Essays on Matching and Obvious Dominance

Mariya Halushka

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Doctorate in Philosophy Degree in Economics

Department of Economics
Faculty of Social Sciences
University of Ottawa

Abstract

This thesis presents three chapters.

In Chapter 1, I propose a simple one-to-one matching model, where individuals on one side have private information that affects the preferences of the individuals on the other side. I show the existence of the stable and strategy-proof mechanism in this environment. I present an algorithm that defines this mechanism—the Serial Dictatorship algorithm with cutoffs. I also consider the concept of obvious strategy-proofness. I first consider the case where only preferences, but not experience levels, are sellers' private information. For this case, Serial Dictatorship with cutoffs elicits preferences in an obviously strategy-proof way. On the other hand, when only experience levels, but not preferences are private information, I show that there is no obviously strategy-proof and stable mechanism. A consequence of the latter result is that obvious strategy-proofness is incompatible with stability.

Chapter 2 considers settings with rich private information — an agent's type may include private information other than just his preferences. In such settings, I identify a necessary condition for obviously strategy-proof implementation of social choice rules. I consider applications to strict preferences, matching and object allocation.

The main assumption behind the obvious dominance is that agents might be cognitively limited and can not engage in contingent reasoning at all. This assumption is unreasonably weak compared to the standard assumption that agents can perfectly distinguish contingencies. In Chapter 3, I strengthen it slightly by assuming that agents are able to do at least some contingent reasoning. I define what it means for the strategy to be obviously dominant with respect to a partition of the state space. I call such strategies *partition dominant strategies*. A strategy is an *almost obviously dominant* if, for all possible partitions, but not for the coarsest, it can be identified as being partition dominant. My hypothesis is that even though some agents can not do state-by-state reasoning as rational players do, they are able to do at least some partitioning of the other player's actions and regardless of how the partitioning is done, the agents can identify an almost obviously dominant strategy.

Declaration of Authorship

I, Mariya Halushka, declare that this thesis titled, “Essays on matching and obvious dominance” and the work presented in it are my own.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ____

Date: 2022 May 13

General Introduction

In the beginning, I was interested in private non-preference information revelation problems; meaning private information that does not affect one's own preferences, but can affect the others' preferences. I started with the simple two-sided matching model when one sides' preferences depend on private information of the other side. I was able to give very natural and convincing solution to it if all one asks for is strategy-proofness and stability. Then, I strengthened it to obvious strategy-proofness and things started to fall apart. Therefore, the rest of the dissertation is focused on answering the following questions: Even though things work out really well with strategy-proofness, why isn't it working with obvious strategy-proofness? In what sense an obvious strategy-proofness is too strong in this context? What are some ways in which we can weaken obvious strategy-proofness?

This dissertation includes three chapters.

Agents have private information that reflects some of their intrinsic characteristics and behaviour as well as extrinsic capabilities, qualifications, and experience. These extrinsic capabilities can affect the preferences of other agents. In Chapter 1, I study a one-to-one matching model with this feature. On one side of the market, there are sellers with different levels of experience that provide services. On the other side of the market, there are buyers that have projects and are in need of services. In addition, they have specific requirements regarding these projects reflected in the levels of experience needed to accomplish them. An economy is described by the private information of the sellers—their preferences and levels of experience—and buyers' needs. A mechanism maps each economy to a matching. I propose a strategy-proof mechanism that selects stable, individually rational and Pareto efficient matchings—the SD mechanism with cutoffs.

I show that, when there are no ties in sellers' experience levels, there is a unique stable matching for each economy due to the structure of the preferences. The SD mechanism with cutoffs chooses the allocation picked by the SD algorithm with cutoffs and finds the unique stable matching. Moreover, the SD mechanism with cutoffs is strategy-proof. Despite of the fact that SD mechanism with cutoffs is strategy-proof, it is not straightforward that agents, in particular sellers, will reveal their private information (either preferences or experience levels) truthfully. Therefore, I focus on the much more demanding concept of obvious strategy-proofness (OSP) (Li (2017)).

If a social choice rule elicits private information in a way that no agent ever has any incentive to misreport, then it is strategy-proof.[1] However, in practice it is not enough to make agents understand the rule for them to be truthful (Kagel et al., 1987; Charness and Levin, 2009; Martínez-Marquina et al., 2019a). Therefore, I focus on "obvious dominance" in Chapter 2. It is a much stronger property than weak dominance. An obviously dominant

[1] A social choice rule is strategy-proof if it is a weakly dominant strategy for each agent to reveal his private information truthfully in the normal form game induced by the rule.

strategy is one that can be identified as dominant even by agents who can not engage in contingent reasoning. This means that the worst outcome from the obviously dominant strategy is at least as good as the best outcome from any deviation (Li, 2017). A social choice rule is obviously strategy-proof (OSP) if there exists some extensive game form such that, for every profile of preferences, there is an equilibrium in obviously dominant strategies and the outcome coincides with the outcome of the rule.

In my model, there is a finite set of agents and each agent has a type, which is his private information. I present a setting with "rich" type information, where each agent's type includes non-preference information in addition to his preferences over outcomes. A social choice rule maps each profile of types to an outcome.

My goal is to identify a necessary condition for a social choice rule to be OSP in such "rich" type environments. I define a social choice rule as being *invariant (to own non-preference information)* if whenever it responds to an agent's non-preference information, while his preferences are the same, the agent is indifferent between the outcomes. In other words, the agent is not affected by his own non-preference information that the social choice rule elicits. I prove that every OSP social choice rule is invariant. I consider the extensive form game that is not OSP, but does implement a strategy-proof rule (a modified serial dictatorship). I show that *there is no extensive game form that implements the proposed rule in obviously dominant strategies*. However, it is a highly implausible behaviour for an agent. So, obvious strategy-proofness is arguably too strong a criterion as it requires a rule to be immune to even such implausible behaviour.[2]

I consider a few applications, starting with the case of agents having strict preferences over the outcomes. I show that it is not possible to elicit any private information other than preferences in this setting. Furthermore, invariance becomes a simpler property, preference-onlyness: a social choice rule ignores private information other than preferences. Next, I present an application to matching, which is a special case of Halushka (2020). Lastly, I consider a class of sequential dictatorship rules for object allocation.

In Chapter 3, I focus on the fact that obvious dominance is too strong to make a recommendation in many games, since the assumption that agents can not engage in any contingent reasoning at all is too weak. Properties like strategy-proofness or obvious strategy-proofness say that a mechanism is not susceptible to certain kinds of behaviour (full information and full rationality in the case of strategy-proofness, inability to distinguish any contingencies in the case of obvious strategy-proofness). The greater the range of behaviours a property says a mechanism must accommodate, the stronger the property. That is why obvious strategy-proofness is stronger than strategy-proofness. My assumption is that agents are able to do at least some contingent reasoning as long as it is easy to do. Therefore, I propose one (particular) step closer to strategy-proofness from obvious strategy-proofness. That step is to not have to accommodate people who can not do any contingent reasoning at all.

Consider the following game. There are two players and only one desired object. There

[2]Some agents experience cognitive limitations when examining the contingencies (Li, 2017). However, I show that some contingencies are just too clear to be ignored even by cognitively limited agents.

are only two possible outcomes: first player or second one gets the object.[3] First, each player has to report a natural number, starting from 1. Second, the player with the lowest number reported gets the object. Ties are broken in favour of the first player. In this game, the first player has an obviously dominant strategy (ODS), which is to report 1. However, the second player does not have an ODS.

