcomes change when a misreporting strategy is adopted by a larger fraction of the population?

5.6.1 *Datasets*

We generate synthetic preference profiles for one-sided matching tasks. Each profile consists of n agents and n items, where every agent has complete and strict preferences over all items. In the main experiments, we focus on small controlled markets with

$n = 10$. This setting allows us to evaluate truthful efficiency, individual misreporting strategies, and population-level strategic adoption under the same experimental regime, while keeping the counterfactual misreporting experiments computationally tractable.

For the main results, we use uniformly sampled preference profiles, in which each agent's preference ranking is drawn uniformly at random from all strict rankings over the item set. The uniform setting provides a neutral benchmark in which there is no imposed structure in item popularity. We use five independent dataset seeds, each containing 500 evaluation profiles. We also report robustness checks on preference profiles with correlated demand, indexed by a correlation parameter $\alpha \in \{0, 0.5, 0.9\}$, with larger values producing more concentrated demand. These additional experiments, reported in Appendix D.3, check whether the main findings are sensitive to the degree of preference correlation. Each preference profile is represented as a complete bipartite graph between agents and items, with edge weights given by the utility weights $w_{a,i}$ defined in Section 5.3.1.

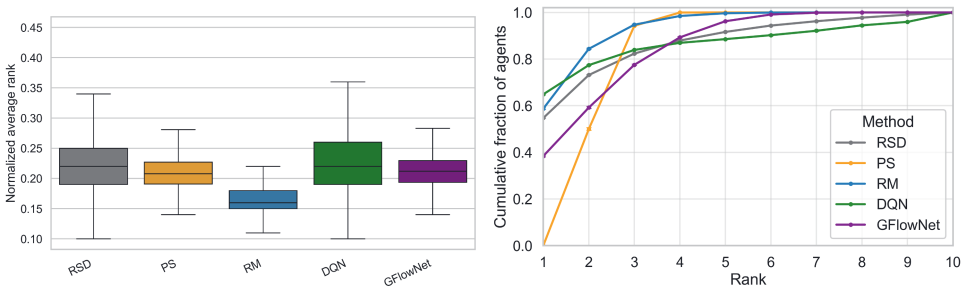
5.6.2 *Evaluation protocol*

We evaluate each method in three stages:

Efficiency under truthful reporting. We first evaluate the outcomes produced when all agents report truthfully. For RM and DQN, each profile produces one deterministic matching. For PS, we evaluate the expected average rank directly from the fractional assignment. For RSD and GFlowNet, which are stochastic, we estimate expected outcomes using Monte Carlo rollouts. We use 10,000 samples per profile for RSD and 50 samples per profile for GFlowNet. The main metric is normalized average rank, defined as the average rank divided by n. Lower values indicate that agents receive items closer to the top of their true preferences.

Vulnerability to individual misreporting. We next evaluate whether individual agents can benefit from unilateral deviations from truthful reporting. For each agent in turn, we hold all other agents truthful, replace the agent's report with a misreport, and compare the agent's expected utility under the two reports. Utility is always evaluated with respect to the agent's true preferences. We use the four misreporting strategies defined in Section 5.5.2: the popularity heuristic, random adjacent swap, best adjacent swap, and best top-h prefix permutation. We set $k = 1$ for the popularity heuristic and $h = 3$ for the top-h permutation. The $k = 1$ setting is aggressive: agents preserve only their top truthful choice before ordering the remaining items by popularity. We choose this setting because in our experiments with $n = 10$ the more efficient methods often assign agents within their top 2 choices; a larger k would leave much of the strategically relevant part of the report unchanged. Because these evaluations require additional counterfactual rollouts, we evaluate them on a fixed subset of 50 profiles per dataset seed. For each strategy, we report the two aggregate measures defined in Section 5.5.1: the fraction of agents with a profitable misreport and the average misreporting gain $\bar{G}^{\sigma}$ as defined in Equation 5.13.

Population-level effects of strategic adoption. Finally, we evaluate what happens when strategic behavior spreads through the population. The main population



(a) Distribution of normalized average rank.

(b) Cumulative rank distribution.

FIGURE 5.3: **Efficiency of evaluated matching methods.** Panel a shows the distribution of normalized average rank across instances; lower values indicate better assignments. Panel b shows the cumulative fraction of agents assigned within each rank threshold.

experiment uses the popularity heuristic, which is the most behaviorally grounded of the four strategies and produces the clearest population-level effects. For each profile, we choose a fraction ρ of agents uniformly at random and have those agents submit the popularity-heuristic report, while the remaining agents report truthfully. We vary $\rho \in \{0, 0.25, 0.5, 0.75, 1.0\}$, and for each adoption rate we repeat the random selection of strategic agents 3 times. Because population-level counterfactuals are more expensive than individual ones, we evaluate this experiment on 25 profiles per dataset seed. We then measure normalized average rank separately for three groups: the overall population, the strategic adopters, and the agents who continue to report truthfully. This decomposition lets us separate individual misreporting opportunity from systemic effects on truthful agents.

We also report population-level adoption results for the other three misreporting strategies in Appendix D.4.1, and robustness checks under correlated preference distributions in Appendix D.3.

5.7 RESULTS

We report results for truthful efficiency, individual misreporting strategies, and population-level effects under strategic adoption. For efficiency, we report normalized average rank, where lower values indicate better assignments. For misreporting strategies, we report both the fraction of agents with a profitable misreport and the normalized average misreporting gain. For population-level strategic adoption, we report the normalized average rank for the overall population, for strategic adopters, and for agents who remain truthful.

5.7.1 *Efficiency under Truthful Reporting*

Figure 5.3 reports efficiency and Table 5.2 gives the numerical values. RM achieves the lowest normalized average rank, as expected, because it directly minimizes the

Method	Normalized average rank (↓)
RSD	0.2228 ± 0.0005
PS	0.2088 ± 0.0003
RM	$\mathbf{0.1641 \pm 0.0003}$
DQN	0.2255 ± 0.0009
GFlowNet	0.2115 ± 0.0002

TABLE 5.2: Truthful efficiency. Lower values of normalized average rank indicate better assignments; the best result is shown in bold.

	Popularity		**Random adj.**		**Best adj.**		**top-h**	
Method	Frac.	Gain	Frac.	Gain	Frac.	Gain	Frac.	Gain
GFlowNet	**0.001**	**0.000**	0.045	0.002	0.307	0.016	0.244	0.015
DQN	0.030	0.009	**0.016**	0.006	**0.121**	0.046	**0.120**	0.048
RM	0.338	0.052	0.022	0.003	0.172	0.028	0.142	0.017
PS	0.004	**0.000**	0.018	**0.000**	0.151	**0.003**	0.143	**0.003**

TABLE 5.3: Effect of individual misreporting strategies. For each method and strategy, the fraction of agents with a profitable misreport (Frac.) and the normalized average misreporting gain (Gain) are reported. Lower values indicate greater robustness; the best result per strategy and metric is shown in bold.

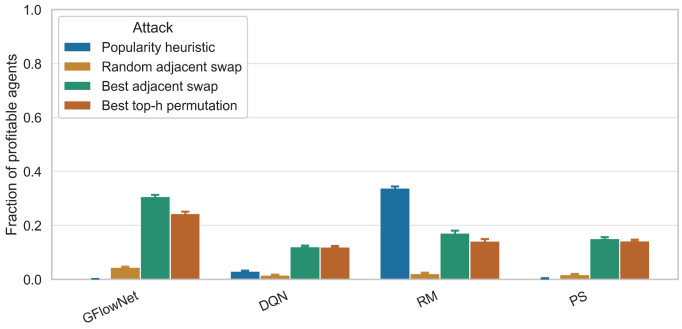
sum of assigned ranks. PS and GFlowNet are close to one another at an intermediate level of efficiency. RSD and DQN have higher median ranks and wider distributions.

The cumulative rank distribution in Figure 5.3b reveals where these differences come from. RM assigns a large fraction of agents to one of their top few choices. PS catches up quickly after the first few ranks, but assigns almost no agents to their first choice, an artifact of its fractional construction. GFlowNet improves more gradually across rank thresholds: it assigns fewer agents to the very top ranks than RM or PS, but reaches a high cumulative fraction by the middle of the ranking. DQN has a flatter curve, indicating that its assignments are more spread across ranks.

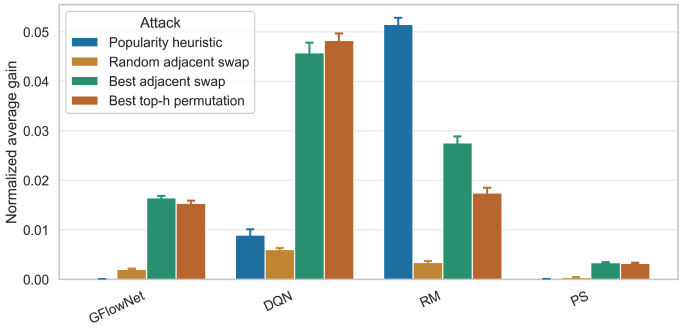
5.7.2 *Vulnerability to Individual Misreporting*

We next evaluate individual misreporting incentives using the four strategies defined in Section 5.5. For each agent, all other agents remain truthful, and the agent's gain is measured against the agent's true utility. Figure 5.4 reports both the fraction of agents with a profitable misreport and the normalized average misreporting gain. RSD is omitted because it is strategyproof. The corresponding numerical results are reported in Table 5.3.

Vulnerability depends strongly on the misreporting strategy. Under the popularity heuristic, RM has by far the largest fraction of profitable agents, consistent with the



(a) Fraction of profitable agents.



(b) Normalized average misreporting gain.

FIGURE 5.4: **Effect of individual misreporting strategies.** Panel a reports how often agents can benefit from each misreporting strategy. Panel b reports the average normalized misreporting gain.

fact that RM directly optimizes reported ranks and is therefore sensitive to structured changes in the reported preference profile. GFlowNet and PS are much less affected by the popularity heuristic, with under 1% of agents able to gain. Under the best adjacent-swap and best top-h permutation strategies, however, GFlowNet has a much larger fraction of profitable agents than under the popularity heuristic. DQN, RM, and PS also show nontrivial vulnerability under these optimized strategies, but the ordering differs from the popularity-heuristic case.

The same pattern is visible in misreporting gain, but with a different ranking. For the popularity heuristic, RM again stands out: it has both the largest fraction of profitable agents and the largest average misreporting gain. For the optimized local strategies, DQN exhibits relatively large gains, especially under best adjacent swaps and best top-h permutations. GFlowNet has a substantial fraction of profitable agents under these optimized strategies, but the average gains are smaller than under DQN. PS has relatively low misreporting gains across all four strategies.

Overall, RM achieves the best efficiency under truthful reporting but is highly exposed to the popularity heuristic. GFlowNet is comparatively robust to the popularity heuristic but less robust to optimized local strategies. DQN combines weaker truthful efficiency with sizeable misreporting gains under stronger strategies. PS is not

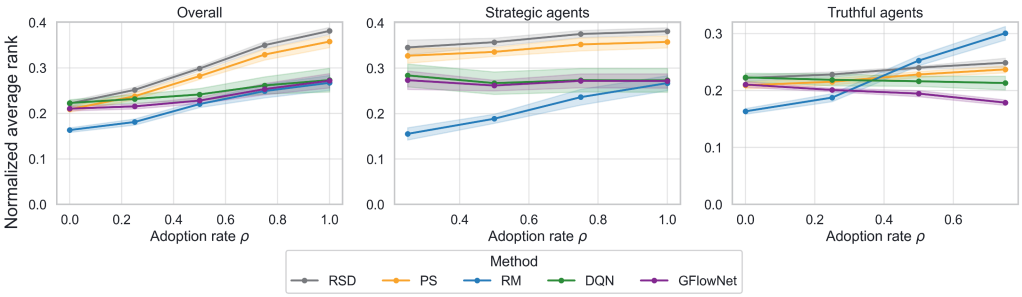


FIGURE 5.5: **Population-level adoption of the popularity heuristic.** Normalized average rank as the adoption rate ρ increases. The left panel reports the overall population outcome, the middle panel reports outcomes for strategic agents, and the right panel reports outcomes for truthful agents. Lower values indicate better assignments. The strategic-agent panel excludes $\rho = 0$, where no strategic agents exist; the truthful-agent panel excludes $\rho = 1$, where no truthful agents exist.

strategyproof, but in our experiments its misreporting gains are consistently small. Appendix D.3 reports the same strategies under correlated preference distributions.

5.7.3 *Population-Level Effects of Strategic Adoption*

The individual misreporting strategies evaluate one deviating agent at a time. We now ask what happens when strategic behavior spreads through the population. For each profile, a fraction ρ of agents is selected uniformly at random to submit the popularity-heuristic report, while the remaining agents report truthfully. We vary $\rho \in \{0, 0.25, 0.5, 0.75, 1.0\}$.

Figure 5.5 reports the resulting normalized average rank. The left panel shows the overall population outcome, the middle panel shows outcomes for adopters of the popularity heuristic, and the right panel shows outcomes for agents who remain truthful. The strategic-agent panel omits $\rho = 0$, where no strategic agents exist; the truthful-agent panel omits $\rho = 1$, where no truthful agents exist. Numerical values are reported in Table D.3 in the appendix.

Increasing adoption of the popularity heuristic worsens overall efficiency for every method, but to varying degrees. The deterioration is strongest for RSD, PS, and RM, whose normalized average rank increases substantially as ρ approaches one. DQN and GFlowNet are more stable in the aggregate, with only modest increases in normalized average rank as adoption grows.

The subgroup panels show that the aggregate pattern hides important differences between adopters and non-adopters. Under RM, strategic adopters initially receive favorable assignments, but their outcomes worsen as more agents adopt the same strategy. At the same time, truthful agents become worse off as adoption increases. RM is therefore sensitive to widespread popularity-based reporting on both sides: the strategy becomes less advantageous when it is common, while also imposing costs on truthful agents. Under GFlowNet and DQN, outcomes are more stable for both

groups: strategic adopters see only modest changes as ρ increases, and truthful agents remain relatively stable, with slightly improving outcomes in some cases. Under PS and RSD, strategic adopters fare poorly throughout, and truthful agents deteriorate gradually as adoption increases.

The same adoption process therefore has qualitatively different effects across methods. RM shows strong redistribution between strategic and truthful agents, GFlowNet and DQN are comparatively stable for both groups, and PS and RSD deteriorate for both. Appendix D.4.1 reports population-level adoption experiments for the other three misreporting strategies.

5.7.4 *Summary of Findings*

The experiments support three main findings:

First, *robustness to preference misreporting is multi-dimensional.* A method that resists one type of misreport could be vulnerable to another one. GFlowNet, for example, is almost entirely robust to the popularity heuristic but substantially exposed to the best adjacent swap. Worst-case bounds such as the incentive ratio [89] summarize vulnerability into a single number, but this hides the structural differences in how methods respond to specific deviation models.

Second, *learned methods provide greater stability than classical methods under population-level strategic adoption.* As the popularity heuristic spreads through the population, GFlowNet and DQN remain notably more stable in average rank of matches than RSD, PS, and RM, and this stability holds for both strategic adopters and truthful agents.

Third, *efficiency under truthful reporting is not a reliable indicator of robustness to misreporting.* PS and GFlowNet achieve similar truthful efficiency in our experiments, but PS is more exposed to optimized local strategies while GFlowNet is more exposed to behavioral heuristics; methods with similar rank performance differ substantially in which misreports they are vulnerable to.

5.8 DISCUSSION

Evaluating learned matching methods. Our findings have implications for how learned matching methods should be evaluated for preference misreporting. Preference misreporting has both an extensive and an intensive margin: a method may give many agents small profitable misreports, or fewer agents larger ones, and these have different implications for the strategic environment a method induces. Evaluating both quantities, together with multiple misreporting strategies that vary in plausibility and search strength, gives a more informative picture than any single number can. Worst-case bounds such as the incentive ratio [89] summarize vulnerability into a single number but cannot distinguish methods that fail in structurally different ways. Our framework is complementary, providing a tractable empirical approach for methods that lack analytical strategyproofness results.

Relative stability of learned methods under population adoption. The population-level stability we observe for the learned methods is the most distinctive empirical pattern in our results. This could be caused by the implicit regularization

provided by neural-network approximation: smooth responses to small input perturbations may keep outputs close to what the network learned on similar truthful profiles. It could also be due to partial matchings induced by sequential construction, which constrains later assignments by the graph structure already built rather than by the full reported profile. Our experiments do not isolate the cause, which we leave for future work.

Learned methods are more stable specifically under the popularity heuristic, where every adopter shifts probability mass toward the same items, generating systematic pressure on the matching method. The swap-based and top-h strategies, by contrast, produce individually optimized misreports that do not aggregate, and their population-level effects (Appendix D.4.1) are small. This could have practical implications: behavioral coordination on common information about popular items is the strategic pattern most likely to diffuse in real markets. Practical deployment of the learned methods would, however, raise additional concerns our experiments do not address, such as scalability and explainability.

The case for learning-based alternatives. These findings also speak to a longer-standing tension in the design of one-sided matching methods. The field is shaped by impossibility results: no method can be both efficient and strategyproof, and similar tensions exist between other desirable properties such as efficiency and envy-freeness, or between strategyproofness and stability in related settings. Classical methods occupy specific points on these trade-offs, but the space between them remains largely unexplored. Learning-based methods could in principle reach intermediate positions in this space. What they sacrifice is analytical guarantees: without formal proofs of strategyproofness or efficiency, their behavior must be characterized empirically. Our framework is an initial step in this direction, providing a tractable way to evaluate misreporting behavior; analogous frameworks could be developed for other design properties such as stability and envy-freeness. Two further directions also seem promising: extending the framework to richer allocation settings, such as many-to-one markets with capacities and priorities, and extending the misreporting strategies beyond fixed behavioral models to adaptive settings in which agents learn strategic behavior through repeated interaction with the matching method.

Limitations. Several limitations qualify our analysis. The main experiments focus on $n = 10$ markets, which supports controlled counterfactual evaluation. Larger markets may exhibit different strategic effects, and practical deployment at larger scales would encounter scaling challenges specific to the learned methods, since the matching MDP has n^2 edges and training cost grows accordingly. The experiments also use synthetic one-to-one matching markets, abstracting away from institutional features of real allocation systems such as capacities, priorities, incomplete lists, and placement guarantees. The learned methods are trained only on uniformly distributed truthful profiles; explicitly incorporating misreports or correlated-demand profiles during training might affect both their truthful efficiency and their robustness to misreporting. The strategic evaluation is also only based on a fixed set of misreporting strategies, which keeps the analysis tractable but does not characterize all possible misreports. We do not report systematic sensitivity analyses for the reward-scaling parameter β, the popularity-heuristic prefix length k, or the top-h permutation length

h; these were chosen based on preliminary validation runs, and a rigorous sensitivity study is left for future work.

5.9 CONCLUSION

We formulate matching as a sequential decision problem, introduce two learned methods based on Generative Flow Networks and Deep Q-Networks, and compare them against classical methods under truthful reports, individual misreporting strategies, and population-level strategic adoption. We report three main findings. First, robustness to preference misreporting is multi-dimensional: a method that resists one structured misreport can be exposed to another, and worst-case bounds such as the incentive ratio summarize vulnerability in ways that hide these structural differences. Second, learned methods are more stable under population-level strategic adoption than the classical methods we tested, even when their individual misreporting incentives are comparable. Third, efficiency under truthful reporting is not a reliable indicator of robustness: methods with similar rank performance can differ substantially in their vulnerability to specific misreporting strategies. The mechanism behind the learned methods' stability under population adoption remains an open question.

