Market Power and Dropping Strategies: A New Approach to Stable Matching Mechanisms

Itai Shapira[1]

[1]Harvard University
March 24, 2023

Abstract

In this paper, we present the dropping lemma and the concept of market power to simplify the analysis of one-to-one matching markets. Building on the approach first introduced by Kojima and Pathak [2009], we develop a constructive and intuitive framework for analyzing the game induced by an arbitrary stable mechanism. Applying this framework, we offer a novel proof for the one-sided incentive compatibility of the men-proposing deferred acceptance algorithm and establish the necessary and sufficient conditions for an agent to effectively manipulate stable mechanisms through market power. Moreover, we introduce a rejection chain algorithm to identify players capable of manipulating stable mechanisms and investigate the associated equilibria. Our contributions provide valuable insights for computer science, particularly in the design and analysis of matching mechanisms, and pave the way for further research on the interplay between strategy, market power, and incentives in matching markets.

1 Introduction

The study of matching mechanisms has come a long way since Gale and Shapley's groundbreaking 1962 paper (Gale and Shapley [2013]), which introduced the stable matching problem. In two-sided markets, matching mechanisms are used to pair agents from one side of the market with those on the other side, such as firms and workers or men and women. A centralized clearinghouse facilitates the matching process based on agents' preferences. The fundamental theoretical standpoint of matching theory is the strategic behavior of agents when a stable matching mechanism is employed.

This paper presents a novel approach to analyzing matching theory, which leverages the concept of *dropping strategies* to offer a more intuitive and streamlined analysis of agents' strategic behavior in the context of stable mechanisms. In the one-to-one simple model, a dropping strategy involves declaring some agents on the other side of the market as unacceptable, while maintaining the relative order of acceptable agents. The idea of dropping strategies was touched upon in a different context by Kojima and Pathak [2009]. By considering only dropping strategies as optimal strategies for players in a game induced by a stable mechanism, we can more constructively derive several well-known results in matching theory.

One of the key contributions of this paper is the introduction of the concept of *market power*, which provides a new perspective on matching theory and offers a powerful and economically meaningful interpretation of the ability to manipulate stable mechanisms. Our approach sheds new light on matching theory by offering a more focused analysis and greater insight into agents' strategic behavior in the context of stable mechanisms. The results in this paper can be viewed as both providing new, intuitive, and constructive proofs for existing results and introducing new results that enhance our understanding of strategic behavior in stable matching mechanisms. For instance, we present an alternative proof for the established result that the men-proposing Deferred Acceptance (DA) mechanism constitutes a dominant strategy for each man to reveal his true preferences. While the original proof for this statement was notably complex and intricate, our novel approach offers a more streamlined and accessible demonstration. This refined proof serves as a noteworthy and valuable addition to the literature, even though the underlying theorem itself is not a new discovery.

2 Preliminaries

In this section, we briefly introduce the essential concepts and terminology related to matching theory and the Deferred Acceptance (DA) algorithm. We present a simplified two-sided one-to-one matching model, playfully referred to as the *stable marriage problem*, which consists of non-empty sets of men M and women W. Each man $m \in M$ and woman $w \in W$ has a strict preference relation (strict total order relation) over the set of potential partners and the option to remain unmatched.

A *preference profile* is a sequence of preference lists $P = (P_i)_{i \in M \cup W}$, and a *marriage market* is a tuple $\Gamma = (M, W, P = (\succ_i)_{i \in M \cup W})$. A matching μ is a mapping that satisfies specific conditions, and it is said to be *individually rational* if it meets certain criteria. Since only rankings of acceptable mates matter for us, we often omit unacceptable mates, and write e.g: $\succ_m$: w_1 w_2, meaning man m prefers w_1 the most, then w_2 and w_1 and w_2 are the only acceptable women. A matching is a mapping $\mu : M \cup W \longrightarrow M \cup W$ such that (i) for every $x \in M$, $\mu(x) \in W \cup \{x\}$ and for every $x \in W$, $\mu(x) \in M \cup \{x\}$; and (ii) $\mu^2 = id$, i.e for every $x \in M \cup W$, $\mu(\mu(x)) = x$. For some $m \in M$, if $\mu(m) = w \in W$, we say that w is the wife of m and m is the husband of w. If $\mu(i) = i$, we say that agent i remains single or unmatched.