Consider the game from the second player's prospective. Suppose he can bipartition the set of all natural numbers with respect to the first player's actions; meaning he can not perform the state-by-state reasoning in the same way as a rational player can, but he can do at least some contingent reasoning. Regardless of what the partitioning is, for the partition component containing 1 and some other natural numbers, the second player can not distinguish between contingencies. He can not make a decision based on obvious dominance.[4] Moreover, for the other component of the partition (the one not containing 1): the second player is always better off reporting 1 (he always gets the object), than any other natural number.

So, the second player can identify that reporting 1 is an *almost obviously dominant strategy (AODS)* for him, even though it is not an ODS, as long as he can perform at least some contingent reasoning: not the finest partitioning as rational player does, but also not the coarsest as Li (2017) suggests. So, the strategy is an *almost obviously dominant strategy (AODS)* if, for all possible partitions, but not for the coarsest, it can be identified as being *partition dominant strategy (PDS)*; meaning, for all possible partitions, but the coarsest, there exists at least one component of the partition, over which it obviously dominates all other strategies (the preference-worst outcome following this strategy is at least as good as a preference-best outcome following any deviation) and there is no other component of the partition over which it is obviously dominated. Furthermore, the second player splits all the first player's possible actions into two sets and regardless of how he does this partitioning, as long as both sets are not empty, he can identify a PDS. I show that considering finer partitions does not change anything.

[3]There is another possible outcome when nobody gets the object. However, I do not consider it, assuming efficient allocation of the object.

[4]If the component of the partition includes only 1: the second player's worst possible outcome from reporting 1 (he does not get the object) is the same as the best possible outcome from reporting any natural number greater than 1 (he does not get the object). So, reporting 1 does not obviously dominate and is not obviously dominated.

Acknowledgments

I am particularly indebted to Vikram Manjunath whose insightful comments led to a significant improvement of the contents of this thesis. I am grateful to Vikram Manjunath for his invaluable advice, patient reviews and constructive discussions. I own special gratitude to Aggey Simons and Radovan Vadovič for helpful feedback and very useful suggestions. I thank Yuki Tamura as well as William Thomson and the reading group at the University of Rochester for their helpful comments on Chapter 2. I am very grateful for insightful comments and relevant further research ideas to Szilvia Papai and Guidon Fenig.

Contents

Chapter 1

Revealing more than just preferences: one-to-one matching when one sides' preferences depend on private information of the other side

Abstract

I propose a simple one-to-one matching model, where individuals on one side have private information that affects the preferences of the individuals on the other side. I show the existence of the stable and strategy-proof mechanism in this environment. I present an algorithm that defines this mechanism—the Serial Dictatorship algorithm with cutoffs. I also consider the concept of obvious strategy-proofness. I first consider the case where only preferences, but not experience levels, are sellers' private information. For this case, Serial Dictatorship with cutoffs elicits preferences in an obviously strategy-proof way. On the other hand, when only experience levels, but not preferences are private information, I show that there is no obviously strategy-proof and stable mechanism. A consequence of the latter result is that obvious strategy-proofness is incompatible with stability.

Keywords: non-preference revelation, interdependent preferences, stability, strategy-proofness, serial dictatorship, obvious strategy-proofness.

1.1 Introduction

Agents have private information that reflects some of their intrinsic characteristics and behaviour as well as extrinsic capabilities, qualifications, and experience. These extrinsic capabilities can affect the preferences of other agents. I study a one-to-one matching model with this feature. I consider the process of matching within the firm; meaning the match maker is needed. For example, the process of matching of computer programmers to hiring companies within the recruiting firm. On one side, there are hiring companies with certain certification and experience requirements that have to be satisfied in order for programmers to be hired. On the other side, there are computer programmers that have certain levels of certification or *experience levels*. There are many other applications of this model, such as: general contractors and customers, consulting or accounting firms and clients, cleaning firms and homeowners, beauty salons and clients, etc.

In my model, on one side of the market, there are sellers with different levels of experience that provide services. On the other side of the market, there are buyers that have projects and are in need of services. In addition, they have specific requirements regarding these projects reflected in the levels of experience needed to accomplish them.

Each seller privately knows his own experience level, which determines his ability to complete a given project. Sellers have strict preferences over the buyers and the option of being unmatched. So, for each seller, the level of experience and preferences are private information. Each buyer has a project that requires a minimum level of experience. Buyers' preferences are completely determined by this minimum level and the sellers' (privately known) experience levels. I assume that each buyer cares primarily about the completion of his project and not about the identities of the sellers. Thus, the minimum qualification to complete his project defines a cutoff for experience levels of the sellers. Given two qualified sellers, I assume that a buyer prefers the one with greater level of experience. Thus, each buyers' preferences are dependent on private information of the sellers. For each buyer, the minimum level of experience is private information.

A matching is a mapping of sellers to buyers, including the option of being unmatched. An economy is described by the private information of the sellers—their preferences and levels of experience—and buyers' needs. A mechanism maps each economy to a matching.

One of the properties of a matching that I am interested in is stability. A blocking pair is any seller and any buyer who would both prefer to be matched with each other than with their assigned partners. A matching is stable if there is no blocking pair. Next, I require a matching to be individually rational: each agent finds his assignment at least as good as being unmatched; meaning each seller and buyer is acceptable to his partner. Another requirement is that a matching is Pareto efficient: meaning there is no other matching that makes at least one person better off without making another one worse off. So, I am interested in mechanisms that select stable, individually rational and Pareto efficient matchings. Another axiom that I consider is strategy-proofness: no seller can benefit from manipulation of his preferences and level of experience and no buyer can be better off by misreporting the level of experience needed for his project. It is important to capture the situation when sellers can not falsify certificates or other documents proving a higher level of experience than the real one, but allow them to withhold such documents. I propose a strategy-proof mechanism that selects stable, individually rational and Pareto efficient matchings—the SD mechanism with cutoffs.

This mechanism is described by the Serial Dictatorship (SD) algorithm with cutoffs. The algorithm works as follows. First, break all ties in preferences of the buyers for the sellers with the same reported level of experience if there are any. Next, starting from the seller with the highest level of experience and proceeding until the seller with the lowest one, assign every seller his most preferred available buyer among those he is qualified for. If he does not find any such buyer acceptable, he is not matched.

I show that, when there are no ties in sellers' experience levels, there is a unique stable matching for each economy due to the structure of the preferences. The SD mechanism with cutoffs chooses the allocation picked by the SD algorithm with cutoffs and finds the unique stable matching. Moreover, the SD mechanism with cutoffs is strategy-proof.

Since the SD mechanism with cutoffs finds a stable, individually rational and Pareto efficient matching, which is generically unique, it is reasonable to suggest an experience monotonicity requirement. I define it as follows: increasing the experience level for one of the sellers does not make any other seller better off and any other buyer worse off. Similarly, decreasing the experience level for one of the buyers does not make any other buyer better off and any other seller worse off. Even though I find that the SD mechanism with cutoffs is buyer experience monotonic, it is not seller experience monotonic. However, when there is a single buyer it is seller experience monotonic.