These findings position learning-based methods as a research direction worth pursuing alongside classical approaches. The field of one-sided matching is shaped by sharp impossibility results, and the standard classical methods occupy specific points in the corresponding trade-offs. Learning-based methods could in principle reach positions between these points, but doing so requires rigorous empirical evaluation in place of analytical guarantees.

6

REFLECTIONS ON CIVIC AI RESEARCH

Based on
Tasnim, M., Ghebreab, S.
AI, interdisciplinarity and civic engagement: Insights from Amsterdam school choice. Cambridge Forum on AI: Culture and Society **1**, e5 (2025) [23].

The previous four chapters approached one-sided matching from formal, institutional, behavioral, and computational angles, but each chapter rested on choices that go beyond what any single discipline could justify: which trade-offs to take seriously, which actors to model, which outcomes to count as harms, and which interventions to put on the table. The Amsterdam school choice case made these choices unavoidable, because the system shapes outcomes, incentives, administrative burdens, and institutional trust at the same time, and no one of these can be addressed without engaging the others. This raises a question that is no longer about matching specifically, but about the research practice that produced the previous chapters in the first place:

Research Question 5: *What do these findings imply for responsible AI research for civic purposes?*

We address this by reflecting on five years of work in the Civic AI Lab, with the Amsterdam school choice project as the central case. We argue that civic-centered AI research operates along three dimensions: between science and society, between fundamental and applied work, and between quantitative and qualitative methods, with the human at the intersection of all three. The previous chapters illustrate this in practice: the formal analysis in Chapters 2 and 3 raised questions that motivated the behavioral study in Chapter 4, the behavioral findings in turn motivated the computational exploration in Chapter 5, and at each step the framing of the research question came from sustained engagement with parents, school boards, and advocacy organizations rather than from within any single discipline. Drawing on this experience, we argue that interdisciplinarity in AI research is rarely sustained on its own terms; it is sustained by civic engagement, which gives researchers a concrete reason to cross disciplinary boundaries and a partner who holds them to it. Rather than treating civic engagement as a downstream concern once the science is settled, we position it as the practical condition that makes interdisciplinary AI research impactful.

6.1 INTRODUCTION

Artificial Intelligence (AI) has always been more than a purely technical endeavor. From its earliest days, it was deeply connected to questions about human intelligence, perception, and reasoning [106, 107]. Researchers drew inspiration from fields like psychology, philosophy, and neuroscience, aiming to understand and replicate the ways humans think [108]. Over time, as computational power increased and data became more abundant, AI evolved from rule-based logic systems to complex statistical and machine learning models [109, 110]. What began as an exploration of cognition has grown into a vast, interdisciplinary field touching every aspect of modern life.

Today, AI drives innovation across scientific research, business processes and public services [111–113]. With this growing presence comes growing responsibility. Ethical questions around fairness, transparency and accountability are faced by institutions around the world [114–116]. The question is no longer whether AI will shape the future, but whether AI will shape the future in the right way, and whether we are guiding AI's development in the right direction.

It is within this context that our research lab, the Civic AI lab, was established in 2020 as a partnership between universities and public institutions designed to address these urgent concerns. The lab was founded on a simple but powerful principle: AI should be developed not just *for* people, but also *with* and *by* people. This draws from the ethos of transdisciplinary research, which argues that knowledge production must occur through iterative engagement with affected communities, blurring the boundaries between science and society [117]. Rather than building systems in isolation, the lab aims to connect technological development with the lived experiences of the communities it affects. Our work is grounded in the belief that civic engagement and interdisciplinarity are essential for meaningful AI innovation.

One of the lab's earliest and most revealing projects focused on a seemingly local issue: school choice in Amsterdam [19]. Each year, thousands of primary school students are matched to secondary schools using a matching algorithm. The city uses the Deferred Acceptance algorithm [40, 118], adapted to use random lottery numbers due to the absence of schools' preferences over students as a result of Amsterdam's free school choice policy. This adaptation created unintended consequences: students with worse lottery numbers found themselves placed in schools far down their preference list, sometimes even as low as their 18th choice.

This problem is not unique to Amsterdam. It illustrates a broader tension at the heart of AI and algorithm design: the gap between theory and reality. The Deferred Acceptance algorithm is celebrated for its theoretical properties, but when adapted to the real-world setting of an open school choice policy, its effectiveness was compromised. Real-world deployment requires contextual awareness, stakeholder involvement, and a deep understanding of the social systems into which algorithms are introduced. Interdisciplinary collaboration, particularly between technical and social domains, has increasingly been positioned as necessary to bridge this divide [119]. However, interdisciplinarity itself is not without problems. It has often been critiqued for remaining superficial, poorly supported by institutional incentives, or blocked by entrenched disciplinary norms [119, 120]. This underscores the need to examine how interdisciplinary efforts can be integrative and impactful.

In this chapter, we reflect on five years of research at our research lab, drawing lessons from our work on school choice and other civic challenges. We found that trying to solve a concrete, socially impactful problem inevitably pushed us to cross disciplinary boundaries and engage directly with communities. Through this process, we came to see how community-informed, interdisciplinary approaches enrich scientific understanding, generating insights that feed back into existing theory. We observe that it is through the transdisciplinarity of civic engagement that interdisciplinary research can truly take shape and overcome many of the barriers that typically limit it. We argue that impactful AI must operate across three critical dimensions: between science and society, between fundamental and applied research, and between quantitative and qualitative methodologies. At the center of all three is the human, whose life, choices, and dignity must remain at the core of our work.

6.2 OUR RESEARCH PRINCIPLES

Our research work has been shaped by five guiding principles that define not just *what* we build, but *why* and *how* we build it. These principles challenge conventional AI development, which often prioritizes efficiency and profit over ethics and equity. Instead, we center human values frequently deemed elusive or hard to define, diverse perspectives, and participatory research as core elements of responsible AI.

The first principle, "plan and purpose," asks the fundamental question: why do we develop AI? Too often, AI is designed to maximize efficiency or profit, sidelining human well-being. Many scholars have long highlighted how top-down technology deployments often fail to account for local needs, reproduce global power imbalances, or reinforce existing social hierarchies [121–123]. Scholars of postcolonial and feminist epistemologies have similarly examined how technological interventions can perpetuate colonial or developmentalist agendas, treating communities as passive sites of experimentation [124, 125]. We aimed to take a different approach, embedding compassion, equity, and social good into the very design of our models. This goes beyond adjusting the objective function of an AI algorithm for balancing efficiency and fairness; it means working with communities from the start, setting research questions and priorities together, and designing systems alongside the people who will be most affected.

The second principle, "people and perspective," considers who develops AI and who is affected by it. AI is never neutral; it reflects the perspectives of those who create it. To ensure inclusive, ethical technology, we must build diverse teams that bring a range of lived experiences to the table. Without this balance, AI risks reinforcing exclusion rather than mitigating it.

The third principle, "parts and parcels," focuses on what goes into AI—its data, methodologies, and computational structures. AI systems are only as good as the data that train them. However, too often data are drawn from narrow populations, specific historical moments, or biased sources, leading to models that are neither fair nor representative. Ethical AI demands a critical examination of data quality, algorithmic bias, and the sustainability of computational infrastructures.

The fourth principle, "processes and procedures" asks how AI is developed. Transparency and accountability are often cited as priorities in AI governance but remain largely aspirational. True accountability means more than just publishing reports; it

requires open processes, clear explanations of decision-making, and mechanisms for redress when AI systems cause harm.

Finally, the fifth principle, "participation and practice," ensures that AI is not just built for people, but with and by people. Those who first-hand experience the consequences of AI must have a role in shaping its design. Inclusive, participatory research fosters more robust, trustworthy AI systems that reflect real-world needs and mitigate harms such as bias and discrimination.

These five principles have not only guided our broader work at our research lab but have also been central to our research on school choice and allocation algorithms. By embedding ethics, participation, and interdisciplinary collaboration into AI development, we strive to create technology that serves the public good, rather than just the bottom line.

6.3 THE CASE OF AMSTERDAM SCHOOL CHOICE

Our research began with a simple question: how can we use responsible AI to reduce inequality? Our education project centered around the city's secondary school admissions system. At first glance, this appears to be a technical optimization problem, matching students to schools based on their preferences. But a complex socio-technical challenge lies underneath the surface involving students, parents, policymakers, schools, and municipal authorities. It is a problem where technical solutions meet real-world stakes, where fairness and organizational trust matter just as much as efficiency.

Amsterdam's school choice policy is unique. Since 2015, the city has used an open school choice system, allowing students to apply to any secondary school in the city regardless of their residential location. This policy was implemented to promote equal access and parental choice, reflecting a broader commitment to educational equity [10, 11]. However, it also created significant challenges: with only a few schools being highly sought-after, demand became extremely skewed, making allocation both politically sensitive and socially impactful [12].

The current allocation mechanism is based on the Deferred Acceptance (DA) algorithm, a well-established method in the field of market design and school choice [28]. DA has been lauded for being strategyproof and fair under certain assumptions. However, in Amsterdam's case, the DA algorithm had to be adapted to fit the open-choice setting. Because schools do not rank students—either due to policy constraints or institutional indifference—student preferences are resolved through random lottery numbers. This modification, while necessary, significantly undermines some of DA's most desirable theoretical properties.

Most students do receive one of their top school choices. However, each year, a small but significant group of students are placed in schools they ranked extremely low, sometimes as far down as their 18th choice. These students and their families often express frustration and disillusionment with the system, sometimes even filing court cases [15]. For parents, the school choice process is not just a logistical hurdle; it is an emotionally taxing, high-stakes decision that shapes their child's future. Being placed far down the list feels like a personal loss and an institutional failure.

In the first phase of the school choice project, our work focused on analyzing this problem and testing alternatives. Our first study examined whether a different algorithm could better meet the goals of the system [19]. We proposed a method called

Rank Minimisation, which uses the Hungarian algorithm to minimize the average rank of school assignments. Unlike DA, this method is not strategyproof - it assumes that some students might game the system. However, our hypothesis was that improving average outcomes might justify the potential for manipulation.

To test this, we analytically derived a best-response strategy for rational agents and conducted simulations under different scenarios. The results were surprising. In 14 out of 16 scenarios, the Rank Minimisation method performed better or comparably to DA, even in the presence of strategic behavior. Only in two scenarios did it result in worse outcomes. This suggested that, under real-world constraints, relaxing the strategyproofness requirement might be a trade-off worth considering.

Yet, despite these findings, publishing this work proved challenging. Reviewers questioned its novelty and criticized its specificity to Amsterdam. Some valued its practical relevance, while others dismissed it as too applied or lacking theoretical innovation. Our work also had social and political implications, so we also had to manage complexities on the political and governance fronts. Our work advocated the exploration of non-strategyproof school choice methods, while the dominant view among both senior academics and policymakers (advised often by these same academics) was that this was too risky to attempt. On the other hand, when communicating our research to parents who were affected by the school choice system, we received feedback that alternative algorithms are worth at least a public discussion.

This experience raised broader questions about the role of applied research in AI and the value of studying local, context-dependent problems, as well as its role in navigating complex organizational landscapes to create lasting policy changes. It became clear to us that designing fair and effective algorithms required not only theoretical insight but also an understanding of how those algorithms are used, perceived, and experienced. We also understood the importance of such work to be disseminated among academics, policymakers and citizens. This realization marked a shift in our approach—from designing AI *for* people to designing AI *with* people. In the second phase of the school choice project, we turned our attention to understanding how parents engage with the school choice system. We designed a behavioral experiment to study the conditions under which parents might act strategically. The study was embedded in a survey and distributed to parents in Amsterdam who were preparing for their child's school allocation process.

In the experiment, we presented three scenarios. In the first, parents were asked to submit their preferences without any additional information. In the second, we provided probabilities of success for different manipulative strategies, with high risk and low reward. In the third, the probabilities remained risky but promised higher rewards. We also collected background information such as income, education, and aspiration levels, with the hypothesis that strategic behavior might correlate with these variables.

Conducting this study required sustained engagement with the local community. We worked closely with OCO (Onderwijs Consumenten Organisatie), a non-profit organization that supports parents navigating Amsterdam's education system. Our collaboration began during the early stages of our first project, when OCO reached out to learn more about our research on the city's school choice system. Given that much of OCO's advising work centers on the transition from primary to secondary education, they were deeply interested in how the school choice algorithm influenced matching

outcomes. In these initial conversations, OCO shared a range of concerns: the lack of transparency in the decision-making processes of school boards, and the broader social inequalities that persist irrespective of any particular matching algorithm. They highlighted how parents from lower socioeconomic backgrounds often face systemic disadvantages, including challenges in compiling complete preference lists or effectively navigating the school choice process. Throughout the first project, we maintained an ongoing dialogue with OCO, incorporating their perspectives into how we framed our research questions. This engagement was instrumental in shaping the second phase of our work. Together with OCO, we recognized the importance of capturing diverse parental perspectives and agreed to collaborate on a survey that OCO would help disseminate. Partnering with OCO not only enriched the design of our study but also helped us navigate a common barrier in academic research: public mistrust toward researchers, especially on sensitive topics like strategic school choice. Through OCO's established relationships and legitimacy within the community, we were able to reach parents more effectively and conduct the study with greater sensitivity and trust.

Our survey design had to balance academic rigor with real-world constraints: parents were often overwhelmed by the information and unsure about the implications of their choices. One key insight from this work was the uneven cognitive and emotional burden placed on parents. Those from higher-income backgrounds were more confident and engaged, while parents from marginalized communities expressed anxiety, distrust, and confusion. These reactions reflected broader systemic inequalities that any AI-based solution must acknowledge and address. Our observation and acknowledgment of these concerns not only deepened our understanding of the problem as a whole, it also gained the trust of parents and their support in our AI research.

Reflecting on both phases of this project, we found that the technical problem of school allocation cannot be solved through algorithms alone. The first phase taught us that improving technical performance is important, but insufficient if the solution doesn't account for human behavior and lived experience. The second phase reinforced the idea that research done together with communities improves trust and relevance and surfaces new insights that no algorithm can reveal alone. Building on this work, we are preparing a follow-up study that uses unstructured narrative interviews (the Biographical Narrative Interpretive Method, BNIM) [126] to engage more deeply with parents' experiences.

Together, these projects illustrate the need for a more expansive approach to AI research that embraces theoretical development, interdisciplinary knowledge, and civic engagement as complementary rather than competing forces.

6.4 REFLECTING ON OUR EXPERIENCE BUILDING CIVIC-CENTERED AI

In this section, we reflect on what our work on school choice taught us about doing responsible, civic-centered AI research, and how these lessons may be relevant to other domains. At our research Lab, we began our work—by design—with a diverse group of researchers and a strong objective: to develop AI that is responsible, inclusive, and driven by civic values.

From the onset, we were thus engaged with the fundamental questions of *who* builds AI and *why*: questions that lie at the core of our guiding principles of people and perspective and plan and purpose. As shown in Figure 6.1, these foundational

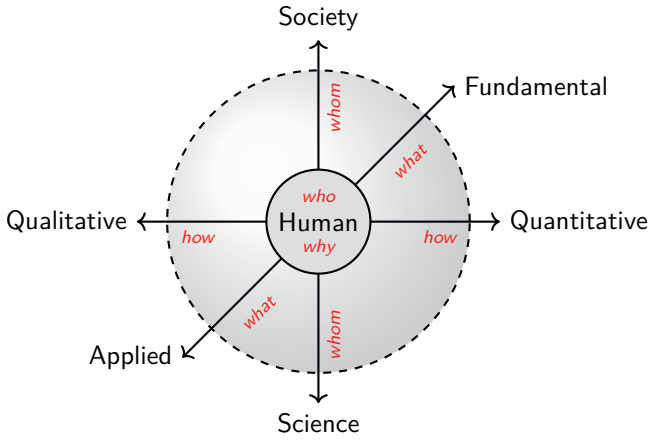


FIGURE 6.1: A spherical conceptual model representing the three dimensions of the practice of interdisciplinarity in civic-centered AI. Each axis spans a spectrum (quantitative–qualitative, science–society, fundamental–applied), tied to the following core questions: *how*, *what*, and *for whom*. The *who* and *why* reside at the center: the Human.

questions sit at the center of the conceptual space where all dimensions intersect: the human.

Particularly for the school choice project, our journey began with a technical question: could we improve an inefficient algorithm used in school allocation? It quickly became apparent that a technical lens alone was insufficient. What counts as a "good" allocation when systems operate under imperfect social and organizational conditions? The insights gained from our first project informed our second: we designed a behavioral experiment with parents to understand how they engage with the school choice system. Here, the lens widened again. The algorithm was not just a function to be optimized; it was embedded in a complex socio-technical system and needed to be designed accordingly.

We now understand our work as operating across three core dimensions, although we did not name them as such at the time. In our other projects and use cases—covering various social domains such as mobility [127], environment [128], governance [129], law [130] and co-creation [131]—these three dimensions of AI design also emerged organically: answering questions on *what*, *how*, and for *whom* we design AI. In Figure 6.1, these dimensions offer a framework through which to understand our interdisciplinary journey.

The first is the relationship between science and society, which relates to the question of: for *whom* do we design AI? Traditional views often hold that science must remain neutral and objective, keeping a deliberate distance from society while observing and analyzing it from a distance. AI, as a tool born out of scientific innovation, is also seen as separate and distinct from society. However, our experience reinforced that science and AI are both inevitably embedded in social contexts. Algorithms are not neutral artifacts; they reflect the assumptions and priorities of their creators. Civic-centered AI recognizes this and embraces the responsibility that comes with it. This reflection

resonates with the principle of *participation and practice*, which urges that those impacted by AI systems must be part of the design process from the start. Research should not be done merely on communities, but with them.

The second dimension is the interplay between fundamental and applied research, which connects to the question of: *what* are we designing? We often think of this as a one-way street: fundamental research lays the groundwork, and applied research implements it. But in our case, applied work raised new questions for theory, challenged core assumptions, and suggested new problem framing. While this bi-directional relationship between applied work and theoretical development emerged from our specific case, we believe that it carries a broader methodological lesson for AI research. Applied research enriches existing theories by revealing overlooked complexities, challenging core assumptions, and stimulating new lines of foundational inquiry. We believe that iterative movement between theory and application is therefore essential for progressing AI development.

The third dimension is the integration of qualitative and quantitative methods, which informs the question of: *how* should we design AI? AI has historically favored quantitative approaches; statistical models, optimization algorithms and predictive accuracy. But societal problems are rarely reducible to numbers alone. In our second project, we learned that qualitative insights were not secondary or supplemental; they were essential to understanding how people actually interact with algorithmic systems. Combining both lenses offered a more holistic understanding, capturing not just what people do, but why they do it.

At the center of all three dimensions is the human. Whether we are speaking about disciplines, methodologies, or research paradigms, all paths ultimately converge on human experiences, values, and relationships. The challenge of building civic-centered AI is, at its core, a human challenge. It requires us to engage across disciplines, collaborate with communities, and reflect on our own positions as researchers.

And yet, despite the increasing urgency of these issues, academia remains largely organized around disciplinary silos. Computer science, economics, psychology, and sociology each have their own languages, norms, and reward structures. Interdisciplinary collaboration is often encouraged in principle but under-supported in practice. In our own experience, we've seen how difficult it can be to publish interdisciplinary work or to gain recognition for research that straddles the line between fundamental and applied. But if we are to respond meaningfully to the societal challenges posed by AI, we must break down these silos. Interdisciplinarity has been widely critiqued for staying superficial, lacking proper incentives, or being blocked by disciplinary boundaries [119, 120]. In our case, it was not interdisciplinarity on its own that made our work effective, but rather the transdisciplinary nature of civic engagement. When we grounded our research in the real concerns of communities, through partnerships and ongoing dialogue, disciplinary boundaries became flexible out of necessity. The drive to address human needs brought together technical, behavioral, and ethical perspectives in a way that individual disciplines could not achieve alone. This leads us to believe that community and societal engagement will be the driving force behind the interdisciplinary creation of AI and the breaking of disciplinary silos as we know them today.