A matching μ is said to be *individually rational* if for all $i \in M \cup W$, $\mu(i) \succeq_i i$. A pair $(m, w) \in M \times W$ is said to a blocking pair if $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. A matching μ is said to be *stable* if it is both individually rational and there exists no blocking pairs. If there exists a stable matching μ such that $\mu(x) = y$ and $x \neq y$, we say that y is a *stable spouse* for x, and vice versa. Stability of a spouse depends upon the preference profile P that is reported. If we wish to emphasize this fact we can say that y is a P-stable spouse of x.

The *stable matching problem* involves finding a stable matching for a given market. Gale and Shapley introduced this problem and developed the men-proposing Deferred Acceptance (DA) algorithm to find stable matchings. The algorithm iteratively processes proposals until every man is either matched or has been rejected by all potential partners. Gale and Shapley proved that the men-proposing DA algorithm generates men-optimal and women-pessimal stable matchings. A stable mechanism is a mechanism that produces a stable matching for every market based on reported preferences. We denote a stable mechanism by φ and the resulting matching by $\varphi(\Gamma)$. Investigating the stable mechanism induced by the DA algorithm can reveal underlying properties of any stable mechanism and the concept of stability itself.

Theorem 1 (Rrural Hospitals Theorem (RHT))**.** *In all stable matching of a given market $\Gamma = (M, W, P)$, the set of agents who remain single is the same. Formally, the function from the set of stable matchings to the set of subsets of $M \cup W$ that outputs the set of unmatched agents is constant.*

In practice, preferences are private information, and matching clearinghouses that employ a stable mechanism must elicit preferences from participants. This process introduces strategic aspects to matching theory, as each agent in the marriage market must decide on a preference ordering to report to the mechanism. We define a normal form game in which the set of players consists of all agents, and their strategies involve preference lists. The stable mechanism specifies a stable matching for any reported preferences, and we define manipulation as the ability to misstate preferences and benefit when others report truthfully.

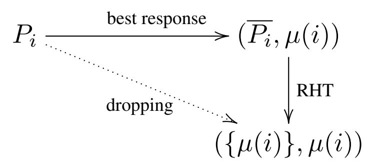
The scope of potential manipulations can be quite large in any sizable market. We now provide the formal definition of two important classes of strategies. One specific way an agent i can misreport their preferences is by simply declaring some agents who are acceptable under the true preference list as unacceptable without changing the relative ordering of acceptable agents. We call this type of reported list a *dropping strategy* (Kojima and Pathak [2009]). A *truncation strategy* of agent i is an order list $\succ'_i$ obtained for $\succ_i$ by making a tail of agents unacceptable, i.e., a dropping strategy that drops agents from the bottom.

3 Strategic Analysis Using the Dropping Technique

3.1 The Dropping Lemma

In this section, we focus on the expected behavior of a single individual in a marriage market. Analyzing the incentives of an agent $i \in M \cup W$ to misrepresent their preferences can be complex, as it requires considering all possible strategies. The *dropping lemma* plays a crucial role in simplifying this analysis by showing that the set of optimal strategies always intersects the set of dropping strategies.

Theorem 2 (The General Dropping Lemma)**.** *Let φ be a stable mechanism, $i \in M \cup W$, and P_{-i} the strategies of other players. Then, there exists a best response strategy P_i which is a dropping strategy. Moreover, a singleton preference list can replicate any best response.*



Proof. Let $\overline{P_i}$ be a best response to P^*_{-i}. Denote by $\mu := \varphi(\overline{P_i}, P^*_{-i})$ the resulting matching. If i is single under μ, then the empty strategy (reporting every agent as unacceptable) is a best response which is a dropping strategy. Otherwise, consider the strategy:

$$P^*_i : \mu(i)$$

i.e a reported list in which i list only his/her match under μ to be acceptable. $\mu(i)$ is acceptable with respect to the true preference P_i and thus P^*_i is a dropping strategy. We assert that the same match μ is still stable with respect to (P^*_i, P^*_{-i}). Since the preferences of all agents but i have not changed, every agent but i still finds his/her spouse to be acceptable, and i declares $\mu(i)$ to be acceptable. Hence μ is individually rational. If (m, w) blocks μ with respect to the preference profile (P^*_i, P^*_{-i}) then $i \in \{m, w\}$ because otherwise (m, w) blocks μ with respect to $(\overline{P}_i, P^*_{-i})$, contradicting stability. But $i \notin \{m, w\}$ since i is matched to $\mu(i)$, his/her most preferred spouse when reporting P^*_i. Thus μ is still stable.