Despite of the fact that SD mechanism with cutoffs is strategy-proof, it is not straightforward that agents, in particular sellers, will reveal their private information (either preferences or experience levels) truthfully. Therefore, I focus on the much more demanding concept of obvious strategy-proofness (OSP) (Li (2017)). A mechanism is OSP-implementable if there exists some extensive game form such that for every profile of preferences and, in my model, for every profile of experience levels, there is an equilibrium in obviously dominant strategies (ODS) and coincides with the outcome as if the mechanism was applied. I show that SD mechanism with cutoffs is OPSP-implementable, which is the weakening of OSP where the designer can condition the game on experience levels of the sellers. So, each seller has an ODS to truthfully pick the best available buyer whom he is qualified for (in accordance to his or her preferences). On the other hand, I prove that SD mechanism with cutoffs is not OESP-implementable, which is another weakening of OSP where the designer can condition the game on preferences of the sellers. So, I show that there does not exist an extensive game form such that for every profile of experience levels there is an equilibrium in ODS and it coincides with the outcome as if the SD with cutoffs was applied. The mechanism is OSP-implementable only if it is both OPSP- and OESP-implementable. Since the SD mechanism with cutoffs is not OESP, it is not OSP. In other words, certain cognitively limited agents, the sellers in my model, do not understand that it is in their best interest to reveal their private information truthfully; meaning if the seller can not engage in contingent reasoning, it results in untruthful information revelation.

1.2 Related Literature

A substantial part of the literature in one-to-one matching is devoted to the preference revelation problem, considering only preferences as a private information of the individuals. Sönmez (1999) shows that if there is a unique stable matching—the core is a singleton—then the mechanism is strategy-proof. This result holds for the setting where agents have only preferences as their private information. Even though the SD mechanism with cutoffs selects a unique stable allocation for each economy with no ties and is strategy-proof, my result is not implied by Sönmez (1999), since agents reveal more private information than just preferences.

Chakraborty et al. (2010) introduce incomplete information and interdependent valuations on one side of the market in two-sided matching. They study the notion of stability for matching market between colleges and students in which colleges have imperfect information and interdependent values over students. In other words, colleges receive partially informative signals about students. Furthermore, they focus on the amount of information about matchings available to colleges to define stability. Moreover, their results suggest that a stable matching mechanism does not generally exist, when the entire matching outcome is observable. However, a stable mechanism exists when colleges observe only their own matches and students have identical preferences over colleges. It may not exist if students have different preferences, even though colleges observe only their own matches. In my model one sides' preferences depend on private information—experience levels—of the other side. The existence of the stable mechanism does not depend on the available amount of information about the matching.

My paper is also related to the studies of preference interdependence in a matching context and the standard interdependent values models (Jehiel and Moldovanu (2001), Bikhchandani (2006) and Che et al. (2015)). In one of the recent works, Fujinaka and Miyakawa (2017) introduce private information on the quality of a house, known only to the initial owner, and present a model of housing markets with interdependent values where ex-post preference of an agent depends on the private information of the other agents with regard to the quality of houses. Similarly, I introduce private information on the experience levels of the agents, known only to them. However, I work with a two-sided matching model where one sides' preferences depend on private information of the other side.

Chakraborty et al. (2015) study two-sided many-to-one matching markets with interdependent valuations and imperfect information held by one side of the market. They formalize a 'modified serial dictatorship' mechanism that implements group stable matchings. I also present an algorithm that defines the stable and strategy-proof mechanism—Serial Dictatorship algorithm with cutoffs—that is also a modified serial dictatorship. However, I do not focus on group stability and I work in a different economic setting, where preferences of the one side of the market totally depend on the private information of the other side of the market. My results contribute to the literature on serial dictatorship mechanisms and their characterizations (Gibbard (1977), Bogomolnaia and Moulin (2001), Katta and Sethuraman (2006), Bogomolnaia et al. (2005), Dutta et al. (2007), Nandeibam (2013), Bade (2020), Nesterov (2017), and many others).

Obvious strategy-proofness (Li (2017)) has been studied in recent years. A direct mechanism, even though it is strategy-proof, may not give the incentives to the agents to reveal their private information truthfully in practice, depending on the implementation details, due to certain cognitive limitations. If the mechanism is obviously strategy-proof, then it can be implemented in a way that agents do reveal private information truthfully, even among agents with such cognitive limitations. Ashlagi and Gonczarowski (2018) showed that there does not exist an OSP mechanism that can return a stable matching. This implies that the Gale-Shapley matching algorithm is not OSP even though it is apparently simple. I also prove that SD mechanism with cutoffs is not OSP-implementable. Since it returns a stable matching, my findings go in hand with the ones of Ashlagi and Gonczarowski (2018). There are many closely related papers on OSP mechanisms such as those of Troyan (2019), who has identified that the acyclicity condition is both necessary and sufficient to implement Top Trading Cycles in an obviously strategy-proof way; Pycia and Troyan (2019), who have introduced the approach to simplicity to provide characterizations of simple mechanisms in general social choice environments both with and without transfers, including obvious strategy-proofness; Bade and Gonczarowski (2017), who have investigated whether some Pareto optimal and strategy-proof social choice functions are OSP in domains such as object assignment, single-peaked preferences, and combinatorial auctions.

The paper is organized as follows: I formalize the model in Section 1.3. I define the axioms in Section 1.4. I present the algorithm and the mechanism in Section 1.5. I provide the results in Section 1.6. I focus on obvious strategy-proofness in Section 1.7. I conclude in Section 1.8.

1.3 The Model

Let $S = \{s_1, ..., s_n\}$ be a finite set of sellers, with an individual seller denoted by $s_i \in S$ for each $i \in \{1, ..., n\}$; n is the total number of sellers. Let $B = \{b_1, ..., b_m\}$ be a finite set of buyers, with an individual buyer denoted by $b_j \in B$ for each $j \in \{1, ..., m\}$; m is the total number of buyers.

For each $s_i \in S$, let $l_{s_i} \in \mathbb{R}_+$ be the level of experience of s_i. For each $b_j \in B$, let $l_{b_j} \in \mathbb{R}_+$ be the minimum level of experience needed for b_j's project. Let l_S be the profile of experience levels of the sellers and l_B be the profile of experience needs for the buyers. Let l_{-s_i} denote the list of other sellers' levels of experience and l_{-b_j} denote the list of other buyers' needs.

For each $i \in S \cup B$, I use R_i to denote i's preference relation. I use P_i to denote strict preference, and I_i to denote indifference.

Let $R_S = (R_{s_1}, ..., R_{s_n})$ be the profile of preferences of the sellers, where R_{s_i} is the preference relation of the seller s_i over the buyers. For each $s_i \in S$ R_{s_i} is a strict preference over $B \cup \{s_i\}$. Let $\mathcal{R}_{s_i}$ be the set of all preference relations for s_i, I assume that $\mathcal{R}_{s_i}$ contains all strict orderings over $B \cup \{s_i\}$. [1] Let $\mathcal{R}_S$ be the set of all preference profiles for the sellers.

Each $b_j \in B$ has preference relation over the sellers, R_{b_j}. Conditional on l_s, l_{b_j} induces

[1] I assume strict preferences for simplicity. The mechanism that I present in the next section can be adapted along the lines of Bogomolnaia et al. (2005) to accommodate indifferences in sellers' preferences.

R_{b_j} over $S \cup \{b_j\}$. That is, for $\forall\ s_i, s_j \in S$:

1. $s_i\ R_{b_j}\ b_j$ if and only if $l_{s_i} \geq l_{b_j}$; meaning s_i is acceptable to b_j if and only if s_i's level of experience is at least as high as the level of experience needed for b_j's project. If $l_{b_j} > l_{s_j}$, then s_j is unacceptable for b_j. So, b_j strictly prefers being unmatched to being matched to the seller s_j—$b_j\ P_{b_j}\ s_j$.