In this spirit, we must also reflect on the principle of people and perspective, which asks: who builds AI? This question goes beyond academic discipline and touches on

identity, representation, and power. During our school choice work, when we publicly questioned the limitations of the existing algorithm, we were met with institutional skepticism: not on technical grounds, but based on who we were perceived to be. This moment made clear that credibility in AI is not just about evidence or argument, but about whose voice is valued. Our AI research must actively challenge these hierarchies by uplifting the knowledge and authority of those who have been historically excluded from technological development, especially those embedded in the very communities AI aims to serve.

It is also worth noting that our reflections do not map neatly onto a single stage of the AI lifecycle. Many of the most important insights emerged in the ideation phase, and subsequently carried over into the design, deployment and impact assessment phases.

The three proposed dimensions cut across the typical technical pipeline, emphasizing that centering human needs and civic perspectives can fundamentally reshape whether, how, and why AI systems are pursued in the first place.

Our reflections go beyond the Amsterdam school choice case. They are a call to researchers, policymakers, and communities to embrace a broader vision of AI research that is interdisciplinary, context-aware, and grounded in civic responsibility. We also underline the supreme importance of civic engagement: to invite citizens' involvement not as data points, but as co-designers and co-owners of the innovations. This, combined with more transparency in AI research and policy decisions, will help build trust between academia, businesses, government, and citizens. This can be a very uncomfortable space for researchers; indeed, this requires skills many scientists are not trained for in their careers. But we argue that true growth and progress can only come when we allow ourselves to be uncomfortable in our working space.

When AI research becomes flexible across all these dimensions, centered on people and engaged with society, it will become not just innovative but transformative.

6.5 CONCLUSION

Our experience working on school choice in Amsterdam has been more than a case study in algorithm design—it has been a broader lesson in what it means to build responsible, civic-centered AI. We began with a technical question and ended with a new appreciation for the role of interdisciplinarity, applied relevance, and community engagement in shaping impactful research. Along the way, we profoundly experienced that designing algorithms is never just a technical task—it is also a social, political, and ethical one.

As we reflect on this journey, we recognize that we are not alone in pursuing this vision. Around the world—in the West, the Global South, and Asia, research communities are beginning to grapple with similar questions about how to center civic values in AI. These communities, like ours, are pushing against the boundaries of disciplinary silos and advocating for research that is not only technically rigorous but also socially grounded. Importantly, they recognize that interdisciplinarity alone does not drive social impact; rather, it is the pursuit of social impact that fosters interdisciplinarity and generates new scientific knowledge.

This convergence of thinking signals a paradigm shift in the field. AI is no longer seen solely as a tool to optimize performance or automate decision-making; it is increasingly

understood as a co-actor in human systems—one that must be accountable, inclusive, and aligned with civic values. Building such systems requires collaboration across disciplines and sectors, and genuine partnerships with the communities affected by AI.

Our hope is that this reflection contributes to that growing movement. We invite fellow researchers, practitioners, and institutions to reimagine the goals and methods of AI research. Let us move beyond narrow definitions of impact and success. Let us invest in relationships, learn from those with lived experience, and pursue questions that matter in the world. Only then can AI fulfill its promise—not just as a technology, but as a force for equitable and collective progress.

7

CONCLUSION

This thesis examined strategic behavior in one-sided matching, with a focus on school choice systems where strategyproofness is valued but can come at a substantial cost in allocation quality.

The opening question was a simple one: how can people be assigned to resources on the basis of their preferences when improving the allocation may also create incentives to misreport those preferences? The preceding chapters approached this question from five angles. Chapter 2 studied the strategic risk of welfare-oriented matching and contributed a best-response heuristic for rank-minimizing matching that requires only knowledge of which schools are popular. Chapter 3 examined how institutional policy can reintroduce incentives even when the underlying mechanism is strategyproof, showing how Amsterdam's placement guarantee revokes the strategyproofness of the underlying mechanism. Chapter 4 asked whether real participants act on formal incentives, and developed an SP/A-based behavioral model of parental decision-making together with survey evidence that transparency about strategic risk can act as a deterrent rather than an enabler. Chapter 5 asked whether learning-based methods can occupy new positions in the design space, and evaluated learned matching policies under a suite of misreporting strategies. Chapter 6 drew broader lessons from the Amsterdam case for AI research in public contexts, proposing a conceptual model and operational principles for civic AI research.

The main argument of the thesis is that one-sided matching cannot be treated as an algorithmic problem alone. When matching algorithms are used to allocate public resources, they become part of a sociotechnical system shaped by the algorithm itself, by the behavior that emerges around it, and by the institutional setting in which it is embedded. The conclusions of this chapter are organized around these three components, as visualized in Figure 7.1.

7.1 SUMMARY OF FINDINGS

This section returns to the research questions of Chapter 1. The overarching question of the thesis was:

How should one-sided matching systems be evaluated and designed when attempts to improve allocations incentivize strategic misreporting of preferences?

The thesis-level answer is that when strict strategyproofness becomes too costly, matching systems should be evaluated not only on allocation quality but on the strategies they create, whether those strategies are behaviorally plausible, what happens when they are adopted widely, and how the surrounding institution shapes incentives, communication, and trust. The five specific research questions developed this answer in turn.

The first question concerned the most direct version of the problem: *if a more welfare-oriented matching rule is used in place of a strategyproof one, what strategies*

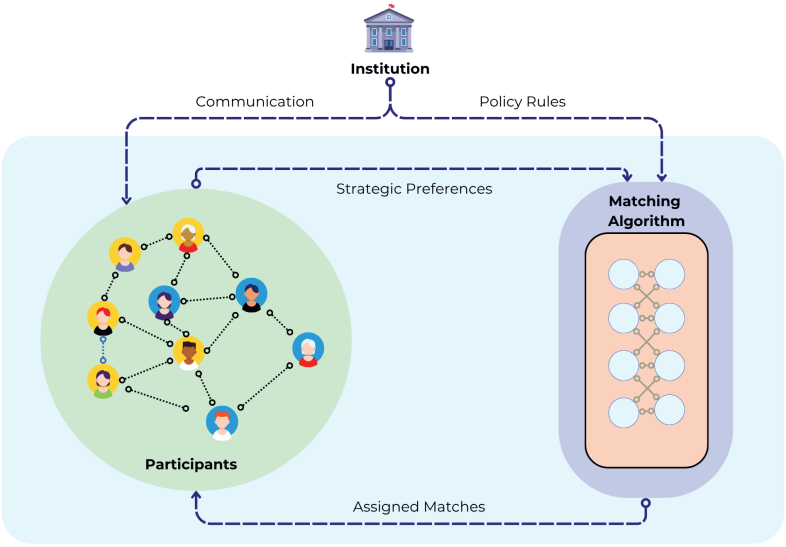


FIGURE 7.1: **One-sided matching as a sociotechnical system.** Participants submit preferences to a matching algorithm, which produces matches. The institution shapes this process through additional policy rules and communication, affecting how participants understand and respond to the system.

become available, how severe is the resulting risk, and to whom? Chapter 2 addressed this by studying Rank Minimization implemented via the Hungarian Algorithm with linear utilities over rank (RM-HAL) as an alternative to RSD. The chapter first derived a best-response strategy under a one-step cost function, then extended it to a heuristic suited to the linear cost case. The heuristic has a simple structure: a participant keeps their top-n preferences truthful and pads the remainder with popular schools, which act as a high-cost "wall" the algorithm avoids when assigning low-ranked positions.

The strategy requires only knowledge of which schools are popular, not full preference information about the other participants. This makes it accessible to ordinary applicants and more practically dangerous than results based on full information might suggest, weakening prior "not obviously manipulable" framings of rank-efficient mechanisms. Although RM-HAL still produces better outcomes than RSD in most high-competition markets even when the heuristic is widely adopted, strategic participants consistently obtain better assignments than truthful ones when both coexist. The welfare gain offered by RM-HAL is therefore real but conditional, and it accrues unequally to those who recognize the strategy.

The second question concerned a less obvious source of strategic behavior: *if the base matching method is strategyproof, can institutional modifications create strategic vulnerabilities?* Chapter 3 addressed this by analyzing the *placement guarantee* policy in Amsterdam school choice, a proportional capacity expansion of 4% designed to ensure that any applicant with a sufficiently long preference list receives a seat at one of their listed schools. The chapter showed that the addition of this guarantee revokes the strategyproofness of the underlying Random Serial Dictatorship (RSD) method.

For applicants with fewer genuinely preferred schools than the required list length, listing popular schools after their sincere preferences weakly dominates filling the remaining positions randomly. The chapter establishes that this strategy is weakly dominant under mild assumptions.

This finding led to a concept the thesis names a *promise gap*: the distance between what an algorithmic safeguard publicly assures and what it can sustain once applicants adapt to it. The placement guarantee illustrates the gap on two levels. At the system level, widely adopted strategic padding inflates the population of guarantee-eligible students beyond the added capacity, rendering the guarantee infeasible. At the distributional level, applicants with longer sincere lists or better information about which schools are popular benefit disproportionately. Institutional behavior supports the analysis: Amsterdam reduced the required list length from twelve to nine in 2025 and restored it to twelve in 2026 after the shorter list was found to compromise the guarantee. The placement guarantee is therefore a clear case of a safeguard that can induce, rather than prevent, the strategic behavior it was meant to forestall.

The first two questions established that strategic incentives can arise from both the matching rule and the policy rules layered around it. The third question shifts focus from formal analysis to behavioral response in practice and asks: *if strategic incentives exist formally, how do real participants respond to them?* Chapter 4 addressed this through a survey of 140 Amsterdam parents during the 2025 application cycle, conducted under three transparency conditions about strategic risk. Choices were modeled using Security/Potential Aspiration (SP/A) theory, which fit substantially better than expected utility and prospect theory baselines. SP/A is therefore a methodological contribution to behavioral matching research: a theoretical framework that captures parental decision-making under matching uncertainty better than the dominant alternatives.

The chapter's main empirical finding reverses a common assumption in the school-choice literature. Transparency about strategic risk acted as a deterrent rather than an enabler. When the costs and failure modes of strategic reporting were made explicit, parents who could have benefited from misreporting became less likely to use it. This effect was clearest for parents in the most competitive academic track (VWO), who were more risk-averse rather than more strategically aggressive. The chapter also identified a bimodal distribution of risk preferences at the population level, indicating that aggregate averages mask two distinct decision-making types. Strategic incentives do not, on their own, predict strategic behavior. What matters is how risk is communicated, how the strategy is understood, and the level of competition in the matching market.

The fourth question turns from analyzing existing mechanisms to widening the design space itself: *can learning-based methods explore alternative tradeoffs between efficiency and strategyproofness?* Chapter 5 formulated one-sided matching as a sequential decision problem and trained Deep Q-Networks (DQN) and Generative Flow Networks (GFlowNet), comparing them to RSD, Probabilistic Serial, and rank-minimizing matching. The chapter introduced four misreporting strategies designed to apply different levels of strategic pressure: the popularity heuristic from Chapter 2, a random adjacent swap, a best adjacent swap, and a best top-h permutation found by local search. These strategies give a multi-dimensional view of robustness rather than a single robustness score, and they vary in behavioral plausibility.

This chapter has three prominent findings. First, robustness to misreporting is multi-dimensional: vulnerability depends on the misreporting strategy, and a method that resists one type of misreport can be exposed to another. GFlowNet, for instance, sat between Probabilistic Serial and RM on efficiency, was robust to the popularity heuristic, but remained vulnerable to optimized local misreports. A single robustness number is not sufficient to capture this. Second, individual incentives and population-level effects come apart: the existence of a profitable misreport for a single agent does not predict what happens when many agents adopt the same misreport simultaneously, which justifies population-adoption experiments as a distinct part of the evaluation pipeline. Learned methods are notably more stable than classical methods under population-level strategic adoption, with stability holding for both strategic adopters and truthful agents. Third, allocation quality under truthful reporting is not a reliable proxy for robustness to misreporting: methods that look similar under truthful reports can differ substantially under strategic stress. Learning-based methods do not solve the efficiency-strategyproofness tradeoff, but they can offer different solutions compared to classical mechanisms, and they must be evaluated under explicit strategic stress.

The four questions above examined the efficiency-strategyproofness tradeoff through formal, institutional, behavioral, and computational lenses. The final question takes a step back to ask *what do these findings imply for responsible AI research for civic purposes?* Chapter 6 addressed this by reflecting on the Amsterdam case and on the broader practice of AI research for public systems. The chapter proposed a conceptual model organized around three dimensions, science and society, fundamental and applied, and quantitative and qualitative, with the human participant at the center. The model is paired with five operational principles for civic AI work, addressing the planning, framing, decomposition, conduct, and participation of such research.

The chapter's central argument concerns the relationship between interdisciplinarity and civic engagement. Interdisciplinarity on its own tends to be additive: methods are stacked, framings are juxtaposed, and integration is left to the reader. Civic engagement forces integration because real institutions and affected communities do not accept disjoint perspectives. Transdisciplinarity, in this sense, is achieved through civic engagement rather than the other way around. The chapter also makes a more uncomfortable reflexive point: institutional credibility in this kind of work is often determined by the perceived identity and disciplinary background of the researcher rather than by the evidence produced. This is an obstacle to applied work that is not adequately recognized in current research practice. Responsible civic AI research therefore requires both intellectual flexibility across the three dimensions and sustained engagement with the public institutions and communities affected by the systems under study.

The remainder of this chapter draws broader conclusions around the three components of the sociotechnical system: the algorithm, the behavior that emerges around it, and the institution in which both are embedded.

7.2 MAIN CONCLUSIONS

The summary above followed the five research questions. The conclusions below reorganize those findings around the three components of the sociotechnical system

in Figure 7.1: the algorithm, the behavior that emerges around it, and the institution in which both are embedded.

C1: Strategic behavior should be part of the design problem

Strategyproofness has obvious appeal in one-sided matching. If truthful reporting is optimal, applicants can state what they want without calculating how the system will respond. This lowers the cognitive burden of participation and makes submitted preferences a more reliable basis for allocation. In public systems such as school choice this matters: families should not need technical knowledge or strategic advice to take part in a public education system.

The thesis has shown that in one-sided matching, strategyproofness can come at a substantial cost in allocation quality, and when that cost is high, a strategyproof mechanism may not be enough for families, policymakers, or institutions to regard the system as satisfactory. In such settings, the presence of strategic incentives should not automatically make a matching method unacceptable. It should instead become a central object of design and evaluation.

This requires a more precise account of strategic behavior than the binary between strategyproof and manipulable allows. A profitable misreport may exist formally without being behaviorally plausible. A policy may appear harmless in theory, yet create a strategy that is easy to understand, easy to communicate, and likely to spread. The behavioral work in this thesis shows that participants do not necessarily act strategically whenever a profitable misreport exists; their responses depend on how risky the strategy appears, how clearly the cost of failure is communicated, and how the system as a whole is understood. So rather than only asking how much efficiency or allocation quality is lost by insisting on strategyproofness, we then have to ask who bears that loss, which strategic opportunities are likely to be used, and whether the resulting risk is acceptable given its behavioral plausibility, distributional impact, and institutional context. When plausible misreporting strategies can be identified, methods can be evaluated against those strategies rather than against the worst case alone. This yields a more realistic account of strategic risk and supports the design of systems robust to the misreports participants are actually likely to use.

C2: Learning-based methods can explore alternative design tradeoffs

Classical one-sided matching mechanisms occupy a small number of familiar points in the tradeoff between allocation quality and strategic risk. Learning-based policies make it possible to search beyond this fixed menu. A policy can be trained, evaluated, and compared by its behavior under truthful reports and under behaviorally plausible misreports.

The results of this thesis show that learning-based methods produce matching policies with efficiency and strategic-risk profiles that classical mechanisms do not match. They do not eliminate the underlying impossibility, nor remove the need for careful evaluation, but they add empirical points of comparison, and they show interesting results when a particular misreporting strategy is adopted at the population level.

This matters because strategic behavior in public systems is rarely an isolated individual misreport. When many participants adopt the same strategy, the system-level effect can differ sharply from the effect of a single misreport. Sequential learned policies can condition later assignments on earlier ones, which gives them a flexibility that standard mechanisms lack in the same form. In the experiments of this thesis, this flexibility allowed GFlowNet to maintain high allocation quality while remaining more robust than rank-minimizing matching against the popularity heuristic, even though it remained vulnerable to optimized local misreports.

Learning-based policies may create incentive structures that are harder for participants to reason about than the structures created by simple deterministic rules. On the other hand, this can increase the opacity of the matching method, which raises serious concerns about transparency and accountability in public systems. However, as also noted in several research works studying the impact of machine-learning methods for public purposes [132–134], transparency in public systems is often less about what is happening inside a black box algorithm, and more about how people view and trust institutions. At the very least, learning-based methods give researchers a way to map a wider range of allocation rules and incentive structures than standard mechanisms can produce. Even when a trained policy is too opaque or too complex to run in practice, it still serves as evidence about the design space: what welfare is attainable under given constraints, where genuine tradeoffs lie, and what structural patterns characterize good mechanisms. These lessons can feed back into theory and into the design of simpler mechanisms that are deployable.

C3: Algorithmic guarantees are conditional on institutional behavior

One-sided matching for public resources cannot be understood through the algorithmic lens alone.

In school choice, the formal matching rule is one part of a larger system: the rule is implemented by institutions, explained to families, modified through policy, and judged through the outcomes it produces. A formally attractive mechanism can fail in practice if it is modified in practice, poorly communicated, if its outcomes are experienced as unfair, or if participants do not trust the institution implementing it. A mechanism with weaker formal guarantees may be more acceptable when its tradeoffs are clearly explained and when affected communities have meaningful opportunities to question them.

This institutional perspective is especially important when strategic behavior is possible. When a mechanism rewards strategic reporting, the burden of strategy is not distributed evenly. Some participants are better placed to recognize the strategic opportunity, act on it, and absorb the risk of failure. Others report truthfully and are disadvantaged by those who do not. The promise gap analyzed in Chapter 3 is a sharper version of the same point: an institutional safeguard can encode a particular promise about outcomes while creating incentives that make the promise impossible to keep, and the resulting failure tends to fall on those least equipped to anticipate it.

The Amsterdam case makes this visible. The problem was never only which algorithm was used, but also how the algorithm was adapted, how families understood the rules, how policymakers responded to dissatisfaction, and how additions such as the placement guarantee changed the strategic environment. This came on top of a school

choice system that already puts a significant burden on parents, especially parents from marginalized backgrounds, who may not be familiar with the procedures. Designing public matching systems therefore requires attention not only to optimization but to communication, legitimacy, institutional context, and the lived experience of participants.

7.3 FUTURE DIRECTIONS

These conclusions point to three sets of future directions, following the same three components: the algorithm, the behavior that emerges around it, and the institution in which both are embedded.

7.3.1 *Evaluating Matching Methods Under Strategic Behavior*

Much of the matching literature treats manipulability as a binary property, and fewer studies measure how much can be gained through plausible misreports, how often such misreports are available, or how outcomes change when strategies are adopted at the population level. Building on RQ1 and RQ4, future work could develop richer best-response analyses that compare unrestricted optima with simpler, behaviorally plausible strategies, and that report both individual gains and system-level effects. Where possible, such analyses should use real preference data rather than synthetic distributions, since the structure of preferences shapes which misreports are profitable.