Since no agent but $\mu(i)$ is acceptable, i is either unmatched or matched to $\mu(i)$ in $\varphi(P^*_i, P^*_{-i})$. Since the set of unmatched agents is the same in all stable matching and μ is a stable matching in which i is matched, $\varphi(P^*_i, P^*_{-i})(i) = \mu(i)$. □

This lemma has several important implications. First, it significantly reduces the complexity of our analysis by allowing us to focus on a smaller class of strategies—the dropping strategies. This reduction in complexity is evident when considering that the number of strategies of size one or less for one man in a market with five women is only $n+1 = 6$, compared to 320 other different strategies. Furthermore, one can extend this result by arguing that a truncation strategy that drops agents only from the end of the preference list can weakly improve any strategy.

The argument presented in this section can be generalized to many-to-one settings. Kojima and Pathak [2009] have proven the following statement for a general many-to-one model:

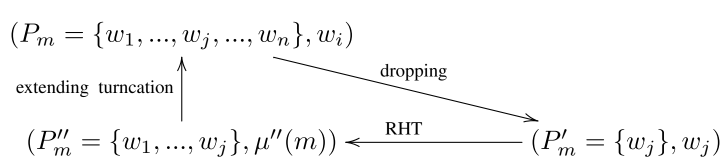
Corollary 3 (The Dropping Lemma, Kojima and Pathak [2009])**.** *Let $i \in M \cup W$. Suppose that i can successfully manipulate by reporting $\overline{P_i}$. Then i can successfully manipulate by reporting a dropping strategy. In particular, i can report a singleton.*

3.2 Applying the Dropping Lemma to DA Men-Proposing

We now turn our attention to the men-proposing Deferred Acceptance (DA) algorithm and the incentives for men in this mechanism. The DA men-proposing procedure yields the most preferred stable matching for the men. Consequently, a strategy P'_i is weakly better than P''_i if the set of stable wives for i in the former contains the set of stable wives in the latter. Using the dropping lemma, we prove the following theorem:

Theorem 4. *DA men-proposing makes it a dominant strategy for each man to state his true preferences.*

This theorem was first proven by Roth [1982] and Dubins and Freedman [1981]. Our approach differs from theirs as we apply the dropping lemma to identify a sequence of weakly-improvement strategies that lead to truthful revelation. This alternative method emphasizes the use of the dropping lemma and its simplifying effect on the analysis of incentives. The result can be generalized to a more general model of many-to-one matching.



proof of theorem 4. Let m be some man and fix P_{-m} a reported lists of other players. Suppose that by reporting his true preference list $P_m = \{w_1 \succ_m ... \succ_m w_n\}$, m is matched to $w_i \in W \cup \{m\}$. Suppose by contradiction there exists a preference list P'_m and $w_j \in W$ such that $\varphi(P'_m, P_{-m})(i) = w_j \succ_m w_i$. By the dropping lemma, we can assume $P'_m : w_j$. Consider the preference list:

$$P''_m : w_1, ..., w_j$$

i.e, a truncation of the real preference list, P_i at w_j. Consider the resulting (P''_m, P_{-m})-stable matching $\mu'' := \varphi(P''_m, P_{-m})$. μ'' is (P'_m, P_{-m})-pairwise stable, since the preferences of all other agents have not changed and the pair (m, w_j) would have (P''_m, P_{-m})-blocked μ'' if it were a blocking pair. μ'' is individually rational in (P'_m, P_{-m}) to all agents other than m. If m is single in μ'', then μ'' is (P'_m, P_{-m})-individual rational, and hence stable. Therefore, by the rural hospital theorem (RHT), m is not single at μ''. By truncation it follows that:

$$\mu''(m) = \varphi(P''_m, P_{-m}) \succeq_m w_j \succ_m w_i$$

As mentioned above, in the deferred acceptance men-proposing algorithm, extending a man's truncated list in compliance with the true preferences is always weakly beneficial. It follows:

$$w_i = \varphi(P) \succeq_m \mu''(m) \succ_m w_i$$

A contradiction, as desired. □

In conclusion, the dropping lemma significantly simplifies the analysis of incentives in a marriage market by allowing us to focus on a smaller class of strategies. By applying the dropping lemma to the men-proposing DA algorithm, we demonstrate that it is a dominant strategy for each man to state his true preferences. This novel approach highlights the power of the dropping lemma in understanding incentives and strategies in matching markets.

4 Market Power

4.1 The Market Power Lemma

In this section, we explore the concept of market power and its role in understanding the ability of agents to manipulate stable mechanisms. We present the necessary and sufficient conditions for an agent to effectively manipulate stable mechanisms.

Consider an instance of a marriage market and an agent $i \in M \cup W$. Let φ denote an arbitrary stable mechanism, and let us analyze the extent of possible manipulations in the game induced by φ. We assume that all agents reveal their true preferences. If agent i has at most one stable spouse, the Rural Hospital theorem implies that any stable mechanism must yield the same outcome for i. Conversely, if i has two or more stable spouses with respect to the preference profile P, the stable mechanism chooses a P-stable spouse for i from multiple options.

In the latter case, the decision-maker has the liberty to choose a stable spouse for i among different options. However, if i declares only their most preferred stable-spouse to be acceptable and all other stable spouses to be unacceptable, they effectively reduce the options available to the mechanism, forcing it to match i with their most preferred spouse. Consequently, having more than one stable partner can be considered as having *market power* - an attribute of a player (given the preferences of others) that enables them to profitably manipulate any mechanism. An agent with market power can guarantee themselves the best (truth-preferences)-stable partner. The lack of manipulability stems from having no market power, meaning that strategically rejecting possible mates cannot improve a player's match.

To formalize this concept, we define market power as follows:

Definition 5. A player $i \in M \cup W$ has P-*market power* or market power with respect to a preference profile P, if i has more than one P-stable spouse.

The Market Power Lemma states:

Corollary 6 (The Market Power Lemma). *Let (M, W, P) be a marriage market, φ be any stable mechanism, and $x \in M \cup W$. Then x cannot benefit by misreporting if x has no market power. Conversely, if x has market power, x can obtain any P-stable spouse, and the optimal strategy yields the best P-stable spouse for x.*

Proof. Without lost of generally, assume $x \in M$ (the proof is the same for $x \in W$) and denote by y his only stable spouse if exists, and $y = x$ otherwise. Suppose by contradiction x can report P'_x and $y' := \varphi(P'_x, P_{-x})(x) \succ_x y$. By the dropping lemma, we can assume $P'_x : y'$. Since there exists a stable matching in which x is matched and the set of unmatched agents is the same in all (P'_x, P_{-x})-stable matching, under the men-proposing DA matching mechanism, x is matched to y' when reporting P'_x. By assumption, x is matched to y under the DA algorithm when reporting P_x, contradicting theorem 4.

Conversely, assume there is a P-stable matching μ in which $x \neq y = \mu(x)$. Consider the strategy $P'_x : y$. The matching μ is (P'_x, P_{-x})-stable, and hence $y = \varphi(P'_x, P_{-x})(x)$. In addition, if the optimal strategy for x yields a matching μ such that $\mu(x)$ is strictly better for x than the best stable spouse for x, then the resulting matching form the DA-algorithm when x is in the proposing side is weakly better for x than μ, implying that x, a member of the proposing side, has successfully manipulate the DA algorithm, a contradiction. □

The market power lemma serves three main purposes from different theoretical standpoints. First, it differentiates between players who can manipulate the mechanism and those who cannot. Second, it establishes a limit to the degree a player can benefit from misreporting, indicating that the optimal strategy yields the most favorable stable spouse. Third, the lemma offers an intriguing economic interpretation. An agent can only gain advantage by misreporting if they possess sufficient popularity to have multiple stable partners, thereby enabling them to compel the mechanism to match them with their top-choice spouse by strategically limiting alternative matchings from which the mechanism must select. Moreover, the Market Power Lemma demonstrates that the men-proposing Deferred Acceptance (DA) algorithm is optimized for men. Consequently, the strategic-proofness outcome bears resemblance to the properties expected from an incentive-compatible mechanism in mechanism design, as outlined in Nisan et al. [2007] (p. 226).