2. If $s_i\ R_{b_j}\ b_j$ and $s_j\ R_{b_j}\ b_j$, then $s_i\ R_{b_j}\ s_j$ if and only if $l_{s_i} \geq l_{s_j}$; meaning if s_i and s_j are both acceptable to b_j and have same level of experience $l_{s_i} = l_{s_j}$, then b_j is indifferent between them, $s_i\ I_{b_j}\ s_j$. If s_i and s_j are both acceptable to b_j and $l_{s_i} > l_{s_j}$, then b_j strictly prefers seller s_i (with higher level of experience) to seller s_j—$s_i\ P_{b_j}\ s_j$.

Let $R_B = (R_{b_1}, ..., R_{b_m})$ be the profile of preferences of the buyers.

A matching, $\mu : B \cup S \to B \cup S$, is such that for each $s_i \in S$, $\mu(s_i) \in B \cup \{s_i\}$, for each $b_j \in B$, $\mu(b_j) \in S \cup \{b_j\}$, and for each pair $s_i \in S$ and $b_j \in B$, $\mu(s_i) = b_j$ if and only if $\mu(b_j) = s_i$. This means that μ is a mapping of sellers to buyers, such that only one seller can be assigned to one buyer. Let $\mathcal{M}$ be the set of all possible matchings. Let an economy be described by the preferences, the experience levels of the sellers, and the needs of the buyers: (R_S, l_S, l_B). Let $\mathcal{E}$ be the set of all economies. A mechanism, $\varphi : \mathcal{E} \to \mathcal{M}$, maps each economy to a matching.

1.4 Axioms

Here I define some axioms with respect to a rule φ.

The first requirement is that a matching is chosen only if each agent finds his assignment at least as good as being unmatched; meaning each seller and buyer is acceptable to his partner. Since the experience levels are private information, buyers' evaluation and comparisons happen only after the matching, and after buyers observe the true experience level of the assigned sellers.

Individual Rationality. For each $(R_S, l_S, l_B) \in \mathcal{E}$, and each $i \in S \cup B$,

$$\varphi_i(R_S, l_S, l_B)\ R_i\ i.$$

The next axiom is that a rule selects a matching that is not blocked by any individual or pair of seller and buyer; meaning there does not exist any seller and any buyer who would both prefer to be matched with each other than with their assigned partners. The buyer makes comparisons only after he observes the true experience level of both the assigned seller and the one he can block the matching with.

Stability. For each $(R_S, l_S, l_B) \in \mathcal{E}$, $\varphi(R_S, l_S, l_B)$ is individually rational and there is no $(s_i, b_j) \in S \times B$, such that:

$$b_j \; P_{s_i} \; \varphi_{s_i}(R_S, l_S, l_B)$$
$$\text{and}$$
$$s_i \; P_{b_j} \; \varphi_{b_j}(R_S, l_S, l_B).$$

The next property is that a matching is chosen only if there is no other matching that makes at least one person better off without making another worse off.

Pareto Efficiency. For each $(R_S, l_S, l_B) \in \mathcal{E}$, there is no $\mu \in \mathcal{M}$ such that, for each $s_i \in S$ and each $b_j \in B$, $\mu(s_i) \; R_{s_i} \; \varphi_{s_i}(R_S, l_S, l_B)$ and $\mu(b_j) \; R_{b_j} \; \varphi_{b_j}(R_S, l_S, l_B)$, and, for some $s_i \in S$, $\mu(s_i) \; P_{s_i} \; \varphi_{s_i}(R_S, l_S, l_B)$ or, for some $b_j \in B$, $\mu(b_j) \; P_{b_j} \; \varphi_{b_j}(R_S, l_S, l_B)$.

Remark 1. Stability implies individual rationality and Pareto efficiency only when there are no ties; meaning every seller has a different experience level. If there are ties, stability implies only individual rationality, but does not imply Pareto efficiency (Appendix 1.9.A).

For each $s_i \in S, l_{s_i}$ and R_{s_i} are s_i's private information; for each $b_j \in B, l_{b_j}$ is b_j's private information. Therefore, the next requirements are that no seller or buyer can benefit by misreporting this private information.

Let T and F, placed above each $l_{s_i} \in l_S$ and each $l_{b_j} \in l_B$, denote a "true" $(\overset{T}{l}_{s_i}, \overset{T}{l}_{b_j})$ and a "false" $(\overset{F}{l}_{s_i}, \overset{F}{l}_{b_j})$ level of experience of the seller and buyer, respectively.

First, I consider two axioms that concern manipulation of the levels of experience.

Seller Experience Strategy-proofness. For each $(R_S, l_S, l_B) \in \mathcal{E}$ and each $s_i \in S$, there is no $l'_{s_i} < l_{s_i}$, such that:

$$\varphi_{s_i}(R_S, \overset{F}{l'_{s_i}}, l_{-s_i}, l_B) \overset{T}{P_{s_i}} \varphi_{s_i}(R_S, \overset{T}{l_{s_i}}, l_{-s_i}, l_B).$$

Note that I restrict reported levels for s_i to be no greater, than the truth, that is: $\overset{F}{l_{s_i}} < \overset{T}{l_{s_i}}$. If s_i falsely reports $\overset{F}{l_{s_i}} > \overset{T}{l_{s_i}}$, then any buyer b_j he is matched to with $\overset{F}{l_{s_i}} \geq l_{b_j} > \overset{T}{l_{s_i}}$ will fire him. This is reasonable as s_i lacks both skills and qualifications for b_j's project. This captures situations when sellers can not falsify such evidence proving a higher level of experience than the real one, but they can withhold such evidence. If one does not restrict reported levels for s_i, $\overset{F}{l_{s_i}} < \overset{T}{l_{s_i}}$, then there does not exist a stable and strategy-proof mechanism (proof in the Appendix 1.9.B).

Buyer Requirement Strategy-proofness. For each $(R_S, l_S, l_B) \in \mathcal{E}$ and each $b_j \in B$, there is no $l'_{b_j} \in \mathbb{R}_+$, such that:

$$\varphi_{b_j}(R_S, l_S, \overset{F}{l'_{b_j}}, l_{-b_j}) \overset{T}{P_{b_j}} \varphi_{b_j}(R_S, l_S, \overset{T}{l_{b_j}}, l_{-b_j}).$$

Second, I look at an axiom concerning the possibility of the sellers benefiting from misreporting their preferences. Let T and F, placed above each $R_{s_i} \in \mathcal{R}_S$, denote a "true" ($\overset{T}{R_{s_i}}$) and "false" ($\overset{F}{R_{s_i}}$) preference relation of the seller, respectively.

Seller Preference Strategy-proofness. For each $(R_S, l_S, l_B) \in \mathcal{E}$ and each $s_i \in S$ there is no $R'_{s_i} \in \mathcal{R}_S$, such that:

$$\varphi_{s_i}(\overset{F}{R'_{s_i}}, R_{-s_i}, l_S, l_B) \overset{T}{P_{s_i}} \varphi_{s_i}(\overset{T}{R_{s_i}}, R_{-s_i}, l_S, l_B).$$

Finally, I present an axiom regarding the joint manipulation of both preferences and the level of experience by the sellers.