A second direction is to design or train matching methods explicitly against behaviorally plausible misreports. Rather than treating all possible misreports as equally relevant, future methods could focus on the strategies that participants are most likely to recognize and use, drawing on the behavioral evidence developed for RQ3. This applies both to learning-based methods and to the design of classical mechanisms. It would allow methods to be evaluated against realistic misreports rather than worst-case adversaries. However, this also opens up research questions surrounding which misreports can be considered behaviorally plausible, and this is discussed further in the following section.

A third direction is to evaluate matching methods, learned and classical, against a broader set of criteria. Efficiency and strategic risk are central, but they are not exhaustive. Justified envy, envy-freeness, distributional fairness, priority-respecting constraints, transparency, and robustness to population-level behavioral change are also relevant. This is especially important for learning-based methods, where strong performance on one metric may hide undesirable behavior on another. Strategic behavior should be treated as a design problem, but not as the *only* design problem.

7.3.2 *Understanding Strategic Behavior in Practice*

This thesis shows that participants do not necessarily act strategically whenever misreporting is possible. What remains less clear is when they do, why they do, and which strategies they recognize as available.

Future research should study strategic behavior in matching through behavioral experiments, field studies, and mixed-methods work. Such studies could vary how

mechanisms are explained, the perceived risk of misreporting, the complexity of the strategy, and the stakes of the decision. They could also examine how participants learn from one another, how advice circulates, and how strategic behavior spreads through social networks and school-choice communities. Extending RQ3, this could include how participants receive advice from institutions and peer networks, and how that advice interacts with the formal information made available by the system.

This work matters because mechanism design tends to treat manipulability as a formal possibility. In real systems, behavior depends on interpretation. A strategy that is profitable in theory may not be used if it is too hard to understand. A strategy that gives only a small gain may still spread if it is easy to communicate. Future work should connect formal models of misreporting with empirical models of how participants actually reason about allocation systems and about each other.

7.3.3 *Improving Institutional Design Around Matching Systems*

If matching algorithms are used to allocate public resources, then the surrounding institutional design matters as much as the algorithm itself. Future research should study how algorithmic tradeoffs are communicated to citizens, how policy rules around the mechanism change incentives, and how institutions can reduce the cognitive burden placed on participants.

This involves studying better communication of algorithmic tradeoffs. Resource allocation is a challenging problem, and policymakers often struggle communicating the constraints and tradeoffs they are dealing with. In such cases, institutional communication should go beyond procedural instructions and offer a clear account of what the system can and cannot guarantee. This is especially important when strategyproofness is relaxed or when additional policy rules create new incentives, as in the placement guarantee case studied for RQ2. Clearer communication may reduce unnecessary strategic behavior by making risks legible, and may improve trust by making tradeoffs more explicit, consistent with the deterrent effect found for RQ3. It is also an interesting research direction to investigate how to communicate the procedural details of a matching system alongside its incentives in a way that does not invite further gaming.

A second direction is to create more deliberative processes around public allocation systems. Allocation rules encode difficult tradeoffs, and those tradeoffs are not purely technical: they reflect contested judgements about which goals matter and which burdens are acceptable. Affected communities should have opportunities to contribute to decisions about the priorities of the system, the constraints it should respect, and the tradeoffs it should accept. This can have the added benefit that participants may be more accepting of difficult tradeoffs when they have some agency over the design of the system. This does not imply that every technical design choice should be settled through participation, but that deliberation should serve a specific function: clarifying public values, identifying burdens hidden by formal metrics, and adjudicating tradeoffs where formal theory cannot.

A third direction is to build stronger bridges between academic research and the institutions that implement matching systems. One risk in applied algorithmic research is that formal results are interpreted too narrowly or implemented without sufficient attention to context. Conversely, institutional adaptations may alter the properties of

a mechanism in ways that are not fully understood by policymakers, as the placement guarantee illustrates. Closer collaboration between researchers, policymakers, schools, and affected communities can reduce this gap and support systems that lower, rather than raise, the strategic and cognitive burden placed on citizens.

7.4 FINAL REFLECTION

I began this thesis treating one-sided matching as an algorithmic problem and quickly came to recognize it as a sociotechnical one. The system I had set out to study turned out to consist of algorithms, participants, institutions, and the relationships among them, and no single discipline could give a satisfactory account of how it works in practice. What seemed at first like a clean technical question, whether a more efficient matching method could replace a strategyproof one, opened onto a series of less tractable ones about institutional adaptation, behavioral response, and the conditions under which formal guarantees actually survive contact with the world.

The core of this thesis will always be rooted in Amsterdam school choice. It was inspired by the city and remains relevant to the city's current matching method, as described in Chapter 3. This specificity does not, however, diminish the transferability of the findings. If anything, the opposite holds: applied research of this kind is crucial for the progress of fundamental research, because it surfaces the boundaries of theoretical guarantees and forces the field to formulate more precise questions and more realistic evaluation criteria. This, however, requires flexibility: between theory and application, between quantitative and qualitative methods, and between academic expertise and civic engagement.

The current structure and incentives of a computer science doctorate do not always accommodate such flexibility. I have felt this throughout the work, and I suspect many others have too. The pressure to publish leaves little room for the slower work of civic engagement: building a two-way relationship with citizens and stakeholders, beyond the one-time interaction of collecting data and presenting findings. Nor does the field reward the hours spent translating academic results for policymakers, or working with journalists to bring an issue into public view. These are activities that take time, do not produce papers, and yet are exactly what civic-facing computer science needs to do more of.

I suspect, however, that the recent sweeping entry of AI into social domains and individual lives acts as a crucible for the field of computer science itself. We are going through a transformation where our methods and techniques now need a nuanced sociotechnical lens, with no clear solutions in sight. This may be the moment to push for a version of the field that takes its public responsibilities seriously: that values engagement with affected communities, that recognizes communication as part of the work, and that makes room in its doctoral training for people who want to do research differently.

It is a challenging future for research, but one I believe worth the trouble.

A

APPENDIX: CHAPTER 2

A.1 RM USING HUNGARIAN ALGORITHM

This work implements the rank minimization matching mechanism using the Hungarian algorithm. The Hungarian algorithm has been chosen in consideration of transparency in real-world matching scenarios for the following reasons:

A.1.1 *Simplicity*

Although not as straight-forward as RSD, the Hungarian algorithm consists of only a few deterministic steps. The theory behind the algorithm should be understandable to any experts with a background in linear algebra and graph theory. This simplicity is valued by policymakers, who may be obliged to explain the procedure to parents. In addition, from a scientific perspective, simple procedures are valued in Mechanism Design because all agents must understand the payoffs for every action at every step.

A.1.2 *Reproducibility*

The matching solution is reproducible. The requirements for this are: a fixed, initial ordering of agents and items, for example sorted alphabetically.

In comparison, mixed-integer programming algorithms do not necessarily give a reproducible solution. The solution found can depend on the computing platform, specifically: the chip instruction set, operating system and compiler used.

A.1.3 *Availability*

Many open-source versions of the Hungarian algorithm are available. Here the `scipy` implementation [135] is used. We thus trust the full algorithm to be auditable, and as such accountability of any matching can be provided to the general public.

A.2 EXAMPLE APPLICATION OF BEST RESPONSE STRATEGY

The application of the best-response strategy is illustrated in Figure A.1, in an example with 5 agents matched to items A, B and C, with capacities $|A| = 1$, $|B| = 1$ and $|C| = 3$. All agents apply the best response. Each preference list has a length of 2. The sink item K, appended to each preference list, has a capacity of 2. Two fake agents (S_6, S_7) are added to the cost matrix to make it square. Items A and B are popular, and are the first choices of agents S_1, S_2 and S_3, S_4 respectively, with popular items B and A as second choice. Item C is not popular, and is the first choice of agent S_5, with item A as the second choice. This results in the cost matrix of Figure A.1a.

Cost matrix 2 in Figure A.1b illustrates the vertical and horizontal lines of step 1 of the proof in Section 2.4.1. Popular items A and B are covered with vertical lines. The unmarked items in the residual cost matrix are C and K, which have spare capacity and C has the largest penalty. In cost matrix 3 of Figure A.1c, illustrating steps 2 and 3 of the proof, the cost reduction step has been applied to the matrix, which makes the costs corresponding to the sink item zero. All zeros are now covered by seven lines (equal to the total number of rows or columns) and a match is made.

Cost matrix 4 in Figure A.1d shows the final match of agents to the items. All actual agents are allocated to either their first choice or the sink item.

A.3 ACCEPTANCE OF TOP-n MATCHES

In a survey performed by OSVO [14] the percentages of parents and students who find being matched to the top-n school choices acceptable are measured for different values of n. The values for parents and students, shown in Table A.1, are averaged and used in the experiments.

Rank	**Finds top-n Acceptable (%)**		
	Parents	**Students**	**Average**
Top 1	100	100	100
Top 2	83	80	81.5
Top 3	63	55	59
Top 4	39	33	36
Top 5	30	24	27
Top 6	14	12	13
Top 7	9	7	8
Top 8	7	5	6
Top 9	5	4	4.5
Top 10	5	4	4.5

TABLE A.1: Percentages of parents and students in Amsterdam who find allocation to a certain rank acceptable.

A.4 AMSTERDAM DATASET

The data from Amsterdam used in the experiments is not public due to privacy concerns and can be obtained from OSVO upon request. The dataset consists of two tables described in Table A.2.

Attribute	Description
Advies	Type of secondary education students recei ve.
Lotnummer	Lottery number assigned to student
Geplaatst op	Name of school student was placed at by allocation algorithm.
Positie	Position at which student ranked their allocated school
Voorrang/Hardheid eerste voorkeur	Denotes if student has priority at any school. These students were excluded from experiments.
Voorkeur 1	Name of student's 1st preferred school
...	...
Voorkeur 34	Name of student's 34th preferred school

(a) Students

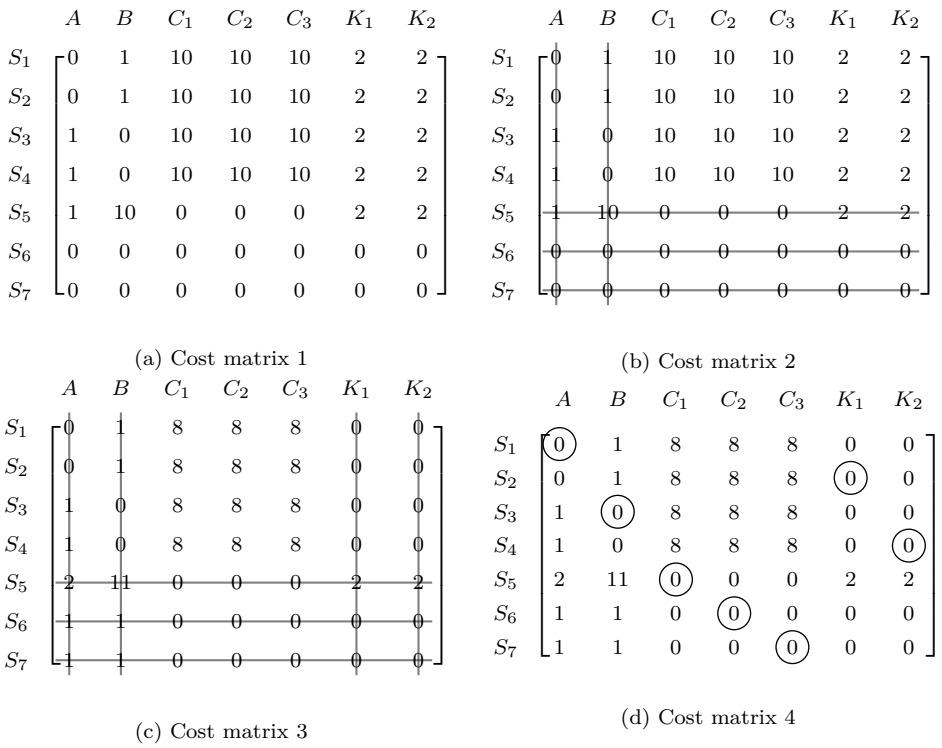
Attribute	Description
Key	Unique key combining school name and education type.
VO-school	Name of School
Capaciteitsgroep	Education type offered by school
Maximale capaciteit definitieve matching	Number of seats available at school for this education type

(b) Schools

TABLE A.2: Metadata from Amsterdam School Allocation Dataset.

A.5 ADDITIONAL RESULTS

Some additional results are presented which extend the results shown in Section 2.6.3. Matching performance is tested for half of all agents reporting truthful preferences, and the rest applying the first-order heuristic strategy for $n \in [1, 3]$. The results in Figure A.2 show that for most markets strategic agents (orange lines) are provided better matches than truthful ones (blue lines). These results confirm the findings of Section 2.6.3, that regardless of demand, agents have the incentive to be strategic in their reported preferences.



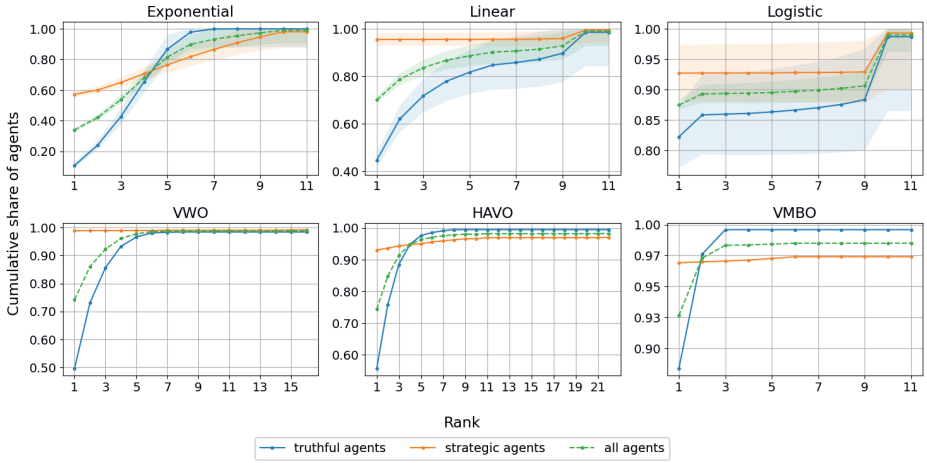
(a) Cost matrix 1

(b) Cost matrix 2

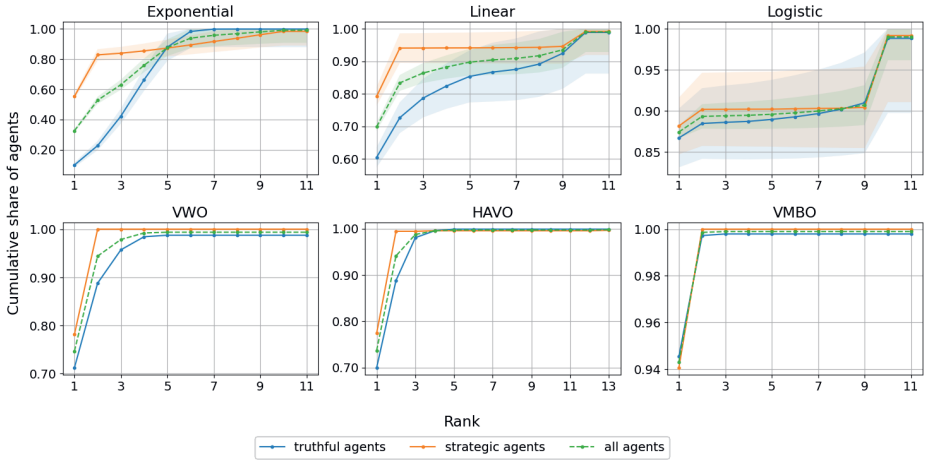
(c) Cost matrix 3

(d) Cost matrix 4

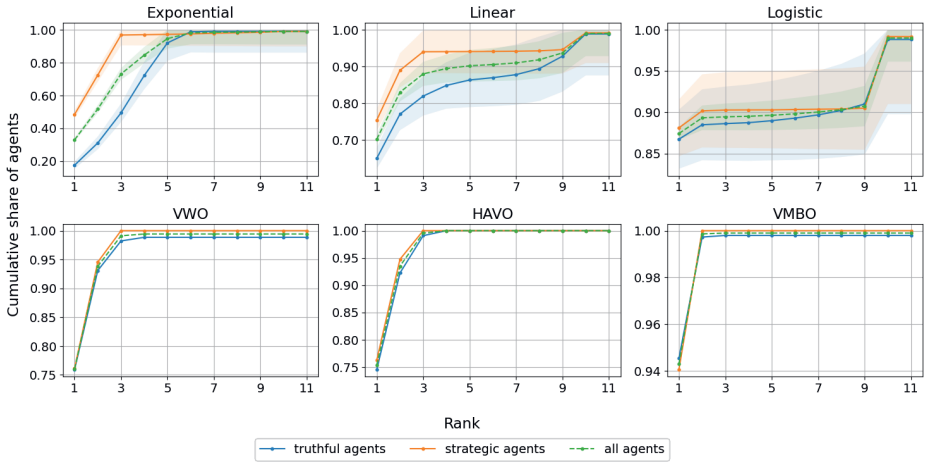
FIGURE A.1: Illustration of the best-response strategy of the Theorem, applied to an example cost matrix.



(a) $n = 1$



(b) $n = 2$



(c) $n = 3$

FIGURE A.2: Cumulative share of matched agents over ranks in RM-HAL for truthful agents vs. strategic agents applying the first-order heuristic strategy for different values of n.

B

APPENDIX: CHAPTER 3

B.1 RSD WITH PLACEMENT GUARANTEE ALGORITHM

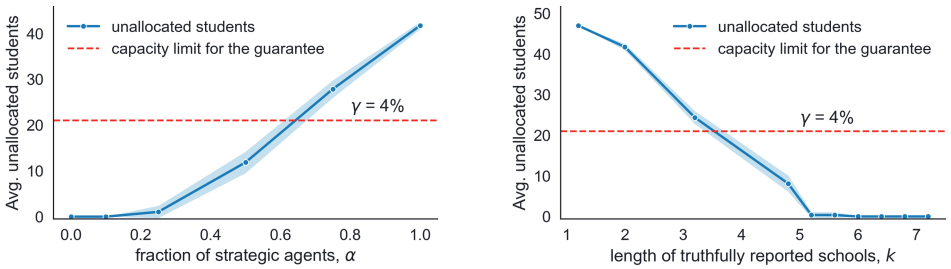
Algorithm 4 gives a step-by-step description of the allocation procedure used in the model and simulations. The procedure first runs Random Serial Dictatorship (RSD) [1] using the submitted lists and the original school capacities. Students who cannot be assigned to any school on their submitted list in this main round remain unassigned. Among these students, only those who submitted a list of length at least l are eligible for the placement guarantee. The guarantee stage is then modelled as an ex-post capacity expansion, where each school s receives $\lfloor \gamma c_s \rfloor$ additional seats. If the number of guarantee-eligible students exceeds the total number of additional seats, the guarantee cannot be satisfied for all eligible students under this aggregate capacity benchmark.