4.2 Rejection Chains

We introduce the concept of rejection chains to understand how we can identify players who can manipulate stable mechanisms. The proof of the market power lemma, a general property of any arbitrary stable mechanism, was essentially constructed by "embedding" the given mechanism within the DA matching mechanism and deriving the desired result using the one-sided incentive compatibility property of the DA mechanism. This technique is similar to the arguments presented in Immorlica and Mahdian [2005] and Storms [2013], where a general property in matching theory is derived by analyzing the one-sided favorite, the DA algorithm.

Similarly, we can utilize this line of reasoning to design an algorithm that checks whether an agent i can benefit by unilaterally misreporting. A straightforward approach would be to compute the men-proposing DA and the women-proposing DA, and then compare them to see if i has more than one stable spouse. However, both theoretically and computationally, this method can be cumbersome. Instead, we propose the following procedure:

Algorithm 1 (The Rejection Chain Algorithm)**. Input:** some $w \in W$. Denote $k = |W|$.

1. Initialization: run the man-proposing algorithm. For each $m \in M$, let U_b the set of women he proposed to. If w is single, terminate.
2. Let $b = \mu(w)$, and let w reject b.
3. (a) If $|U_b| \geq k$, terminate. Else:
   (b) Select w' to be b's most preferred women among those he hasn't proposed to yet, add w' to U_b.
4. (a) If w' has received a proposal from a man she likes better than b, she rejects b and the algorithm return to step 3.
   (b) If not, w' accepts b.
      i. If $w' = w$, the algorithm terminates.
      ii. Else, if w' was previously married, set b to be her previous husband, and return to step 3.
      iii. If w' was previously unmarried, terminate the algorithm.

Given that w is married after step 1, this algorithm terminates either at step $4(b)i$, at step $3(a)$ or at step $3(b)iii$. We say that the rejection chain algorithm returns to w if it terminates at step $4(b)i$.

Theorem 7 (Rejection Chain Lemma)**.** *Let $w \in W$. If the rejection chain algorithm does not returns to w, then w cannot misreport any stable mechanism.*

This result was used in Roth and Peranson [1999], Immorlica and Mahdian [2005], Kojima and Pathak [2009] and Storms [2013]. The following is a straightforward proof based on the dropping lemma and the market power lemma.

Proof. w can misreport a stable mechanism implies she has market power. In particular, she can obtain a different stable spouse by misreporting the men-proposing DA algorithm using a dropping strategy. The process described in the rejection chain algorithm is induced by this dropping strategy. □

In conclusion, the concept of market power and the Rejection Chain Algorithm allow us to identify players who can manipulate stable mechanisms and better understand the implications of such manipulation.

5 Equilibrium Analysis

In this section, we investigate equilibria in the game induced by a stable mechanism. We begin with an arbitrary stable mechanism φ and continue to the men-proposing DA algorithm. In the latter case we assume each man chooses his dominant strategy to state his true preferences. One should note that there are other types of equilibria in this game, e.g. reporting the empty order list.