Seller Strategy-proofness. For each $(R_S, l_S, l_B) \in \mathcal{E}$ and each $s_i \in S$, there is no pair $l'_{s_i} < l_{s_i}$ and $R'_{s_i} \in \mathcal{R}_S$, such that:

$$\varphi_{s_i}(\overset{F}{R'_{s_i}}, R_{-s_i}, \overset{F}{l'_{s_i}}, l_{-s_i}, l_B) \overset{T}{P_{s_i}} \varphi_{s_i}(\overset{T}{R_{s_i}}, R_{-s_i}, \overset{T}{l_{s_i}}, l_{-s_i}, l_B).$$

The situation when no agent can benefit from misreporting any kind of his private information is considered in the next axiom.

Strategy-proofness. A mechanism φ is strategy-proof if it satisfies buyer requirement strategy-proofness and seller strategy-proofness.

In the next section I define a mechanism that finds the unique stable matching for economies with no ties in experience levels of the sellers. Therefore, it is reasonable to suggest an experience monotonicity requirement. First, that increasing the experience level for one of the sellers does not make any other seller better off and any other buyer worse off.

Seller Experience Monotonicity. Let $(R_S, l_S, l_B) \in \mathcal{E}$. For each $s_i \in S$ if $l'_{s_i} > l_{s_i}$, then for each $s'_i \in S \backslash \{s_i\}$:

$$\varphi_{s'_i}(R_S, l_S, l_B) \; R_{s'_i} \; \varphi_{s'_i}(R_S, l'_{s_i}, l_{-s_i}, l_B),$$

and for each $b_j \in B$:

$$\varphi_{b_j}(R_S, l'_{s_i}, l_{-s_i}, l_B) \; R_{b_j} \; \varphi_{b_j}(R_S, l_S, l_B).$$[2]

Similarly, decreasing the experience level needed for one of the buyers does not make any other buyer better off and any other sellers worse off.

Buyer Experience Monotonicity. Let $(R_S, l_S, l_B) \in \mathcal{E}$. For each $b_j \in B$ if $l'_{b_j} < l_{b_j}$, then for each $b'_j \in B \backslash \{b_j\}$:

$$\varphi_{b'_j}(R_S, l_S, l_B) \; R_{b'_j} \; \varphi_{b'_j}(R_S, l_S, l'_{b_j}, l_{-b_j}),$$

and for each $s_i \in S$:

$$\varphi_{s_i}(R_S, l_S, l'_{b_j}, l_{-b_j}) \; R_{s_i} \; \varphi_{s_i}(R_S, l_S, l_B).$$

[2] Note that R_{b_j} is dependent on the experience levels of the sellers. So, when both sides involve the same seller, the buyer compares the seller with the new experience level and the same seller with the old experience level.

1.5 The mechanism

In this model, the seller- and buyer-optimal stable matchings are the same only when there are no ties. If there are ties, there might not be seller-optimal stable matching.[3] This is a result of the correlation of the buyers' preferences with the sellers' levels of experience—a specific feature of the model.

In the next section I show that SD algorithm with cutoffs finds the unique stable matching when there are no ties. There is a particular algorithm that finds this matching—Serial Dictatorship (SD) algorithm with cutoffs (Algorithm 1).

For each $s_i \in S$, and each $b_j \in B$, s_i is **qualified** for b_j if and only if s_i's level of experience is at least as high as the level of experience needed for b_j's project, $l_{b_j} \leq l_{s_i}$.

For each $s_i \in S$, and each $b_j \in B$, b_j is **acceptable** for s_i if and only if b_j is over the cutoff in the preferences of s_i, $b_j \; P_{s_i} \; s_i$.

Algorithm 1: SD algorithm with cutoffs.

Step 1: Break all ties in preferences of the buyers for the sellers with the same level of experience, if there are any, as follows:

$$s_i \; R_b \; s_j \text{ if and only if } l_{s_i} > l_{s_j}$$

$$\text{or } l_{s_i} = l_{s_j} \text{ and } i < j.$$[4]

Step 2: The seller with the highest level of experience is matched to his most preferred buyer among those he is qualified for. If he does not find any such buyer acceptable, he is not matched.

. . .

Step n: The seller with the lowest level of experience is matched to his most preferred

[3]Consider the following example. Let $S = (s_1, s_2)$ and $B = (b_1)$. Let preferences and levels of experience be as follows:

$l_{b_1} = 1$	$l_{s_1} = 2$	$l_{s_2} = 2$
R_{b_1}	R_{s_1}	R_{s_2}
$\{s_1, s_2\}$	b_1	b_1
b_1	s_1	s_2

So, $\mu_1 : (s_1, b_1), (s_2)$; $\mu_2 : (s_1), (s_2, b_1)$. There are two stable matchings, but neither of them is a seller-optimal stable matching.

[4]I use index as a tie breaker in buyers' preferences for the sellers of the same level of experience. One can randomize over tie-breakers as long as the ties in every buyers' preferences are broken in the same way.

available buyer among those he is qualified for. If he does not find any such buyer acceptable, he is not matched.

The algorithm terminates. Any buyer for whom no seller is qualified is unmatched. Any seller who is not qualified for any of the buyers also is unmatched.

1.6 Results

In a given economy $(R_S, l_S, l_B) \in \mathcal{E}$, where no two sellers have the same experience level, there can be only one stable matching and it coincides with the matching produced by SD algorithm with cutoffs.

Theorem 1. SD algorithm with cutoffs yields a stable matching.

Proof: Let SD algorithm with cutoffs yield a matching $\mu \in \mathcal{M}$ and $s_i \in S$ be assigned $b_j \in B$, $\mu(s_i) = b_j$. Suppose s_i and $b_{j'} \in B \backslash \{b_j\}$ block μ, so it is unstable. Then, since $(s_i, b_{j'})$ blocks μ, $b_{j'} \; P_{s_i} \; b_j$ and s_i must be qualified for $b_{j'}$, $s_i \; P_{b_{j'}} \; \mu(b_{j'})$. If $\mu(b_{j'}) = b_{j'}$, then $b_{j'}$ is available when s_i chooses and SD algorithm with cutoffs must yield $\mu(s_i) = b_{j'}$, but it yields $\mu(s_i) = b_j$. So, this means s_i is not qualified for $b_{j'}$ and is below the cut off in $b_{j'}$s' preferences, otherwise the algorithm would assign $b_{j'}$ to s_i. If $\mu(b_{j'}) = s_{i'}$, where $s_{i'} \in S \backslash \{s_i\}$, then since $s_i \; P_{b_{j'}} \; s_{i'}$, $l_{s_i} > l_{s_{i'}}$. So, regardless of the tie breaks, s_i chooses before $s_{i'}$. So, again $b_{j'}$ is available when s_i chooses. Following the same reasoning, this contradicts $\mu(s_i) = b_j$. Therefore, $(s_i, b_{j'})$ can not block μ and SD algorithm with cutoffs yields a stable matching. So, μ is stable with respect to the tie-broken preferences and also with the respect to the original preferences. □

Proposition 1. For each economy where there are no ties the SD algorithm with cutoffs finds the unique stable matching.

Proof: By contradiction, suppose there are two stable matchings: μ, produced by SD with cutoffs, and μ'. If $\mu \neq \mu'$, let $s_i \in S$ be the highest priority (most experienced) seller such that $\mu(s_i) = b_j$ and $\mu'(s_i) \neq b_j$.[5] Since s_i is the highest priority seller who gets something different under μ', $\forall \; s'_i \in S \backslash \{s_i\}$ with $l_{s'_i} > l_{s_i}$, $\mu(s'_i) = \mu'(s'_i)$. By definition of SD with cutoffs, b_j is the best buyer for s_i among those he is qualified for after the buyers were assigned to higher

[5] Under μ', s_i is either unmatched or matched to another buyer $b'_j \neq b_j$.

priority sellers. So, either $\mu'(b_j) = b_j$ or $\mu'(b_j) = s_i'$, where $\forall\ s_i' \in S \backslash \{s_i\}$ with $l_{s_i'} < l_{s_i}$. Therefore, the pair (s_i, b_j) blocks μ' and there can not be two separate stable matchings. □

Proposition 2. SD algorithm with cutoffs is Pareto Efficient.