Algorithm 4: RSD with placement guarantee

Require: Students N, schools S, capacities c_s, reported lists L_i, guarantee threshold l, expansion rate γ
Draw a random permutation π of students
Initialise remaining capacities $q_s \leftarrow c_s$ for all $s \in S$
for student i in order π **do**
    Assign i to the highest-ranked school in L_i with $q_s > 0$, if one exists
    If assigned to school s, set $q_s \leftarrow q_s - 1$; otherwise leave i unassigned
end for
Let $U = \{i : i \text{ is unassigned and } |L_i| \geq l\}$
Expand capacities by $\tilde{c}_s = c_s + \lfloor \gamma c_s \rfloor$ for all $s \in S$
Assign students in U using $\tilde{c}_s$ in order of π
If $|U| > \sum_{s \in S} \lfloor \gamma c_s \rfloor$, the guarantee cannot be satisfied

B.2 ADDITIONAL EXPERIMENTAL RESULTS

This appendix provides additional experimental material that supports the main results in Section 5.7. First, we report the synthetic versions of the feasibility and inequality experiments. These experiments preserve the same qualitative patterns as the Amsterdam-based results in the main text, while providing a more controlled environment in which the mechanism can be isolated. Second, we report tables with the exact Amsterdam-based means and 95% confidence intervals underlying the main-chapter figures.



(a) Varying α

(b) Varying k

FIGURE B.1: **Feasibility of the placement guarantee under strategic adaptation in the synthetic setting.** The figure shows the average number of guarantee-eligible students after the main RSD round as the fraction of strategic agents, α, and the number of truthfully reported schools, k, vary. The dashed red line marks the total additional capacity C^{extra} created by the placement guarantee. Error bars show 95% confidence intervals over 10,000 simulation runs.

B.2.1 *Synthetic Feasibility Results*

Figure B.1 reports the synthetic feasibility experiments. As in the Amsterdam-based setting, the average number of guarantee-eligible students increases as the fraction of strategic agents, α, increases, and decreases as the number of sincerely reportable schools, k, increases. The same qualitative mechanism therefore appears in both environments: strategic padding becomes more effective when strategic reporting is more widespread and when students have fewer sincere preferences, and under these conditions the number of guarantee-eligible students can exceed the additional capacity created by the placement guarantee.

B.2.2 *Synthetic Inequality Results*

Figure B.2 reports the synthetic inequality experiments. As in the Amsterdam-based setting, unequal strategic capacity produces unequal outcomes. When one group can rank more schools sincerely, that group obtains more favourable final outcomes. Likewise, when one group has more accurate information about which schools are useful for strategic padding, that group benefits more from the mechanism. The synthetic results therefore reinforce the interpretation of the Amsterdam-based results as reflecting a general incentive pattern induced by the placement guarantee rather than an artefact of one particular dataset.

B.2.3 *Amsterdam-Based Feasibility Results*

Table B.3 reports the exact Amsterdam-based means and 95% confidence intervals corresponding to Figure 3.3 in the main text. The final column reports the excess over capacity, computed as the mean number of guarantee-eligible students minus C^{extra}.

B.2.4 *Amsterdam-Based Inequality Results*

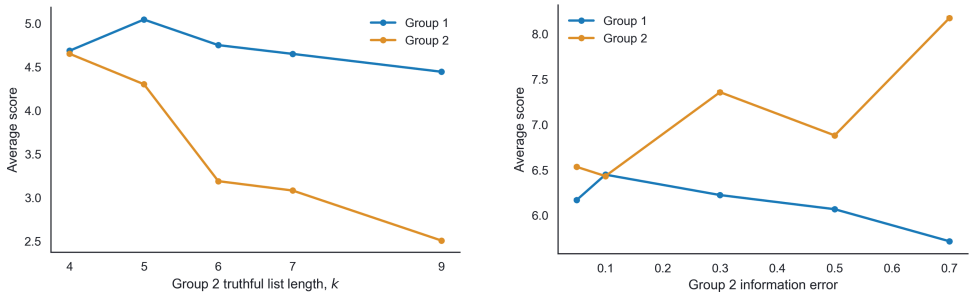
Tables B.1 and B.2 report the exact Amsterdam-based means and 95% confidence intervals corresponding to Figure 3.4 in the main text. In both tables, lower average score indicates a more favourable final outcome.

Group 1 k	Group 2 k	Group 1 score	Group 2 score
4	4	1.892 ± 0.003	1.910 ± 0.003
4	5	1.847 ± 0.003	1.730 ± 0.003
4	6	1.846 ± 0.003	1.587 ± 0.003
4	7	1.837 ± 0.003	1.485 ± 0.003
4	9	1.731 ± 0.003	1.401 ± 0.002

TABLE B.1: Amsterdam-based inequality results for different truthful-list lengths across groups. Values are reported as mean score ± 95% CI.

Group 1 P_{error}	Group 2 P_{error}	Group 1 score	Group 2 score
0.05	0.05	2.097 ± 0.003	1.757 ± 0.003
0.05	0.10	2.000 ± 0.003	1.900 ± 0.003
0.05	0.30	1.655 ± 0.003	2.123 ± 0.003
0.05	0.50	1.486 ± 0.003	2.162 ± 0.003
0.05	0.70	1.486 ± 0.003	2.177 ± 0.003

TABLE B.2: Amsterdam-based inequality results for different levels of popularity-information error across groups. Values are reported as mean score ± 95% CI.



(a) Different truthful-list lengths (k) across groups

(b) Different popularity-information accuracy across groups

FIGURE B.2: **Unequal strategic capacity experiments in the synthetic setting.** Panel (a) shows average score by group as Group 2's truthful-list length, k, varies while Group 1 is held fixed. Panel (b) shows average score by group as Group 2's error in identifying popular schools varies while Group 1 is held fixed. In both panels, the two groups are of equal size and all agents are strategic. Lower average score indicates a more favourable final outcome. Error bars show 95% confidence intervals over 10,000 simulation runs.

Experiment	Setting	Guarantee-eligible students	C^{extra}	Excess over capacity
Varying k	$k = 1$	212.012 ± 0.003	55	157.012
Varying k	$k = 3$	114.518 ± 0.159	55	59.518
Varying k	$k = 4$	85.828 ± 0.169	55	30.828
Varying k	$k = 7$	33.139 ± 0.165	55	-21.861
Varying α	$\alpha = 0.00$	0.319 ± 0.023	55	-54.681
Varying α	$\alpha = 0.10$	1.777 ± 0.059	55	-53.223
Varying α	$\alpha = 0.25$	7.114 ± 0.101	55	-47.886
Varying α	$\alpha = 0.50$	32.571 ± 0.167	55	-22.429
Varying α	$\alpha = 0.75$	71.918 ± 0.181	55	16.918
Varying α	$\alpha = 1.00$	114.503 ± 0.160	55	59.503

TABLE B.3: Amsterdam-based feasibility results. Values are reported as mean number of guarantee-eligible students ± 95% CI.

C

APPENDIX: CHAPTER 4

C.1 SURVEY QUESTIONNAIRE

C.1.1 *Scenario 1: Choice Without Probability Information*

Participants rank three schools in order of preference. A school allocation algorithm determines placement, but no probability information is provided.

Assume the following are the participant's top three choices:

- School A (First choice)
- School B (Second choice)
- School C (Third choice)

Some parents believe that including popular schools (e.g., School X and School Y) may improve placement chances, even if they are not preferred choices. The following lists represent possible preference submissions:

1. List 1:
   - 1st Choice - School A
   - 2nd Choice - School B
   - 3rd Choice - School C
2. List 2:
   - 1st Choice - School A
   - 2nd Choice - School B
   - 3rd Choice - School X
3. List 3:
   - 1st Choice - School A
   - 2nd Choice - School X
   - 3rd Choice - School Y

Question: Without knowing admission probabilities, which list would you submit?

- [ ] List 1
- [ ] List 2
- [ ] List 3

Optional: Briefly explain why you chose this list.

C.1.2 *Scenario 2: Choice With Probability Information*

Now, admission probabilities for each list are provided:

	1st Choice	2nd Choice	3rd Choice	Not Placed
List 1	Low (20%)	High (70%)	Low (10%)	0%
List 2	Low (22%)	High (73%)	0%	Low (5%)
List 3	Low (25%)	0%	0%	High (75%)

Question: Knowing these probabilities, which list would you submit?

- [ ] List 1
- [ ] List 2
- [ ] List 3

Optional: Briefly explain why you chose this list.

C.1.3 *Scenario 3: Adjusting Choices After a Probability Shift*

The probabilities for List 3 have changed, increasing the likelihood of placement in the first-choice school.

	1st Choice	2nd Choice	3rd Choice	Not Placed
List 1	Low (20%)	High (70%)	Low (10%)	0%
List 2	Low (22%)	High (73%)	0%	Low (5%)
List 3	High (90%)	0%	0%	Low (10%)

Question: With these updated probabilities, which list would you submit now?

- [ ] List 1
- [ ] List 2
- [ ] List 3

Optional: Briefly explain why you chose this list.

C.1.4 *Background Information*

To help analyze school choice behavior, respondents are asked to provide demographic information.
1. Has your child gone through or will they go through the school matching process?

- [ ] Yes
- [ ] No
- [ ] Not Applicable

2. What is your child's school advice level?
   - [ ] VWO
   - [ ] VWO/HAVO
   - [ ] HAVO
   - [ ] VMBO
   - [ ] Practical Education
3. What is your highest level of education?
   - [ ] Primary School
   - [ ] High School
   - [ ] Bachelor's Degree
   - [ ] Master's Degree
   - [ ] PhD
4. What is your household's yearly income?
   - [ ] Less than 25,000
   - [ ] 25,000 - 50,000
   - [ ] 50,000 - 100,000
   - [ ] Not Applicable
5. How important do you think school quality is for your child's future success?
   - [ ] Extremely Important
   - [ ] Important
   - [ ] Somewhat Important
   - [ ] Not Very Important
   - [ ] Neutral/No Opinion
6. How important is it to you that your child attends a top-ranked school?
   - [ ] Extremely Important
   - [ ] Important
   - [ ] Somewhat Important
   - [ ] Not Very Important
   - [ ] Neutral/No Opinion

Scenario	Strategy	Probability of being matched to Rank 1	Probability of being matched to Rank 2	Probability of being matched to Rank 3	Probability of being not placed	Potential	Security
Scenario 1	Truthful	Unknown	Unknown	Unknown	Unknown	0.5	0.5
Scenario 1	Mild	Unknown	Unknown	Unknown	Unknown	0.5	0.5
Scenario 1	Aggressive	Unknown	Unknown	Unknown	Unknown	0.5	0.5
Scenario 2	Truthful	0.2	0.7	0.1	0	0.2	1
Scenario 2	Mild	0.22	0.73	0	0.05	0.22	0.95
Scenario 2	Aggressive	0.25	0	0	0.75	0.25	0.25
Scenario 3	Truthful	0.2	0.7	0.1	0	0.2	1
Scenario 3	Mild	0.22	0.73	0	0.05	0.22	0.95
Scenario 3	Aggressive	0.9	0	0	0.1	0.9	0.9

TABLE C.1: Potential and security values for each strategy across the three experimental scenarios. Potential reflects the admission probability for the first-ranked school on each list. Security reflects the probability of being matched to any school on the list (i.e., not remaining unplaced). Values were directly derived from the admission probabilities shown to participants.

C.2 SP/A MODELLING DETAILS

We estimate two variants of the SP/A model:

- **Linear covariate model:** α_i is modeled as a logistic function of respondent characteristics (education, income, school track, aspiration variables).
- **Baseline model:** A constant α is estimated across all participants, assuming no heterogeneity.

Parameters are estimated by maximum likelihood using the BFGS optimizer implemented in SciPy with default convergence settings. Missing covariates were mean-imputed after standardization. We compute parameter uncertainty via the inverse Hessian of the log-likelihood, using the 'numdifftools' library. Due to limited sample size, we report in-sample model comparisons using AIC, BIC, and McFadden's pseudo-R^2. No cross-validation was performed.

C.3 BASELINE MODELS

To benchmark the SP/A model, we estimate two widely used decision-making models under uncertainty:

- **Expected Utility (EU):** Assumes each parent selects a strategy in proportion to its linear expected placement probability, scaled by a single sensitivity parameter.
- **Prospect Theory (PT):** Implements a reduced form of cumulative prospect theory with a single probability weighting parameter γ, using the one-parameter Prelec function:

$$w(p) = \exp\left(-(-\ln p)^{\gamma}\right).$$

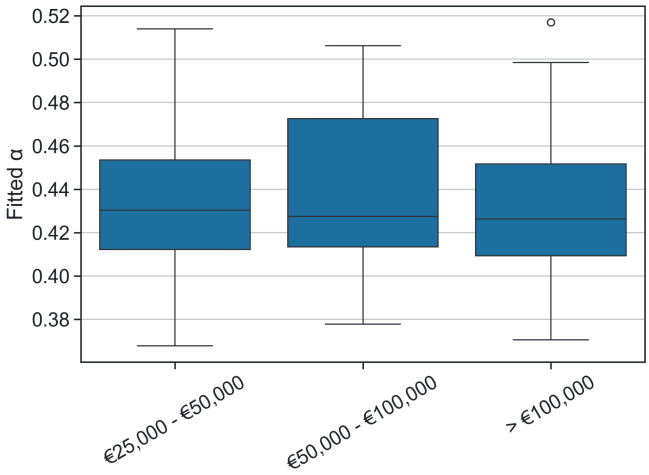


FIGURE C.1: Distribution of fitted risk preference parameters α_i across income level of respondents.

This variant does not model value curvature or loss aversion and is included as a comparison for probability distortion effects only.

Model comparisons demonstrate that the SP/A model, even in its simplest form, provides a substantially better fit to observed behavior than either EU or PT. The linear covariate version of the SP/A model captures some meaningful heterogeneity, though individual covariate coefficients were not statistically significant. Additional robustness checks involving clustering and covariate exclusion are discussed in the results.

C.4 ESTIMATING SECURITY AND POTENTIAL

In each scenario of the survey, participants were shown admission probabilities for each of the three strategy lists: a truthful ranking, a mildly strategic ranking, and an aggressive ranking. Based on these probabilities, we defined:

- **Potential** as the probability of being admitted to the top-ranked school in a given list.
- **Security** as the probability of being matched to any school in the list, i.e., 1 − probability of being unplaced.

The values used in our SP/A model were computed directly from these scenario-specific probabilities and are summarized in Table C.1.

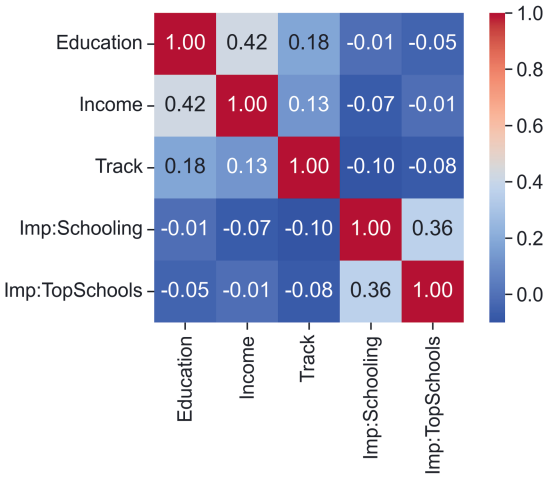


FIGURE C.2: Correlation matrix of standardized parent-level covariates. Education and income show moderate correlation ($r = 0.42$), while other relationships are weak. This supports the use of all covariates in the risk preference model without major concerns about multicollinearity.

C.5 CORRELATION AMONG COVARIATES

Figure C.2 presents the pairwise Pearson correlation coefficients among the standardized covariates used to model heterogeneity in risk preferences: parental education, income, child's school track, and two attitudinal variables (importance of schooling and aspiration for top schools).

The highest observed correlation is between education and income ($r = 0.42$), which is moderate and expected in socioeconomic data. Other correlations are relatively low (all $r < 0.37$), suggesting that multicollinearity is limited and unlikely to bias coefficient estimates substantially. All variables were standardized prior to modeling.

D

APPENDIX: CHAPTER 5

D.1 APPENDIX A: GFLOWNET IMPLEMENTATION

The GFlowNet background, sampling property, and forward-looking detailed balance objective are described in Section 5.3.3. The matching MDP, the terminal score $\mathcal{S}(\mu)$, and the GNN parameterization are described in Section 5.4.2. This appendix provides the implementation details not covered in the main text.

D.1.1 *Reward Shaping*

The forward-looking detailed balance objective requires a reward signal $R(s)$ defined at every state, not just at terminal states. We use an intermediate reward equal to the partial sum of utility weights of the edges that have been selected up to state s:

$$\mathcal{S}(s) = \sum_{(a,i) \text{ selected in } s} w_{a,i}. \tag{D.1}$$

At the initial state s_0 no edges have been selected, so $\mathcal{S}(s_0) = 0$. At a terminal state s_T corresponding to complete matching μ, all selected edges contribute and $\mathcal{S}(s_T) = \mathcal{S}(\mu)$ as defined in Equation 5.7. At intermediate states, $\mathcal{S}(s)$ is the running total of utility weights collected so far.

The GFlowNet reward at any state is derived from $\mathcal{S}(s)$ via an exponential transform controlled by a scaling parameter β:

$$\log R(s) = \beta \cdot \mathcal{S}(s), \tag{D.2}$$

equivalently $R(s) = \exp(\beta \cdot \mathcal{S}(s))$. This is the quantity that enters the FL-DB loss in Equation 5.5. Larger values of β concentrate sampling probability on near-optimal matchings, while smaller values give more uniform sampling. In the main experiments we use $\beta = 500$, held fixed throughout training. The intermediate reward signal makes use of the forward-looking detailed balance objective: at each transition, the credit assignment incorporates the partial score difference $\mathcal{S}(s') - \mathcal{S}(s)$, which equals the utility weight of the selected edge.

D.1.2 *Architecture*

The GFlowNet policy uses a 3-layer Graph Isomorphism Network (GIN) with sum aggregation, learnable ϵ, and batch normalization between layers. Hidden dimension is 128 throughout. Agent and item nodes are distinguished through learnable type embeddings. Each edge (a, i) has a 1-dimensional input feature, the current state entry at (a, i), which is mapped to dimension 128 by a linear edge-feature transform. After three rounds of message passing, edge scores are produced by a 2-layer MLP

that takes the concatenation of the source node embedding, target node embedding, and transformed edge feature (total input dimension 3×128) and outputs a scalar. Scores at infeasible edges are masked to $-\infty$, and the remaining scores are normalized into a categorical distribution by softmax. The flow function $F(s)$ used in the FL-DB objective is computed by a separate GIN with the same architecture as the policy network, with a graph-level scalar output obtained by max-pooling edge scores within each graph.

D.1.3 *Training*

We train the GFlowNet using the forward-looking detailed balance objective defined in Equation 5.5. Training is on-policy: at each iteration, the current policy is rolled out on a batch of profiles to generate transitions, the loss is computed over these transitions, and the buffer is reset before the next iteration. Models are trained on profiles sampled from the uniform preference distribution.

Hyperparameter	Value
Optimizer	Adam
Learning rate (policy)	10^{-3}
Learning rate (flow)	10^{-3}
Training profiles per epoch	10,000
Epochs	1
Rollout batch size	8 profiles
Training steps per rollout	30
Training batch size	64 transitions
Exploration rate p_{rand}	0.3
Reward scaling β	500 (fixed)
GIN layers	3
Hidden dimension	128
Aggregator	sum

TABLE D.1: GFlowNet training hyperparameters.

D.1.4 *Monte Carlo Sampling at Evaluation*

At evaluation time, we sample complete matchings by rolling out the trained forward policy from the initial state. At each state, the model scores feasible edges, masks infeasible actions, and samples the next edge from the resulting categorical distribution. For each evaluation profile, we draw 50 Monte Carlo samples and average outcomes across them.