Corollary 8. *A profile of preference P' is an equilibrium if and only if every matched agent is matched to his most preferred P'-stable spouse.*

Proof. Consider a player $i \in M \cup W$. If all other players follow the strategy P', then by the market power lemma, i has a strategy that matches him/her with the most preferred P'-stable spouse. Equilibrium, therefore, implies he/she is matched with the most preferred stable spouse. Conversely, if P' is not an equilibrium, there exists a player $i \in M \cup W$ who can benefit by deviating from P'. Again, it follows from the market power lemma that i has more than one stable partner and i is not matched to the best one when reporting P'. □

Corollary 9 (Roth's impossibility theorem)**.** *There exists no incentive compatible stable mechanism, i.e., no stable mechanism exists for which stating the ture preferences is always an equilibrium.*

Proof. Truth-telling P is an equilibrium only if each agent is matched to his most preferred stable partner, only if the women-optimal matching is the men-optimal matching, only if there is only one P-stable matching. Hence, every instance of true preferences P in which there is more than one stable matching is a counterexample. □

Theorem 10 (Roth and Sotomayor [1992])**.** *Consider the DA-men proposing mechanism. Suppose each man chooses his dominant strategy and states his true preferences. Then any stable matching can be obtained by an equilibrium. Conversely, any equilibrium produces a matching that is stable with respect to the true preferences.*

Proof. Let μ be a P-stable matching. Suppose each μ-matched woman w chooses the strategy $P'_w : \mu(w)$. Denote by P' this strategy profile and we assert P' is the desired equilibrium. This follows from the fact that μ is the unique P'-stable matching.

Conversely, suppose μ is an individually rational unstable matching. Then there exists a blocking pair (m, w). Hence, m must have proposed to w during the implementation of the men-proposing DA algorithm. It follows that reporting m as her first choice but otherwise agreeing with P_w leads to a strictly preferred outcome for w in which she is matched to m, implying P is not an equilibrium. □

6 Conclusion

In this paper, we delved into an approach to the analysis of matching theory by emphasizing dropping strategies as the optimal strategies for players in games induced by stable mechanisms. While inspired in part by Kojima and Pathak [2009], our exploration allowed us to develop novel insights and results in the field of matching theory, expanding upon the original work.

Our analysis was grounded on the dropping lemma, which states that a reporting of a singleton-preference list can replicate any best response. Combined with the rural hospital theorem, we were able to establish the one-sided incentive compatibility of the men-proposing DA algorithm, which has been proven through other, more laborious means. This new perspective allowed us to delve deeper into the dynamics of matching markets and contribute fresh findings to the existing body of knowledge.

We introduced the concept of market power, which provides sufficient and necessary conditions for an agent to have the ability to manipulate stable mechanisms. The value of adopting the concept of market power lies in its capacity to simplify the analysis of matching theory, provide an economic interpretation of the ability to manipulate stable mechanisms, and expand our understanding of the dynamics in matching markets by offering new insights and perspectives.

It is worth mentioning, however, that our analysis is based on the assumption of perfect information, and the dropping lemma may not be robust to changes in the information structure. The dropping lemma implies a player's optimal strategy is to report only their most preferred stable spouse, but in real-life applications of matching theory, the information required to correctly compute the optimal strategy may be out of reach for the players in some environments. Even if agents only have probabilistic knowledge of other agents' preferences, the expected utility from these strategies could be low, as mistakes may leave an agent unmatched. Despite this limitation, the theoretical virtues of this approach remain valuable.

Finally, we note that most of the results presented in this paper can be generalized to more complex models, such as random markets, many-to-one models, and matching with contracts. Further research in these areas has the potential to yield valuable insights into the practical applications of matching theory and uncover potential improvements to existing mechanisms. The introduction of market power as a new contribution to the theory opens up exciting opportunities for future exploration and can potentially lead to a deeper understanding of the interplay between strategy, market power, and incentives in matching markets.

References

Lester E Dubins and David A Freedman. Machiavelli and the gale-shapley algorithm. *The American Mathematical Monthly*, 88(7):485–494, 1981.

David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 120(5):386–391, 2013.

Nicole Immorlica and Mohammad Mahdian. Marriage, honesty, and stability. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 53–62. Society for Industrial and Applied Mathematics, 2005.

Fuhito Kojima and Parag A Pathak. Incentives and stability in large two-sided matching markets. *American Economic Review*, 99(3):608–27, 2009.

Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007.

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

Alvin E Roth and Elliott Peranson. The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American economic review*, 89(4):748–780, 1999.

Alvin E Roth and Marilda Sotomayor. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541, 1992.

Evan Storms. Incentives and manipulation in large market matching with substitutes. Technical report, Citeseer, 2013.