Proof: Let SD algorithm with cutoffs yield a matching $\mu \in \mathcal{M}$ and $s_i \in S$ be assigned $b_j \in B$, $\mu(s_i) = b_j$. Let s_k be the seller who chooses in the k-th step. Since s_1 is matched to his most preferred buyer, there is no matching that makes him better off. Thus, if μ' Pareto dominates μ, $\mu'(s_1) = \mu(s_1)$. In step k, since s_k is matched to his most preferred remaining buyer that he is qualified for, the only way to make him better off is to match him to a buyer he is not qualified for or to a buyer already assigned in an earlier step. The former will make the buyer worse off and the latter is not feasible. Thus, $\mu'(s_k) = \mu(s_k)$. Therefore, for each $i = 1, ..., n$, $\mu'(s_i) = \mu(s_i)$, contradicting the assumption that μ' Pareto dominates μ. So, SD algorithm with cutoffs yields a Pareto efficient matching. □

Let SD mechanism with cutoffs choose the allocation defined by the SD algorithm with cutoffs.

Proposition 3. SD mechanism with cutoffs is Seller Preference Strategy-proof.

Proof: For the seller $s_k \in S$, at the k-th step of the SD algorithm with cutoffs, the algorithm assigns him his most preferable buyer from among those who have not been matched in previous steps and for whom s_k is qualified or leaves him unmatched if no such buyer exists. So, s_k has no incentive to lie about his top choice among the remaining buyers, because the only way for s_k to ensure getting his top choice is by reporting his top choice truthfully. Therefore, it is a weakly dominant strategy for each seller to state his true preferences. □

Proposition 4. SD mechanism with cutoffs is Seller Experience Strategy-proof.

Proof: Suppose $s_i \in S$ is assigned $b_j \in B$ at the k-th step of the SD algorithm with cutoffs, when s_i reports $\overset{F}{l}_{s_i} < \overset{T}{l}_{s_i}$ (by assumption, any seller reports only his true level $\overset{T}{l}_{s_i}$ or any level that is lower, than the true one). I will show that s_i is still qualified for b_j and b_j is still available for s_i if he reports $\overset{T}{l}_{s_i}$; meaning s_i is at least as well off as when he reports the truth. First, as s_i is qualified for b_j when he reports untruthfully, we have: $l_{b_j} \leq \overset{F}{l}_{s_i} < \overset{T}{l}_{s_i}$. So, if s_i

reports truthfully, he is qualified for b_j. Second, if s_i reports $\overset{T}{l}_{s_i}$, then he has to be assigned no later than k-th step of the SD algorithm with cutoffs. Since none of the sellers chose b_j in steps before k when s_i reported $\overset{F}{l}_{s_i}$, b_j still remains available when s_i reports $\overset{T}{l}_{s_i}$. Therefore, it is a weakly dominant strategy for each seller to state the true level of experience. □

Proposition 5. SD mechanism with cutoffs is Buyer Requirement Strategy-proof.

Proof: Suppose $b_k \in B$ reports $\overset{F}{l}_{b_k} < \overset{T}{l}_{b_k}$ and is assigned to $s_k \in S$ at the k-th step of the SD algorithm with cutoffs. I will show that if b_k reports $\overset{F}{l}_{b_k}$ he is either as well off as when reporting $\overset{T}{l}_{b_k}$ or strictly worse off, depending on l_{s_k}. First, if b_k reports untruthfully and $\overset{F}{l}_{b_k} < \overset{T}{l}_{b_k} \leq l_{s_k}$, then s_k is qualified for b_k and b_k is assigned to s_k at the k-th step of the algorithm. So, since s_k is qualified for both b_k reporting truthfully or not, b_k is as well off as when he reports $\overset{T}{l}_{b_k}$. Second, if b_k reports untruthfully and $\overset{F}{l}_{b_k} \leq l_{s_k} < \overset{T}{l}_{b_k}$, then s_k is qualified for b_k and b_k is assigned to s_k at the k-th step of the algorithm. Even though s_k is qualified for b_k with $\overset{F}{l}_{b_k}$, but s_k is underqualified for b_k's truthful level of experience needed, $\overset{T}{l}_{b_k}$. Therefore, b_k is even worse off than being unmatched.

Next, suppose b_k reports $\overset{F}{l}_{b_k} > \overset{T}{l}_{b_k}$. I will show that b_k will be still assigned to s_k if b_k reports $\overset{T}{l}_{b_k}$, meaning b_k is as well off as when he reports the truth. First, as s_k is qualified for b_k when b_k reports untruthfully, we have: $l_{s_k} \geq \overset{F}{l}_{b_k} > \overset{T}{l}_{b_k}$. So, if b_k reports truthfully, then s_k is qualified for b_k. Second, if b_k reports $\overset{T}{l}_{b_k}$, then he has to be assigned to s_k also in k-th step of the SD algorithm with cutoffs. Since none of the sellers chose b_k in steps before k when b_k reported $\overset{F}{l}_{b_k}$, b_k still remains available for s_k and is assigned to him when b_k reports $\overset{T}{l}_{b_k}$. Therefore, it is a dominant strategy for each buyer to state the true level of needed experience. □

Proposition 6. SD mechanism with cutoffs is Seller Strategy-proof.

Proof: If some seller $s_i \in S$ prefers to be engaged with some buyer $b_j \in B$ and is not qualified for him ($\overset{T}{l}_{s_i} < l_{b_j}$), then there is nothing s_i can do to get b_j. Regardless of R_{s_i}, s_i can not submit $\overset{F}{l}_{s_i} \geq l_{b_j}$, by the imposed restriction. On the other hand, if $\overset{T}{l}_{s_i} \geq l_{b_j}$, s_i can get b_j only if there is no other $s_i' \in S$ with $l_{s_i'} > \overset{T}{l}_{s_i} \geq l_{b_j}$ and for who b_j is the most preferred buyer among those who are available when s_i' chooses. Otherwise, b_j will be matched with

s_i'. Regardless of R_{s_i}, s_i can not lie to have $\overset{F}{l_{s_i}} \geq l_{s_i'}$. By Proposition 4, there is no gain from reporting lower level of experience for sellers. Therefore, there is no possible misreport $\overset{F}{l_{s_i}}$ and $\overset{F}{R_{s_i}}$ for sellers to be better off. □

Proposition 7. SD mechanism with cutoffs is Strategy-proof.

Proof: By Proposition 5 and 6, SD mechanism with cutoffs is Strategy-proof. □

Since SD mechanism with cutoffs finds a unique stable matching, one might expect SD mechanism with cutoffs to satisfy an experience monotonicity requirement. However, SD mechanism with cutoffs is not seller experience monotonic. In addition, when there is a single buyer and $|S| \geq 1$, it is seller experience monotonic (proof in the Appendix 1.9.C).

Proposition 8. When $|S| \geq 1$ and $|B| \geq 2$, SD mechanism with cutoffs is not Seller Experience Monotonic.