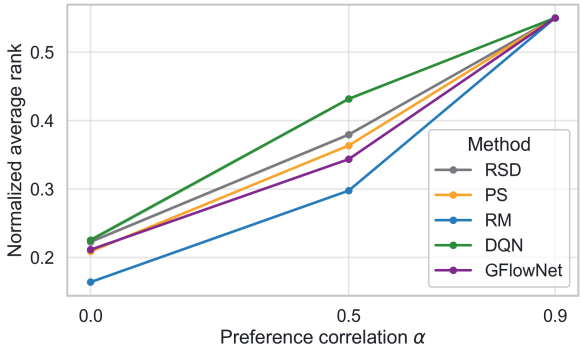


FIGURE D.1: **Truthful efficiency under correlated preference distributions.** Lower normalized average rank indicates better assignments.

D.2 APPENDIX B: DQN IMPLEMENTATION DETAILS

The DQN setup, including the matching MDP, the intermediate reward signal, and the GNN parameterization, is described in Section 5.4.3. This appendix provides implementation details not covered in the main text.

D.2.1 *Architecture*

The Q-network uses a 2-layer GraphConv backbone with hidden dimension 64. Agent and item nodes are first mapped to the hidden space by a linear node encoder using one-hot node-type features. Two GraphConv layers with ReLU activations refine the embeddings via message passing. For each edge (a, i), the source node embedding, target node embedding, original utility weight $w_{a,i}$, and current state entry at (a, i) are concatenated and passed through a 2-layer MLP to produce a scalar Q-value. The total input dimension to the MLP is therefore $2 \times 64 + 1 + 1 = 130$.

D.2.2 *Training and Inference*

The DQN is trained with experience replay and a periodically updated target network. We use Huber (smooth-ℓ_1) loss to stabilize learning under heavy-tailed Q-value errors. Gradients are clipped at norm 5.0 before each optimizer step. Models are trained on profiles sampled from the uniform preference distribution. At evaluation time, the agent selects the feasible edge with the highest predicted Q-value at each step. The resulting matching is deterministic given the trained Q-network.

D.3 ROBUSTNESS UNDER CORRELATED PREFERENCES

The main results in Section 5.7 are reported for the uniform preference setting. To test whether the main patterns are specific to uniformly random preferences, we

Hyperparameter	Value
Optimizer	Adam
Learning rate	3×10^{-4}
Weight decay	10^{-5}
Batch size	16
Replay buffer size	20,000 transitions
Learning starts after	2,000 transitions
Discount factor γ	0.95
Exploration	ϵ-greedy, decayed from 1.0 to 0.05
Target update interval	250 training steps
Gradient clipping	max norm 5.0
Loss	Huber (smooth-ℓ_1)
Episodes	3,000
Rollout batch size	8 profiles
Train every	50 environment steps
GraphConv layers	2
Hidden dimension	64

TABLE D.2: DQN training hyperparameters.

also evaluate the methods under correlated-demand preference profiles. We vary the preference-correlation parameter

$$\alpha \in \{0, 0.5, 0.9\},$$

where larger values indicate stronger agreement among agents about which items are desirable. These experiments use the same $n = 10$ setting, dataset seeds, methods, and misreporting strategies as the main individual-evaluation experiments.

Figure D.1 reports truthful efficiency as preference correlation increases. Normalized average rank generally worsens as preferences become more correlated, since agents increasingly compete for the same highly ranked items.

Figures D.2 and D.3 report the corresponding misreporting strategies across preference-correlation levels. The popularity heuristic becomes less effective as correlation increases: under high correlation, agents already agree on which items are popular, so reordering the report by popularity provides little additional leverage. On the other hand, the effect of the local strategies increases sharply with correlation: as agents compete for the same items, even small local deviations can produce substantial gains. PS generally remains unaffected by misreporting at all correlation levels, consistent with its bounded incentive ratio [93].

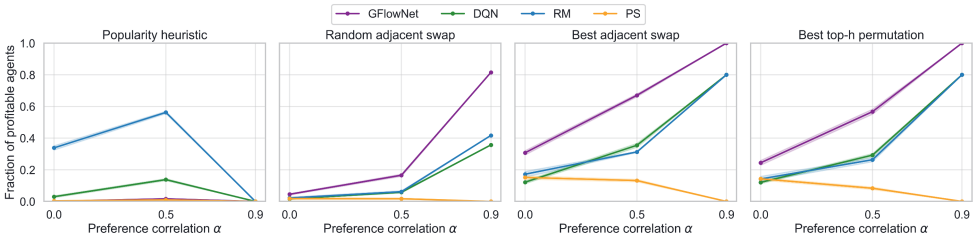


FIGURE D.2: **Fraction of agents with a profitable misreport under correlated preferences.** Each panel corresponds to one misreporting strategy.

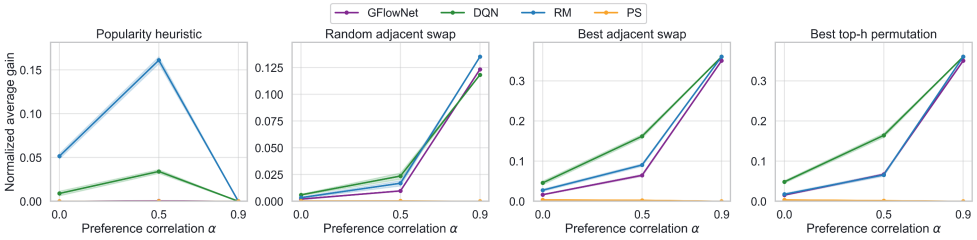


FIGURE D.3: **Normalized average misreporting gain under correlated preferences.** Each panel corresponds to one misreporting strategy.

D.4 POPULATION-ADOPTION RESULTS

Table D.3 reports the numerical values corresponding to the population-level adoption experiment in Figure 5.5. A fraction ρ of agents adopts the popularity heuristic while the remaining agents report truthfully. We report normalized average rank for the full population, for agents who adopt the heuristic, and for agents who remain truthful, with lower values indicating better assignments. The strategic-agent column is undefined at $\rho = 0$ (no adopters); the truthful-agent column is undefined at $\rho = 1$ (no truthful agents).

D.4.1 *Additional Population-Adoption Experiments*

The main population-level experiment in Section 5.7.3 focuses on adoption of the popularity heuristic. We report population-adoption experiments for the other three misreporting strategies in Figures D.4, D.5, and D.6.

The population-level effects of the three alternative strategies are markedly smaller than those of the popularity heuristic. Under all three strategies, the overall and strategic-agent outcomes are essentially flat across ρ for all methods. The only visible movement is in the truthful-agent panels under the best adjacent swap and best top-h permutation, where DQN drifts upward at high adoption rates; under random adjacent swap, all methods remain flat.

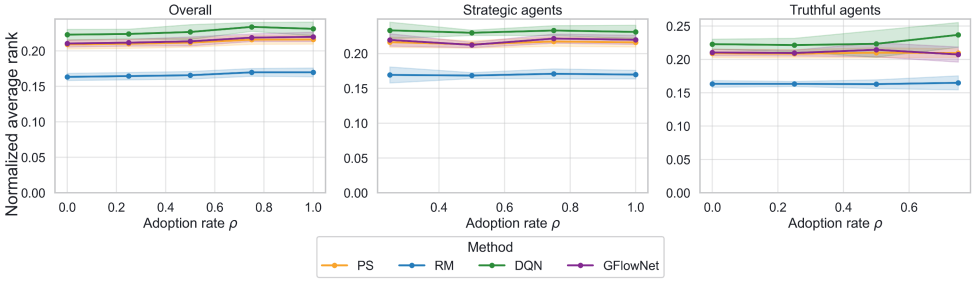


FIGURE D.4: **Population adoption of the random adjacent-swap perturbation.**

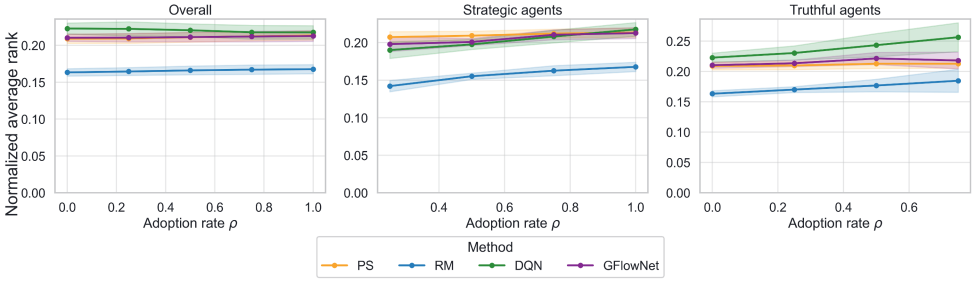


FIGURE D.5: **Population adoption of the best adjacent swap.**

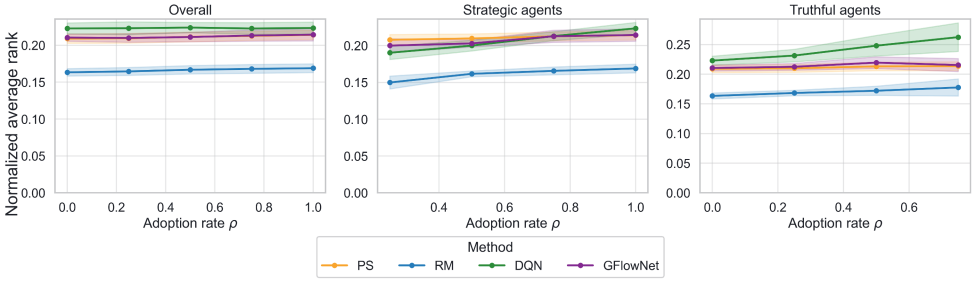


FIGURE D.6: **Population adoption of the best top-h permutation strategy.**

ρ	Method	Overall (↓)	Strategic agents (↓)	Truthful agents (↓)
0.00	DQN	0.2229 ± 0.0033	n/a	0.2229 ± 0.0033
0.00	GFlowNet	0.2104 ± 0.0022	n/a	0.2104 ± 0.0022
0.00	PS	0.2091 ± 0.0028	n/a	0.2091 ± 0.0028
0.00	RM	$\mathbf{0.1634 \pm 0.0022}$	n/a	$\mathbf{0.1634 \pm 0.0022}$
0.00	RSD	0.2218 ± 0.0024	n/a	0.2218 ± 0.0024
0.25	DQN	0.2317 ± 0.0042	0.2841 ± 0.0112	0.2187 ± 0.0047
0.25	GFlowNet	0.2155 ± 0.0025	0.2736 ± 0.0091	0.2009 ± 0.0014
0.25	PS	0.2377 ± 0.0034	0.3274 ± 0.0073	0.2153 ± 0.0025
0.25	RM	$\mathbf{0.1811 \pm 0.0030}$	$\mathbf{0.1553 \pm 0.0060}$	$\mathbf{0.1875 \pm 0.0032}$
0.25	RSD	0.2516 ± 0.0027	0.3457 ± 0.0073	0.2281 ± 0.0020
0.50	DQN	0.2416 ± 0.0059	0.2671 ± 0.0110	0.2162 ± 0.0039
0.50	GFlowNet	0.2282 ± 0.0031	0.2618 ± 0.0052	$\mathbf{0.1945 \pm 0.0024}$
0.50	PS	0.2820 ± 0.0032	0.3359 ± 0.0044	0.2281 ± 0.0024
0.50	RM	$\mathbf{0.2205 \pm 0.0034}$	$\mathbf{0.1886 \pm 0.0045}$	0.2525 ± 0.0040
0.50	RSD	0.2986 ± 0.0023	0.3572 ± 0.0040	0.2401 ± 0.0011
0.75	DQN	0.2614 ± 0.0084	0.2735 ± 0.0116	0.2130 ± 0.0054
0.75	GFlowNet	0.2535 ± 0.0056	0.2722 ± 0.0068	$\mathbf{0.1787 \pm 0.0021}$
0.75	PS	0.3293 ± 0.0054	0.3523 ± 0.0063	0.2370 ± 0.0027
0.75	RM	$\mathbf{0.2491 \pm 0.0067}$	$\mathbf{0.2362 \pm 0.0078}$	0.3005 ± 0.0053
0.75	RSD	0.3500 ± 0.0034	0.3753 ± 0.0035	0.2487 ± 0.0035
1.00	DQN	0.2732 ± 0.0117	0.2732 ± 0.0117	n/a
1.00	GFlowNet	$\mathbf{0.2718 \pm 0.0071}$	$\mathbf{0.2718 \pm 0.0071}$	n/a
1.00	PS	0.3579 ± 0.0059	0.3579 ± 0.0059	n/a
1.00	RM	$\mathbf{0.2672 \pm 0.0073}$	$\mathbf{0.2672 \pm 0.0073}$	n/a
1.00	RSD	0.3813 ± 0.0035	0.3813 ± 0.0035	n/a

TABLE D.3: Population-level adoption of the popularity heuristic. The table reports normalized average rank for the overall population, strategic adopters, and truthful agents. Lower values indicate better assignments; the best value per (ρ, group) cell is shown in bold. The strategic-agent column is not applicable at $\rho = 0$, and the truthful-agent column is not applicable at $\rho = 1$.

SUMMARY

One-sided matching algorithms decide who gets access to scarce public resources, from secondary school seats to social housing to refugee placements. The standard design problem is framed as a trade-off between two desirable properties: matching efficiency, which is whether people are assigned to items they actually prefer, and strategyproofness, which is whether they have any reason to misrepresent their preferences. The classical result is that no mechanism can have both. Policymakers must therefore choose, and either choice comes with a cost: strategyproof mechanisms produce inefficient matches with long tails of unfavorable assignments, while more efficient mechanisms create opportunities for strategic manipulation that may be distributed unequally across participants. This thesis investigates how one-sided matching systems should be evaluated and designed when attempts to improve allocations create new incentives for strategic misreporting of preferences.

In *Chapter 2*, we investigate strategic reporting in the rank minimization mechanism, a welfare-oriented alternative to Random Serial Dictatorship that achieves better rank distributions but is not strategyproof. We derive a best-response strategy for participants seeking to be matched to their top preferred items and extend it to a first-order heuristic strategy that requires only knowledge of which items are popular as first choices. Using both synthetic markets and real preference data from secondary school choice in Amsterdam, we show that the strategy is easy to apply and creates a clear individual incentive to misreport, exposing a distributional gap between those who can act on it and those who cannot.

In *Chapter 3*, we expand the analysis from manipulable mechanisms to strategyproof mechanisms embedded in broader institutional settings. We study the Amsterdam placement guarantee, a policy modification layered on top of Random Serial Dictatorship that promises placement to students who submit a sufficiently long preference list. We characterize a simple reporting strategy for students with fewer sincere preferences than the required list length, and show through simulation that as this behavior spreads, the number of guarantee-eligible students can exceed the additional capacity the policy creates, rendering the guarantee infeasible. We use this case to articulate a *promise gap* in public algorithmic systems: a mismatch between what an implemented system publicly assures and what it can sustain once participants adapt to it.

With *Chapter 4*, the focus shifts from formal incentives to behavioral response. We conduct a behavioral experiment embedded in a survey of 140 parents participating in Amsterdam's 2025 school choice process, and model their decisions using Security-Potential/Aspiration theory. We find that transparency about strategic risk acts as a deterrent rather than an enabler of manipulation, and that parents in the more competitive academic tracks are more risk-averse rather than more strategic. These findings reverse common assumptions in the school choice literature and suggest that the trade-off between efficiency and manipulability may be less sharp in practice than formal analyses imply.

While the previous chapters take the design space of matching mechanisms as given, *Chapter 5* asks whether learning-based methods can reach intermediate positions

that classical mechanisms cannot. We formulate one-sided matching as a sequential decision problem on a bipartite graph and train two learned policies, one based on Generative Flow Networks and one on Deep Q-Networks. Because learned policies admit no analytical strategyproofness guarantees, we develop a framework for empirical robustness evaluation built around four misreporting strategies and a population-level adoption experiment. We find that truthful efficiency is not a reliable indicator of robustness, that individual incentive to misreport does not predict population-level outcomes, and that learned policies remain stable under population-level strategic adoption where classical methods deteriorate.

Finally, *Chapter 6* reflects on what this body of work implies for AI research in civic contexts. We propose a framework for civic AI research organized around three dimensions: between science and society, between fundamental and applied work, and between quantitative and qualitative methods, with the human at the intersection of all three. Drawing on the Amsterdam school choice case, we argue that civic engagement is not a downstream concern but the practical condition that makes interdisciplinary AI research possible in the first place.

This thesis brings formal, behavioral, and computational perspectives to one-sided matching. Our contributions show that algorithmic allocation systems shape outcomes, incentives, administrative burdens, and institutional trust at the same time, and that responsible evaluation of such systems cannot stop at the formal properties of the underlying algorithm but must also examine how participants respond to it and how the design space itself can be widened.

NEDERLANDSE SAMENVATTING

Eenzijdige matchingalgoritmen bepalen wie toegang krijgt tot schaarse publieke voorzieningen, van plaatsen op middelbare scholen tot sociale woningbouw en de plaatsing van vluchtelingen. Het standaardontwerpprobleem wordt geformuleerd als een afweging tussen twee wenselijke eigenschappen: matching-efficiëntie, dat wil zeggen of mensen worden toegewezen aan opties die zij daadwerkelijk prefereren, en strategiebestendigheid, dat wil zeggen of zij enige reden hebben om hun voorkeuren onjuist op te geven. Het klassieke resultaat is dat geen enkel mechanisme aan beide eigenschappen tegelijk kan voldoen. Beleidsmakers moeten daarom kiezen, en elke keuze brengt kosten met zich mee: strategiebestendige mechanismen leiden tot inefficiënte toewijzingen met een lange staart aan ongunstige uitkomsten, terwijl efficiëntere mechanismen ruimte bieden voor strategische manipulatie die mogelijk ongelijk verdeeld is over de deelnemers. Dit proefschrift onderzoekt hoe eenzijdige matchingsystemen geëvalueerd en ontworpen zouden moeten worden wanneer pogingen om toewijzingen te verbeteren nieuwe prikkels creëren tot het strategisch misrapporteren van voorkeuren.

In *Hoofdstuk 2* onderzoeken we strategisch rapporteren binnen het rangminimaliseringsmechanisme, een welvaartgeoriënteerd alternatief voor Random Serial Dictatorship dat betere rangverdelingen oplevert maar niet strategiebestendig is. We leiden een beste-respons strategie af voor deelnemers die geplaatst willen worden bij hun meest geprefereerde opties, en breiden deze uit tot een eerste-orde heuristische strategie die enkel kennis vereist van welke opties populair zijn als eerste keuze. Aan de hand van zowel synthetische markten als echte voorkeursgegevens uit de Amsterdamse middelbare schoolkeuze, laten we zien dat de strategie eenvoudig toepasbaar is en een duidelijke individuele prikkel creëert om voorkeuren onjuist op te geven, waarmee een verdelingskloof zichtbaar wordt tussen wie hiernaar kan handelen en wie niet.

In *Hoofdstuk 3* verbreden we de analyse van manipuleerbare mechanismen naar strategiebestendige mechanismen die ingebed zijn in een bredere institutionele context. We bestuderen de Amsterdamse plaatsingsgarantie, een beleidswijziging bovenop Random Serial Dictatorship die plaatsing belooft aan leerlingen die een voldoende lange voorkeurslijst indienen. We karakteriseren een eenvoudige rapportagestrategie voor leerlingen met minder oprechte voorkeuren dan de vereiste lijstlengte, en tonen via simulatie aan dat naarmate dit gedrag zich verspreidt, het aantal voor de garantie in aanmerking komende leerlingen de extra capaciteit kan overschrijden die het beleid creëert, waardoor de garantie onhaalbaar wordt. We gebruiken deze casus om een *beloftekloof* (*promise gap*) in publieke algoritmische systemen te benoemen: een mismatch tussen wat een geïmplementeerd systeem publiekelijk belooft en wat het kan waarmaken zodra deelnemers zich eraan aanpassen.

Met *Hoofdstuk 4* verschuift de focus van formele prikkels naar gedragsmatige respons. We voeren een gedragsexperiment uit, ingebed in een enquête onder 140 ouders die deelnamen aan het Amsterdamse schoolkeuzeproces van 2025, en modelleren hun beslissingen aan de hand van de Security-Potential/Aspiration-theorie. We vinden dat transparantie over strategisch risico eerder als afschrikking dan als aanmoediging van manipulatie werkt, en dat ouders waarvan de kinderen op de meer competitieve

schoolniveaus zitten juist risico-averser zijn dan strategischer. Deze bevindingen weerspreken gangbare aannames in de schoolkeuzeliteratuur en suggereren dat de afweging tussen efficiëntie en manipuleerbaarheid in de praktijk minder scherp kan zijn dan formele analyses doen vermoeden.

Terwijl de voorgaande hoofdstukken de ontwerpruimte van matchingmechanismen als gegeven beschouwen, stelt *Hoofdstuk 5* de vraag of leergebaseerde methoden tussenliggende posities kunnen bereiken die klassieke mechanismen niet kunnen. We formuleren eenzijdige matching als een sequentieel beslissingsprobleem op een bipartiete graaf en trainen twee geleerde policies, een gebaseerd op Generative Flow Networks en een op Deep Q-Networks. Omdat geleerde policies geen analytische strategiebestendigheidsgaranties bieden, ontwikkelen we een raamwerk voor empirische robuustheidsevaluatie, opgebouwd rondom vier misrapportagestrategieën en een experiment naar populatiebrede adoptie. We vinden dat efficiëntie onder oprechte rapportage geen betrouwbare indicator is van robuustheid, dat de individuele prikkel tot misrapporteren geen voorspeller is van populatiebrede uitkomsten, en dat geleerde policies stabiel blijven onder populatiebrede strategische adoptie waar klassieke methoden achteruitgaan.

Tot slot reflecteert *Hoofdstuk 6* op wat dit geheel aan werk impliceert voor AI-onderzoek in civiele contexten. We stellen een raamwerk voor civiel AI-onderzoek voor, georganiseerd langs drie dimensies: tussen wetenschap en samenleving, tussen fundamenteel en toegepast onderzoek, en tussen kwantitatieve en kwalitatieve methoden, met de mens op het snijvlak van alle drie. Gebaseerd op de Amsterdamse schoolkeuzecasus, stellen we dat maatschappelijke betrokkenheid geen secundaire zorg is, maar juist de praktische voorwaarde die interdisciplinair AI-onderzoek mogelijk maakt.

Dit proefschrift brengt formele, gedragsmatige en computationele perspectieven samen rondom het probleem van eenzijdige matching. Onze bijdragen tonen aan dat algoritmische toewijzingssystemen tegelijkertijd uitkomsten, prikkels, administratieve lasten en institutioneel vertrouwen vormgeven, en dat een verantwoorde evaluatie van zulke systemen niet kan blijven steken bij de formele eigenschappen van het onderliggende algoritme, maar ook moet onderzoeken hoe deelnemers erop reageren en hoe de ontwerpruimte zelf verbreed kan worden.

BIBLIOGRAPHY

1. A. Abdulkadiroğlu & T. Sönmez. *Random serial dictatorship and the core from random endowments in house allocation problems*. Econometrica **66**, 689 (1998).
2. L. Zhou. *On a conjecture by Gale about one-sided matching problems*. Journal of Economic Theory **52**, 123 (1990).
3. A. Bogomolnaia & H. Moulin. *A new solution to the random assignment problem*. Journal of Economic theory **100**, 295 (2001).
4. A. Abdulkadiroğlu, P. Pathak, A. Roth & T. Sönmez. *Changing the Boston school choice mechanism*. 2006.
5. A. Abdulkadiroğlu, P. Pathak & A. Roth. *Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match*. American Economic Review **99**, 1954 (2009).
6. A. Abdulkadiroğlu, Y.-K. Che, P. A. Pathak, A. E. Roth & O. Tercieux. *Minimizing justified envy in school choice: the design of New Orleans' OneApp*. Tech. rep. (National Bureau of Economic Research, 2017).
7. J. Ortega & T. Klein. *The cost of strategy-proofness in school choice*. Games and Economic Behavior (2023).
8. OSVO. *Bijlage II: Procesomschrijving Centrale Loting en Matching Amsterdam*. Kernprocedure Overstap PO-VO Amsterdam 2026–2027. Describes the Amsterdam central lottery and matching procedure as developed according to DA-STB and Random Serial Dictatorship, and states that the LMS application uses the Random Serial Dictatorship calculation method. 2026.
9. Rechtbank Amsterdam. *Ouders uitgelote leerlingen krijgen geen extra informatie*. `https://www.rechtspraak.nl/organisatie-en-contact/organisatie/rechtbanken/rechtbank-amsterdam/nieuws/ouders-uitgelote-leerlingen-krijgen-geen-extra-informatie`. Persbericht. 2017.
10. P. Gautier, M. De Haan, B. Van Der Klaauw & H. Oosterbeek. *Eerste analyse matching en loting voortgezet onderwijs Amsterdam 2016*. Evaluation Paper, URL: http://www. verenigingosvo. nl/wp-content/uploads/2013/09/Rapport-loting-en-matching-2016-beschrijvend-rapport. pdf, Retrieved June **14**, 2016 (2016).
11. M. De Haan, P. Gautier, H. Oosterbeek & B. Van der Klaauw. *The performance of school assignment mechanisms in practice*. Available at SSRN 2619231 (2015).
12. N. Ruijs & H. Oosterbeek. *School choice in Amsterdam: Which schools are chosen when school choice is free?* Education Finance and Policy **14**, 1 (2019).
13. *De vereniging van schoolbesturen in het Amsterdamse voortgezet onderwijs*. `https://www.verenigingosvo.nl/`. Accessed: 2022-08-18. 2022.
14. L. Hermanussen, T. Groot Beumer & A. Grond. *Evaluatie van het plaatsingssysteem* (2019).

15. M. Couzy. *Ouders uitgelote leerlingen stappen naar rechter — parool.nl*. https://www.parool.nl/nieuws/ouders-uitgelote-leerlingen-stappen-naar-rechter~ba569a68/. 2019.

16. Vereniging OSVO. *Wijziging Centrale Loting & Matching 2025-2026*. https://verenigingosvo.nl/2024/10/24/wijziging-centrale-loting-matching-2025-2026/. Vereniging van schoolbesturen in het Amsterdamse voortgezet onderwijs. Published 2024-10-24. Accessed 2026-05-17. 2024.

17. Vereniging OSVO. *Jaarplan 2026*. https://verenigingosvo.nl/wp-content/uploads/2025/12/Jaarplan-2026.pdf. Accessed: 2026-05-22. 2025.

18. Onderwijs Consumenten Organisatie. *OCO Amsterdam*. https://ocoamsterdam.nl/. Accessed: 2026-06-02. 2026.

19. M. Tasnim, Y. Weesie, S. Ghebreab & M. Baak. *Strategic manipulation of preferences in the rank minimization mechanism*. Autonomous Agents and Multi-Agent Systems **38**, 44 (2024).

20. M. Tasnim, J. Trevorrow, S. Ghebreab & N. Ajmeri. *The Promise Gap in Algorithmic School Choice: How Placement Guarantees Fail under Strategic Adaptation*. Under Submission (2026).

21. M. Tasnim, P. Verhagen, T. Blanke, E. Acar & S. Ghebreab. *Modeling Strategic Risk in School Choice: A Case for Transparent Design*. in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* **8** (2025), 2470.

22. M. Tasnim, S. Ghebreab & E. Acar. *Sequential Learning for One-Sided Matching under Preference Misreporting*. Under Submission (2026).

23. M. Tasnim & S. Ghebreab. *AI, interdisciplinarity and civic engagement: insights from Amsterdam school choice*. Cambridge Forum on AI: Culture and Society **1**, e5 (2025).

24. M. Tasnim, Y. Weesie, S. Ghebreab & M. Baak. *Strategic manipulation of preferences in the rank minimization mechanism*. Autonomous Agents and Multi-Agent Systems **38**, 44 (2024).

25. P. Biró. *Applications of matching models under preferences* (2017).

26. F. Bloch & D. Cantala. *Dynamic assignment of objects to queuing agents*. American Economic Journal: Microeconomics **9**, 88 (2017).

27. D. Delacrétaz, S. D. Kominers & A. Teytelboym. *Refugee resettlement*. University of Oxford Department of Economics Working Paper (2016).

28. A. Roth. *Deferred acceptance algorithms: History, theory, practice, and open questions*. international Journal of game Theory **36**, 537 (2008).

29. C. R. Featherstone. *Rank efficiency: Modeling a common policymaker objective*. Unpublished paper, The Wharton School, University of Pennsylvania.[25, 28, 45] (2020).

30. A. Abdulkadiroğlu, P. Pathak, A. Roth & T. Sönmez. *The Boston public school match*. American Economic Review **95**, 368 (2005).

31. S. Aksoy, A. Azzam, C. Coppersmith, J. Glass, G. Karaali, X. Zhao & X. Zhu. *School choice as a one-sided matching problem: Cardinal utilities and optimization*. arXiv preprint arXiv:1304.7413 (2013).

32. P. Troyan. *Non-obvious manipulability of the rank-minimizing mechanism*. arXiv preprint arXiv:2206.11359 (2022).
33. D. E. Knuth. *An exact analysis of stable allocation*. Journal of Algorithms **20**, 431 (1996).
34. A. Nikzad. *Rank-optimal assignments in uniform markets*. Theoretical Economics **17**, 25 (2022).
35. L.-G. Svensson. *Strategy-proof allocation of indivisible goods*. Social Choice and Welfare **16**, 557 (1999).
36. P. Troyan & T. Morrill. *Obvious manipulations*. Journal of Economic Theory **185**, 104970 (2020).
37. O. Lev, R. Meir, S. Obraztsova & M. Polukarov. *Heuristic voting as ordinal dominance strategies*. in *Proceedings of the AAAI Conference on Artificial Intelligence* **33** (2019), 2077.
38. J. Scheuerman, J. L. Harman, N. Mattei & K. B. Venable. *Heuristic strategies in uncertain approval voting environments*. arXiv preprint arXiv:1912.00011 (2019).
39. P. Pathak. *The mechanism design approach to student assignment*. Annu. Rev. Econ. **3**, 513 (2011).
40. D. Gale & L. Shapley. *College admissions and the stability of marriage*. The American Mathematical Monthly **69**, 9 (1962).
41. A. Roth & M. Sotomayor. *Two-sided matching*. Handbook of game theory with economic applications **1**, 485 (1992).
42. P. Pathak. *What really matters in designing school choice mechanisms*. Advances in Economics and Econometrics **1**, 176 (2017).
43. P. Pathak & T. Sönmez. *School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation*. American Economic Review **103**, 80 (2013).
44. F. Kojima & M. Manea. *Incentives in the probabilistic serial mechanism*. Journal of Economic Theory **145**, 106 (2010).
45. P. Pathak & T. Sönmez. *Leveling the playing field: Sincere and sophisticated players in the Boston mechanism*. American Economic Review **98**, 1636 (2008).
46. J. Pais & Á. Pintér. *School choice and information: An experimental study on matching mechanisms*. Games and Economic Behavior **64**, 303 (2008).
47. C. Calsamiglia, G. Haeringer & F. Klijn. *Constrained school choice: An experimental study*. American Economic Review **100**, 1860 (2010).
48. R. Irving, T. Kavitha, K. Mehlhorn, D. Michail & K. Paluch. *Rank-maximal matchings*. ACM Transactions on Algorithms (TALG) **2**, 602 (2006).
49. K. Paluch. *Capacitated rank-maximal matchings*. in *Algorithms and Complexity: 8th International Conference, CIAC 2013, Barcelona, Spain, May 22-24, 2013. Proceedings 8* (2013), 324.
50. J. Munkres. *Algorithms for the assignment and transportation problem*. Journal of the Society for Industrial and Applied Mathematics **5** (1957).

51. R. Burkard, M. Dell'Amico & S. Martello. *Assignment Problems* (2009).

52. R. Jonker & T. Volgenant. *Improving the Hungarian assignment algorithm.* Operations Research Letters **5**, 171 (1986).

53. M. Bazaraa, J. Jarvis & H. Sherali. *Linear Programming and Network Flows 4th ed.* (2010).

54. K. van Beek, E. Bootsma, J. Buma, I. Macdonald, D. Colen & J. Smit. *Stichting Vrije Schoolkeuze Amsterdam.* 2022.

55. A. Abdulkadiroğlu & T. Sönmez. *School choice: A mechanism design approach.* American economic review **93**, 729 (2003).

56. A. Erdil & H. Ergin. *What's the matter with tie-breaking? Improving efficiency in school choice.* American Economic Review **98**, 669 (2008).

57. Y.-K. Che, J. Grenet & Y. He. *Allocating students to schools: theory, methods, and empirical insights.* Handbook of the Economics of Matching **2**, 307 (2025).

58. N. Agarwal & P. Somaini. *Revealed preference analysis of school choice models.* Annual Review of Economics **12**, 471 (2020).

59. Vereniging OSVO. *Procesomschrijving Centrale Loting en Matching Amsterdam.* `https://verenigingosvo.nl/wp-content/uploads/2025/12/Bijlage-2-Procesomschrijving-Centrale-Loting-en-Matching-OSVO.pdf`. Accessed: 2026-05-22. 2026.

60. T. Riehm & D. van de Scheur. *Unintended consequences: Navigating the complexity of school allocation in Amsterdam.* 2023.

61. Vereniging OSVO. *Uitkomst Centrale Loting & Matching 2026.* `https://verenigingosvo.nl/wp-content/uploads/2026/04/Persbericht-Centrale-Loting-Matching-2026.pdf`. Accessed: 2026-05-22. 2026.

62. G. Fack, J. Grenet & Y. He. *Beyond truth-telling: Preference estimation with centralized school choice and college admissions.* American Economic Review **109**, 1486 (2019).

63. A. J. Kapor, C. A. Neilson & S. D. Zimmerman. *Heterogeneous beliefs and school choice mechanisms.* American Economic Review **110**, 1274 (2020).

64. D. A. Díaz, A. M. Jiménez & C. Larroulet. *An agent-based model of school choice with information asymmetries.* Journal of Simulation **15**, 130 (2021).

65. J. S. Hastings, R. Van Weelden & J. M. Weinstein. *Preferences, information, and parental choice behavior in public school choice.* 2007.

66. J. S. Hastings & J. M. Weinstein. *Information, school choice, and academic achievement: Evidence from two experiments.* The Quarterly journal of economics **123**, 1373 (2008).

67. R. Peeters. *The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making.* Information Polity **25**, 507 (2020).

68. J. Zerilli, U. Bhatt & A. Weller. *How transparency modulates trust in artificial intelligence.* Patterns **3** (2022).

69. The Nobel Prize. *The Prize in Economic Sciences 2012 – Press Release*. https://www.nobelprize.org/prizes/economic-sciences/2012/press-release/. Nobel Prize Outreach AB. Accessed 2026-05-17. 2012.

70. M. Moorman. *Column Marjolein Moorman: Van ideeënstrijd naar de beste oplossing*. https://amsterdam.pvda.nl/nieuws/column-marjolein-moorman-van-ideeenstrijd-naar-de-beste-oplossing/. PvdA Amsterdam column. Accessed 2026-05-17. 2017.

71. M. Tasnim, P. Verhagen, T. Blanke, E. Acar & S. Ghebreab. *Modeling Strategic Risk in School Choice: A Case for Transparent Design*. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society **8**, 2470 (2025).

72. Y. He. *Gaming the Boston School Choice Mechanism in Beijing*. Toulouse School of Economics Working Paper. 2012.

73. R. Khaddari. *Weer minder achtstegroepers kunnen naar eerste voorkeursschool: 'Volgend jaar wordt het beter'*. 2022.

74. M. Hulshof. *Opinie: 'Ineens mogen kinderen niet weten hoe groot hun toelatingskans op een scholen is – dat is onmisbare informatie'*. 2025.

75. European Parliament & European Council. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*. OJ (2023).

76. European Parliament & European Council. *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance)*. OJ (2022).

77. Q. Wang, Y. Huang, S. Jasin & P. V. Singh. *Algorithmic transparency with strategic users*. Management Science **69**, 2297 (2023).

78. B. Ustun, A. Spangher & Y. Liu. *Actionable Recourse in Linear Classification*. in *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery, Atlanta, GA, USA, 2019), 10.

79. S. Bronfman, N. Alon, A. Hassidim & A. Romm. *Redesigning the israeli medical internship match*. ACM Transactions on Economics and Computation (TEAC) **6**, 1 (2018).

80. A. Rees-Jones. *Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match*. Games and Economic Behavior **108**, 317 (2018).

81. L. L. Lopes & G. C. Oden. *The role of aspiration level in risky choice: A comparison of cumulative prospect theory and SP/A theory*. Journal of mathematical psychology **43**, 286 (1999).

82. D. Kahneman, A. Tversky, *et al. Prospect theory: An analysis of decision under risk*. Econometrica **47**, 363 (1979).

83. J. Von Neumann & O. Morgenstern. *Theory of games and economic behavior, 2nd rev* (Princeton university press, 1947).

84. J. M. Keynes. *A treatise on probability* (Courier Corporation, 2013).

85. Q. McNemar. *Note on the sampling error of the difference between correlated proportions or percentages*. Psychometrika **12**, 153 (1947).

86. A. Roth. *The economics of matching: Stability and incentives*. Mathematics of operations research **7**, 617 (1982).

87. E. Bengio, M. Jain, M. Korablyov, D. Precup & Y. Bengio. *Flow network based generative models for non-iterative diverse candidate generation*. Advances in Neural Information Processing Systems **34**, 27381 (2021).

88. V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al. Human-level control through deep reinforcement learning*. nature **518**, 529 (2015).

89. N. Chen, X. Deng & J. Zhang. *How profitable are strategic behaviors in a market?* in *Algorithms–ESA 2011: 19th Annual European Symposium, Saarbrücken, Germany, September 5-9, 2011. Proceedings 19* (2011), 106.

90. T. Mennle, M. Weiss, B. Philipp & S. Seuken. *The Power of Local Manipulation Strategies in Assignment Mechanisms.* in *IJCAI* (2015), 82.

91. R. Shirakawa. *Simple manipulations in school choice mechanisms*. American Economic Journal: Microeconomics **17**, 202 (2025).

92. A. Nesterov, O. Rospuskova & S. Rubtcova. *Robustness to manipulations in school choice*. Social Choice and Welfare **62**, 519 (2024).

93. Z. Wang, Z. Wei & J. Zhang. *Bounded incentives in manipulating the probabilistic serial rule*. in *Proceedings of the AAAI Conference on Artificial Intelligence* **34** (2020), 2276.

94. S. S. Ravindranath, Z. Feng, S. Li, J. Ma, S. D. Kominers & D. C. Parkes. *Deep learning for two-sided matching*. arXiv preprint arXiv:2107.03427 (2021).