Proof: Let $S = (s_1)$ and $B = (b_1, b_2)$. Preferences and levels of experience are such that:

$l_{b_1} = 1$	$l_{b_2} = 2$	$l_{s_1} = 1$
R_{b_1}	R_{b_2}	R_{s_1}
s_1	s_1	b_2
b_1	b_2	b_1
		s_1

At (R_S, l_S, l_B) the SD algorithm with cutoffs yields the matching μ: $(s_1,\ b_1)$, b_2.

Suppose now $l'_{s_1} > l_{s_1}$. Since s_1 now is qualified for b_2 and b_2 is available, at (R_S, l'_{s_1}, l_B) the SD algorithm with cutoffs yields μ': $(s_1,\ b_2)$, b_1. So, b_1 is matched to s_1 under μ and unmatched under μ'. Therefore, increasing the experience level for s_1 makes b_1 worse off.

The argument generalizes to the case with $|S| \geq 2$ and $|B| \geq 2$ (more cases in the Appendix 1.9.D). □

Proposition 9. SD mechanism with cutoffs is Buyer Experience Monotonic.

Proof: Suppose $b_k \in B$ reports l_{b_k} and $s_k \in S$ is assigned his most preferred among the remaining buyer b_k in n-th step of the SD algorithm with cutoffs. I will show that if l_{b_k}

changes to $l'_{b_k} < l_{b_k}$, other sellers (buyers) can be either as well off as or better off (worse off) than when he reports l_{b_k}.

First, if $l'_{b_k} < l_{b_k} \leq l_{s_k}$, then s_k chooses at the n-th step of the SD algorithm with cutoffs as before and is assigned b_k. The order in which sellers choose buyers stays same, regardless of the experience level b_k reports between l_{b_k} and l'_{b_k}. So, in this case all other buyers and sellers are as well off as before.

Second, if $l_{s_k} < l_{b_k}$, then s_k is not qualified for b_k in this case. Let $\mu \in \mathcal{M}$ be the matching at $(R_S, l_S, l_{b_k}, l_{-b_k})$. Suppose $\mu(s_k) = b_j$, and that s_k is assigned b_j at the n-th step of the SD algorithm with cutoffs. Suppose that $b_k \ P_{s_k} \ b_j$. When $l'_{b_k} \leq l_{s_k} < l_{b_k}$, let $\mu' \in \mathcal{M}$ be the matching at $(R_S, l_S, l'_{b_k}, l_{-b_k})$. So, since the order of sellers does not change s_k still chooses at the n-th step of the SD algorithm with cutoffs as before, but $\mu'(s_k) = b_k$. Furthermore, b_j becomes available for any other seller $s_j \in S$ who chooses after the n-th step. If $b_j \ P_{s_j} \ \mu(s_j)$ and $l_{s_j} \geq l_{b_j}$, then $\mu'(s_j) = b_j$, s_j is assigned b_j in some step after n-th step, since b_j is available now. Furthermore, regardless of who the seller b_j is assigned to in steps after n-th, b_j is always worse off. Therefore, decreasing the experience level needed for b_k makes b_j worse off and s_j better off.

Moreover, the buyer who s_j was matched to under μ is now available after the step when s_j is assigned b_j for some other seller; this buyer is worse off now. By the same argument, they can be matched under μ' in some step after the step when s_j is assigned b_j; this seller is better off now. Therefore, each buyer who is displaced is worse off and each seller who receives the more preferred buyer is better off, by the same argument. This chain process continues until there is the last buyer who is either matched to a less preferred seller or becomes unmatched under μ'; and there is the last seller who was either unmatched under μ and is matched under μ' or who was matched under μ and is matched to a more preferred buyer under μ'. □

1.7 Obvious Strategy-proofness

One of the important considerations is the incentives given to agents to reveal either their preferences or any other private information truthfully. Although I have proved that SD mechanism with cutoffs is strategy-proof (Proposition 8), it does not mean that agents, in particular sellers, will report their experience levels truthfully in practice. In general, there is experimental evidence (Kagel et al. (1987), Charness and Levin (2009), Martínez-Marquina et al. (2019a)) of the inability of some agents to engage in contingent reasoning, resulting in untruthful information revelation, even though the mechanism is Strategy-proof (SP) and truthful reporting is a weakly dominant strategy. Therefore, I focus now on much more demanding concept of obvious strategy-proofness (OSP) (Li (2017)). To do this I first define an obviously dominant strategy (ODS) of an extensive form game that, in general, is a strategy such that even the best outcome under any possible deviation, at any information set where both strategies first diverge, is no better than the worst outcome under the ODS. An ODS is a strategy that can be identified as optimal (dominant) by an agent who can not perform state contingent reasoning. If a mechanism has an equilibrium in ODSs, it is considered to be obviously strategy-proof.

In addition, OSP is a property of the extensive game form, unlike SP. A mechanism is OSP-implementable if there exists some extensive form such that for every profile of preferences and, in my model, for every profile of experience levels, its outcome has an equilibrium in obviously dominant strategies and coincides with the outcome as if the mechanism was applied.

Suppose buyers' true levels of experience are publicly known information. Only each seller knows his preferences and experience level and buyers' preferences still depend on the sellers' private information.

I consider only game forms with perfect recall and finite depth (Li (2017), Pycia and Troyan (2019)). Let the set of all such games be represented by $\mathcal{G}$. Let $\Gamma \in \mathcal{G}$ denote an extensive game form with outcomes in the set of all possible matchings $\mathcal{M}$; meaning each terminal history ω results in some outcome $\mu(\omega) \in \mathcal{M}$.

For each $s_i \in S$, let $\mathcal{I}_{s_i}$ denote the information sets for s_i, with representative element $I_{s_i} \in \mathcal{I}_{s_i}$. Let $\mathcal{I}$ denote all the information sets for all sellers. Let H be the set of all histories, with representative element h. Let $\succ$ denote the precedence order over histories (Li (2017)). For any $h \in I_{s_i}$ and $h' \in I'_{s_i}$ if $h \succ h'$, then $I_{s_i} \succ I'_{s_i}$.

For each $s_i \in S$, let $\mathcal{A}_{s_i}$ denote the set of all possible actions for s_i. Let $A(I_{s_i}) \in \mathcal{A}_{s_i}$ denote the set of all actions available to s_i at $I_{s_i} \in \mathcal{I}_{s_i}$. Let $\mathcal{A}$ denote all possible actions available for all of the sellers.

A strategy is $\sigma_{s_i} : \mathcal{I}_{s_i} \to \mathcal{A}_{s_i}$, where $\mathcal{A}_{s_i} = \cup_{I_{s_i} \in \mathcal{I}_{s_i}} A(I_{s_i})$ such that for each $I_{s_i} \in \mathcal{I}_{s_i}$, $\sigma_{s_i}(I_{s_i}) \in A(I_{s_i})$. In other words, σ_{s_i} is a complete contingent plan, it chooses, for each information set I_{s_i}, a particular action $a \in A(I_{s_i})$ for any s_i. A strategy profile $\sigma = (\sigma_{s_i})_{s_i \in S}$ specifies a strategy for each $s_i \in S$.

The earliest points of departure $\alpha(\sigma_{s_i}, \sigma'_{s_i})$ are the earliest information sets I_{s_i} at which

σ_{s_i} and σ'_{s_i} diverge ($\sigma_{s_i}(I_{s_i}) \neq \sigma'_{s_i}(I_{s_i})$) (Li (2017)). In other words, $\alpha(\sigma_{s_i}, \sigma'_{s_i})$ are the earliest information sets, where these two strategies choose different actions. Note that there can be no earlier point of departure under perfect recall; meaning that at every previous information set σ_{s_i} and σ'_{s_i} chose the same action.