95. R. Maruo, K. Takeuchi & H. Kashima. *Learning Neural Strategy-Proof Matching Mechanism from Examples*. Frontiers in Artificial Intelligence and Applications (eds Lynce, I., Murano, N., Vallati, M., Villata, S., Chesani, F., Milano, M., Omicini, A. & Dastani, M.) (2025).

96. K. Taywade, J. Goldsmith & B. Harrison. *Multi-agent Reinforcement Learning for Decentralized Stable Matching*. in *Algorithmic Decision Theory* (eds Fotakis, D. & Ríos Insua, D.) (Springer International Publishing, Cham, 2021), 375.

97. Y. Min, T. Wang, R. Xu, Z. Wang, M. Jordan & Z. Yang. *Learn to match with no regret: Reinforcement learning in markov matching markets*. Advances in Neural Information Processing Systems **35**, 19956 (2022).

98. Z. Liu, M. Lu, Z. Wang, M. Jordan & Z. Yang. *Welfare maximization in competitive equilibrium: Reinforcement learning for markov exchange economy*. in *International Conference on Machine Learning* (2022), 13870.

99. M. A. Alomrani, R. Moravej & E. B. Khalil. *Deep policies for online bipartite matching: A reinforcement learning approach*. arXiv preprint arXiv:2109.10380 (2021).

100. P. S. Bagga & A. Delarue. *Solving the quadratic assignment problem using deep reinforcement learning*. arXiv preprint arXiv:2310.01604 (2023).

101. C. Anil & X. Bao. *Learning to elect*. Advances in Neural Information Processing Systems **34**, 8006 (2021).

102. H. W. Kuhn. *The Hungarian method for the assignment problem*. Naval Research Logistics Quarterly **2**, 83 (1955).

103. L. Pan, N. Malkin, D. Zhang & Y. Bengio. *Better training of gflownets with local credit and incomplete trajectories*. in *International Conference on Machine Learning* (2023), 26878.

104. K. Xu, W. Hu, J. Leskovec & S. Jegelka. *How Powerful are Graph Neural Networks?* in *International Conference on Learning Representations* (2019).

105. D. Zhang, H. Dai, N. Malkin, A. C. Courville, Y. Bengio & L. Pan. *Let the flows tell: Solving graph combinatorial problems with gflownets*. Advances in neural information processing systems **36**, 11952 (2023).

106. A. M. Turing. *Computing machinery and intelligence* (Springer, 2009).

107. J. McCarthy, M. L. Minsky, N. Rochester & C. E. Shannon. *A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955*. AI magazine **27**, 12 (2006).

108. A. Newell, H. A. Simon, *et al. Human problem solving*. **9** (Prentice-hall Englewood Cliffs, NJ, 1972).

109. C. M. Bishop & N. M. Nasrabadi. *Pattern recognition and machine learning*. **4** (Springer, 2006).

110. Y. LeCun, Y. Bengio & G. Hinton. *Deep learning*. nature **521**, 436 (2015).

111. A. Agrawal, J. Gans & A. Goldfarb. *Prediction machines, updated and expanded: The simple economics of artificial intelligence* (Harvard Business Press, 2022).

112. V. Eubanks. *Automating inequality: How high-tech tools profile, police, and punish the poor* (St. Martin's Press, 2018).

113. C. Van Ooijen, B. Ubaldi & B. Welby. *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance* (2019).

114. V. Dignum. *Responsible artificial intelligence: how to develop and use AI in a responsible way* (Springer, 2019).

115. A. Jobin, M. Ienca & E. Vayena. *The global landscape of AI ethics guidelines*. Nature machine intelligence **1**, 389 (2019).

116. B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter & L. Floridi. *The ethics of algorithms: Mapping the debate*. Big Data & Society **3**, 2053951716679679 (2016).

117. M. Gibbons & H. Nowotny. in *Transdisciplinarity: Joint problem solving among science, technology, and society: An effective way for managing complexity* 67 (Springer, 2001).

118. A. Abdulkadiroğlu, P. Pathak & A. Roth. *The new york city high school match*. American Economic Review **95**, 364 (2005).

119. *Interdisciplinarity* (eds Barry, A. & Born, G.) (Routledge, London, England, 2013).

120. J. A. Jacobs & S. Frickel. *Interdisciplinarity: A critical assessment*. Annual review of Sociology **35**, 43 (2009).

121. R. Heeks. *Information systems and developing countries: Failure, success, and local improvisations*. The information society **18**, 101 (2002).

122. K. Toyama. in *Proceedings of the 2011 iConference* 75 (2011).

123. A. Escobar. *Encountering development: The making and unmaking of the Third World* (Princeton University Press, 2011).

124. C. Mohanty. *Under Western eyes: Feminist scholarship and colonial discourses*. Feminist review **30**, 61 (1988).

125. S. Harding. *Is science multicultural?: Postcolonialisms, feminisms, and epistemologies* (Indiana University Press, 1998).

126. T. Wengraf. *Qualitative research interviewing: Biographic narrative and semi-structured methods* (2001).

127. D. Michailidis, M. Tasnim, S. Ghebreab & F. P. Santos. *Tackling school segregation with transportation network interventions: an agent-based modelling approach*. Autonomous Agents and Multi-Agent Systems **38**, 22 (2024).

128. T. Alpherts, S. Ghebreab & N. van Noord. *EMPLACE: Self-Supervised Urban Scene Change Detection*. in *Proceedings of the AAAI Conference on Artificial Intelligence* (2025).

129. M. Dankloff, V. Skoric, G. Sileno, S. Ghebreab, J. v. Ossenbruggen & E. Beauxis-Aussalet. *Analysing and organising human communications for AI fairness assessment: Use cases from the Dutch Public Sector*. AI & SOCIETY, 1 (2024).

130. V. Skoric, G. Sileno & S. Ghebreab. *Roles of Standardised Criteria in Assessing Societal Impact of AI*. in *2024 IEEE Conference on Artificial Intelligence (CAI)* (2024), 1240.

131. A. Kalender. *Towards Just Smart City: Community-based Art Organisations as Partners in Design Justice*. Available at SSRN 4850415 (2024).

132. J. Burrell. *How the machine 'thinks': Understanding opacity in machine learning algorithms*. Big data & society **3**, 2053951715622512 (2016).

133. M. Ananny & K. Crawford. *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*. new media & society **20**, 973 (2018).

134. R. A. Johnson & S. Zhang. *What is the bureaucratic counterfactual? Categorical versus algorithmic prioritization in US social policy*. in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (2022), 1671.

135. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. van der Walt, M. Brett, J. Wilson, K. Millman, N. Mayorov, A. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. Quintero, C. Harris, A. Archibald, A. Ribeiro, F. Pedregosa, P. van Mulbregt & SciPy 1.0 Contributors. *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*. Nature Methods **17**, 261 (2020).

ACKNOWLEDGEMENTS

In the Summer of 2020 during the pandemic, I sat down in the garden of my home in Dresden, and tried to make an important life decision. I knew that I loved working in Computer Science, and I loved learning new things. I knew that I was passionate about Education, as it gave me my keys to freedom, and that I believed (quite naively) that Computers and AI could be a key towards sharing this freedom with others. As such, I had almost made up my mind about doing a PhD in AI, with some form of focus towards Education.

However, I also had a chaotic mind, and lacked the skills needed to do a PhD: self-management, consistent focus, discipline, and a thick skin towards rejection. Yet in that moment, I decided I *must* do a PhD, because *surely* this was the best way to learn the skills I lacked! Five years later, I realize there are probably better and less painful ways to learn to self-manage and handle rejections. And yet, I would not exchange this experience for anything.

The reason behind this has little to do with goals related to personal growth, and everything to do with the incredible human beings I have met and worked with throughout this PhD.

First and foremost, to my promotor and supervisor Sennay, thank you for the opportunity of a lifetime, a mentorship that taught me far more than research, and for your relentless support. We often tend to value smooth roads and journeys that do not stagger, but you helped me see that *to struggle is very much the point*, and I am better and stronger for having faced the obstacles that I did. You showed faith in me in the times that I felt lost and overwhelmed, and this led me to keep calm and carry on. This thesis is what it is mainly because of you. I also suspect that the values that I learned from you will guide most of my choices from here on out.

I would also like to deeply thank my co-promotor and daily supervisor Max Baak, whose passion for the Amsterdam School Choice problem led to the beginning of this thesis. Thank you for including me in this amazing project, and then for teaching me how to take ownership of it.

Another person who has provided me with incredible support throughout this PhD is my co-promotor and daily supervisor Erman Acar. Erman, you took me in as your student in my third year of PhD without hesitation, and I have had some of the most exciting and fun months of my PhD working on Chapter 5 with you. Thank you for your kindness and generosity of spirit, and for having an inspiring research vision.

I am also grateful to my promotor Tobias Blanke, for always being available on incredibly short notice whenever I needed feedback, and for inspiring me to think outside of my computer-science-shaped box.

I would also like to express my sincerest gratitude to Fernando P. Santos. Even without officially being my supervisor, Fernando reviewed my drafts before deadlines, helped me find suitable conferences and journals, and connected me to experts who provided crucial input. Fernando, your kindness to me has been unparalleled, and I look up to you for being a model academic and human being.

My sincere thanks also go to the members of my doctoral committee prof. dr. Ulle Endriss, prof. dr. Virginia Dignum, prof. dr. Eelke Heemskerk, prof. dr. Judith Good and dr. R. T. Rădulescu for their efforts in reading and assessing this thesis.

The Civic AI Lab began in the middle of the pandemic with 7 PhD students. In a time of isolation, we met through Zoom and GatherTown, did covid-compliant outdoor picnics at the Bos, and organized co-working sessions and game nights. Through all this, I got to have some lifelong friends, who had an immeasurable impact on me.

I'd firstly like to thank my paranymphs, Dimitris and Sara. To Dimi, I'm not sure you know how much of an anchor you were to so many of us in the lab, but especially to me. On the days I was struggling, I knew if I would just drag myself into the office, I could rely on finding you there, typing away, calmly turning to me and asking "how's life?" And with that one phrase, it would feel like any other day and I'd find it in me and sit down and work for a few hours. This is because you did it every day, without fail. Thank you for being there, thank you for being you, thank you for showing me how this is done, and for making it look easy. To Sara, thank you so much for being my biggest supporter at work and in life. From Day 1, there was never a problem too big or small that I couldn't share with you and feel completely and utterly understood. Not many people understand the precise struggles of moving across the world to an unfamiliar place and forging a life there, but you did, and I am beyond grateful that I had you beside me to share this journey with.

To Mirthe, thank you for being one of the most fun people alive, for making me laugh so hard I could forget all my stresses. Thank you also for being a deeply principled colleague, from whom I learned a lot about confidence and integrity. I love that the topics we can geek out over range from pop culture to Sebastian Sallow to the unequal power structures in academia, and I wouldn't have it any other way. To Ilse, I'm grateful to you for so much. You single-handedly set the group culture of the Civic AI Lab. Thank you for being so generous with your knowledge and support in those early days when none of us knew what to do and who to go to. Thank you for setting up the most awesome office in the building (equipped with a sound system and proper coffee), and for always looking out for me. It is through this text that I'd like to reveal to you, that after 5 long years, I'm finally biking in Amsterdam. And it's every bit as fun and rewarding as you told me it would be! To Tim, with whom I share the title of being the most emotional person I know, thank you for your constant, quiet support, and for your truly underrated wisdom. I still remember the words you said to me after our 14-month assessment, when you advised me to stop being my biggest critic, as there will be no shortage of critics in the world. It was so helpful to have you there for a word of advice whenever I was stuck with something, be it the introduction of a chapter, or how to reach out to schools to do a survey.

Finally, I'd like to thank Vanja, for the endless warmth, inspiration, and encouragement. It was great to have someone as cool as you in the lab to look up to.

It takes a village to do anything of value, and for me my village was my group SIAS. In my short career I've had the opportunity to work in a variety of settings, but none so supportive and caring as SIAS. To Marta, thank you for all the board games and picnics, and for holding my hand as I *swam* through writing my thesis. Getting to work with you in my postdoc has been a dream come true. To João, thank you for your kindness and for all the coffees spent listening to all my troubles. To Adia, thank you so much for all the long lunches, walks and a kind listening ear, and for the

amazing handmade cup I will always treasure. To Alex, thank you for bringing joy and light to all my lunches, and for making me feel old and wise. To Gabi and Angela, thank you for teaching me the importance of movement as I deal with the less fun side of my 30s. To Madhura and Ece, thank you for subsequently making it hard to diet with all the chocolates and delicious treats. To Martin, who helped me reframe the introduction to my first journal submission, thank you for your passion and for always being ready to help. To Tijs, Alexandra, Darcy, and Aida, thank you for the amazing discussions at the Philo-socio-technical sessions, which truly broadened the scope of my thinking about technical systems. A special thanks to Tijs for being a truly inspirational collaborator, and I hope I have the good fortune to work with you again on your amazing projects. To Arjan, thank you for your wisdom and for always being available to give me advice on my research. To Roman, Ege and Satchit, thank you for breathing new life into the group and making the office a *truly cool place*.

To Virginie, thank you for your tireless effort behind helping this new lab and group stand on its own. You were always there for us, no matter how busy, no matter how tired. No matter how big or small the problem, I could bring it to you and be met with warmth and a quick solution. You are the *best*, Virginie.

I am also deeply grateful to Giovanni, who provided me with invaluable research and career advice, and then gave me the opportunity to continue this line of work in a postdoc. Through conversations with you, I learned to step back and view the larger context of my work, and reflect on the limitations of narrow definitions.

And lastly, a great, big thank you to my co-author and work friend Paul, for helping conduct the survey, speaking to parents, and for being a great sounding board for my research ideas. Thank you for all the coffees and conversations, and for allowing me to be free with my thoughts and critiques, which has helped me grow as a researcher, as well as a person.

I also want to thank the people who are no longer at SIAS but who had an impact on me. To Aida, thank you for our long conversations and for helping me learn about the invisible mechanisms behind AI in the public sector. To Hinda, thank you for being the sort of leader who was always in our corner. To Katerina, thank you for your endless kindness to me – every time you saw me, you had nothing but words of encouragement, and they have always meant a lot.

This thesis was shaped by many hands beyond my own. For Chapter 2, I am grateful to Youri Weesie, whose early collaboration helped set this work on firmer ground. Chapter 3 would not exist without Joseph Trevorrow's joint authorship and the thoughtful guidance of Nirav Ajmeri. It is no easy feat to collaborate across countries with mismatched schedules, but you remained patient and believed that this work was something important, and for this I am grateful to you both. Chapter 4 also would not exist without the people who understand Amsterdam's schools far better than I do: Lidewij Koren and Floor Kaspers at OCO — thank you for going out of your way to connect me to the parents of Amsterdam.

To the many parents of high schoolers across Amsterdam who gave me their time and their honest answers: *thank you*, this research rests on your trust. I am also deeply grateful to Nicolas Choquette-Levy and Elisabeth Krueger for providing early feedback and advice for Chapter 4, which helped immensely in shaping its direction. A huge thanks to Eric Dignum, who helped me get my first dataset and subsequently inspired my research with his work on school segregation. Eric, I'm extremely grateful

that I met you while standing awkwardly near the snacks table at my very first IvI All Hands meeting, which turned out to be very lucky for me.

This work would have stayed on the whiteboard without the people at the Municipality of Amsterdam who were willing to engage with it. Thank you to Ralph Rusconi, Frederique van Spijker, and Isabelle Schultz from the education department, for sharing data and context and helping me get my bearings in the early days.

I also had the privilege of supervising a few brilliant students along the way — Aiden Ta, Diederik Salimans, and Mart Becx. Thank you for letting me be part of your work; if I am honest, I came away from our meetings having learned more from you than you probably did from me.

Beyond my own group and institute, I have been lucky to cross paths with some amazing academics across the Netherlands and beyond. First off, I'd like to deeply thank Ulle Endriss. Ulle, your kind invitation to the COMSOC seminar meant that I was exposed to an entirely new domain, but one that was fundamental to my research area. Your exceptionally written lecture notes and books were a huge resource for me. I do not know how you have enough time in a day to be a professor and still remember to email a first year PhD with the link to a relevant workshop — but I hope you know that you had a great impact on my work.

I'd also like to thank the entire C-MAS community, but especially Sebastian Stein and Vahid Yazdanpanah for setting up a community for Citizen-centric research. I'm grateful, too, to my fellow organizers of the Citizen-centric AI systems seminar Jessica Woodgate, Enrico Liscio, Jim Dilkes and Shubhalaxmi Mukherjee. Thank you for volunteering your time to help organize this seminar, exposing me to some amazing research which served as a great inspiration.

A special thanks to Bertilla Fabris, who I met during the Responsible AI retreat in Umeå. We joked that meeting each other was like finding our "research soulmate", and I honestly regret not meeting you sooner. I hope the future will lead to more opportunities for us working together.

A warm thank you, too, to Savvina — in addition to being my singing buddy, you inspired me greatly with the way you approach research. I remember the day you gave a talk at SIAS where you presented your papers, weaving a very cool story about how it all fit together. It helped me untangle my own questions about my research and how they fit together. Savvina, my friend, my wonderful paranymph-in-law, you simply are the *coolest*.

When I moved to Amsterdam in 2021, it was the middle of the pandemic and it was a time of deep isolation. Despite this, I somehow found an amazing group of friends who helped me survive in this new city. To Zuranee, Heetal and Kavita, thank you so much for making me feel like I'd known you all my life. Thank you for the dinners, the birthday celebrations, and for always being available when I needed to vent about my never ending PhD.

It is also never easy to be away from family, more so when you are an entire continent away. Yet, I somehow found a second family in Nixon bhaiya and Sara apu, Tanjina apu and Rashid bhai, Imon and Jannat. Thank you so much for being there for me in the good days, but especially in the really bad days. We lived through some historically bad times for our country, but I'm so grateful we had each other to lean on. I would also like to especially thank Nixon bhaiya's haleem, Tanjina apu's duck

curry and Imon's peanut bhorta–these special tastes of home literally breathed new life into me.

I feel like I am a person representing a sum of all the people that have loved me, but none more so than my best friends, who have shaped me as a person for nearly 25 years. To the incredibly talented Adiba: there is not another person in this world that I can call in the middle of a workday and ask to design a book cover, and then have the audacity to only give them 5 days, but with you, somehow I always took these liberties. Thank you for carrying me, for loving me, for supporting me through everything. To Nabila, the smartest and kindest person I know, thank you for listening to hours worth of rants about my research and then giving me the simplest solutions to all my problems. To Parsa and Vatsala, my fellow academics across the world, thank you for the compassion and the inspiration you gave me when I was stuck. To Aboni and Joyita, thank you for always hyping me up, and for always promising summer trips to Europe that somehow never happened. To Sami, thank you for quietly looking out for me, and for always believing that I could do this.

To my incredible family, to my amazing *Ammu*, *Baba* and my dearest Sakib, who bore the grief of sending me away, and took care of everything so I could have the privilege to live and work here: there will never be enough words to thank you, and I will not even try: this entire work is dedicated to you. I am also here because of my amazing aunts and uncles, who took care of me during high school and tolerated the horrible Dhaka traffic so I could do my SSC exams on time. To my dear *mejho mama* and *mami*, you are not here today, but I am here because of you. I'd also like to thank my dear *Attammu* and *mejho khalu* – thank you all for taking care of me like I was one of your own.

And finally, to Abir. You left your whole world behind to come be with me, and you've stayed beside me through all the best and the worst days of this PhD. How can I even begin to make up for this? I can't, but I will spend the rest of my life trying to. Thank you for everything, *bhalobasha*.

Mayesha
Amsterdam, August 2026