Next, a strategy σ_{s_i} is an ODS if for any alternative strategy σ'_{s_i} the best possible outcome under σ'_{s_i} is no better than the worst under σ_{s_i}, conditional on reaching any earliest point of departure $\alpha(\sigma_{s_i}, \sigma'_{s_i})$. Now I define the ODS formally as follows.

Definition 1. Obviously Dominant Strategy (ODS). For each $(R_S, l_S, l_B) \in \mathcal{E}$ and each $s_i \in S$, given Γ, σ_{s_i} is obviously dominant strategy (ODS) if, for each σ'_{s_i} and for each $I_{s_i} \in \alpha(\sigma_{s_i}, \sigma'_{s_i})$, the R_{s_i}-worst outcome under any terminal history following σ_{s_i} is at least as good as the R_{s_i}-best outcome under following σ'_{s_i}.

Definition 2. Obvious Strategy-proofness (OSP). Let $\phi : \mathcal{E} \to \mathcal{M}$ be a mechanism. Suppose $\Gamma \in \mathcal{G}$ is such that for each $(R_S, l_S, l_B) \in \mathcal{E}$, in Γ every $s_i \in S$ has an obviously dominant strategy $\sigma^{ODS}_{s_i}$. Let $\mu^{ODS}(R_S, l_S, l_B)$ be the resulting outcome. If for each $(R_S, l_S, l_B) \in \mathcal{E}$, $\mu^{ODS}(R_S, l_S, l_B) = \phi(R_S, l_S, l_B)$, then Γ implements ϕ in ODS. If there is $\Gamma \in \mathcal{G}$ that implements ϕ in ODS, then ϕ is OSP.

Suppose $G^E : \mathbb{R}^{\mathbb{S}}_+ \to \mathcal{G}$ is a mapping from profiles of experience levels to extensive game forms; meaning G^E says which game form from the set of all extensive game forms $\mathcal{G}$ to use if the sellers' experience profile is $l_S \in \mathbb{R}^{\mathbb{S}}_+$. If for every profile of experience levels $l_S \in \mathbb{R}^{\mathbb{S}}_+$, $G^E(l_S)$ is such that every seller has an ODS, I say that G^E has an equilibrium in obviously dominant strategies.

Definition 3. Obvious Preference Strategy-proofness (OPSP). Let $\phi : \mathcal{E} \to \mathcal{M}$ be a mechanism. Suppose $G^E : \mathbb{R}^{\mathbb{S}}_+ \to \mathcal{G}$ is such that for each $(R_S, l_S, l_B) \in \mathcal{E}$, in $G^E(l_S)$ every $s_i \in S$ has an obviously dominant strategy $\sigma^{ODS}_{s_i}$. Let $\mu^{ODS}(R_S, l_S, l_B)$ be the resulting outcome. If for each $(R_S, l_S, l_B) \in \mathcal{E}$, $\mu^{ODS}(R_S, l_S, l_B) = \phi(R_S, l_S, l_B)$, then G^E implements ϕ in ODS given experience levels. If there is G^E that implements ϕ in ODS given experience levels, then ϕ is OPSP.

In other words, G^E OPSP-implements some mechanism ϕ if for any economy $e = (R_S, l_S, l_B) \in \mathcal{E}$, the ODS equilibrium outcome of $G^E(l_S)$ coincides with $\phi(e)$.

Next, suppose $G^P : \mathcal{R}^{\mathcal{S}}_+ \to \mathcal{G}$ is a mapping from profiles of preferences to extensive game forms; meaning G^P says which game form from the set of all extensive game forms $\mathcal{G}$ to use if the sellers' preference profile is $R_S \in \mathcal{R}_+^{\mathcal{S}}$. If for every profile of preferences $R_S \in \mathcal{R}_+^{\mathcal{S}}$, $G^P(R_S)$ is such that every seller has an ODS, I say that G^P has an equilibrium in obviously dominant strategies.

Definition 4. Obvious Experience Strategy-proofness (OESP). Let $\phi : \mathcal{E} \to \mathcal{M}$ be a mechanism. Suppose $G^P : \mathcal{R}_+^{\mathcal{S}} \to \mathcal{G}$ is such that for each $(R_S, l_S, l_B) \in \mathcal{E}$, in $G^P(R_S)$ every $s_i \in S$ has an obviously dominant strategy $\sigma_{s_i}^{ODS}$. Let $\mu^{ODS}(R_S, l_S, l_B)$ be the resulting outcome. If for each $(R_S, l_S, l_B) \in \mathcal{E}$, $\mu^{ODS}(R_S, l_S, l_B) = \phi(R_S, l_S, l_B)$, then G^P implements ϕ in ODS given preferences. If there is G^P that implements ϕ in ODS given preferences, then ϕ is OESP.

G^P OESP-implements some mechanism ϕ if for any $e = (R_S, l_S, l_B) \in \mathcal{E}$, the ODS equilibrium outcome of $G^P(R_S)$ coincides with $\phi(e)$.

In order for ϕ to be OSP-implementable, it has to be both OPSP- and OESP-implementable.

Proposition 10. Suppose ϕ is OSP, then ϕ is OPSP and OESP.

Proof: Suppose Γ is an extensive game form that implements ϕ in ODS. Then let $G^E : \mathbb{R}_+^{\mathbb{S}} \to \mathcal{G}$ be such that, for each $l_S \in \mathbb{R}_+^{\mathbb{S}}$, $G^E(l_S) = \Gamma$ and $G^P : \mathcal{R}_+^{\mathcal{S}} \to \mathcal{G}$ be such that, for each $R_S \in \mathcal{R}_+^{\mathcal{S}}$, $G^P(R_S) = \Gamma$. Since Γ OSP-implements ϕ, G^E OPSP-implements ϕ and G^P OESP-implements ϕ. Therefore, ϕ is OPSP- and OESP-implementable. □

I show below that SD mechanism with cutoffs is OPSP-implementable. In the game form that OPSP-implements SD mechanism with cutoffs, sellers select buyers in decreasing order of their experience level, each seller observes what buyers were taken by other sellers.

Let $G^E : \mathbb{R}_+^{\mathbb{S}} \to \mathcal{G}$ be such that, for each $l_S \in \mathbb{R}_+^{\mathbb{S}}$, $G^E(l_S)$ be defined as follows. First, ties are broken for the sellers with the same level of experience if there are any. Second, all sellers are put in order corresponding to their levels of experience from the highest to the lowest and are informed of the order (but not about the experiences). Next, sellers play sequentially, starting from the seller with the highest experience to the lowest. When a seller takes his turn, he is shown the updated list of buyers that have not yet been matched, he picks up to one buyer among those he is qualified for if one exists. Each $s_i \in S$ observe the updated availability status of each b_j before he chooses.

G^E OPSP-implements SD with cutoffs if for each $e = (R_S, l_S, l_B) \in \mathcal{E}$, the ODS equilibrium outcome of G^E coincides with SD mechanism with cutoffs.

Proposition 11. G^E OPSP-implements SD with cutoffs.

Proof: I show that each player has an ODS to truthfully pick the best available buyer whom he is qualified for; meaning the best outcome any seller can get under any deviation strategy is no better than the worst outcome under the truthful strategy.

Consider seller $s_k \in S$ with the level of experience l_{s_k}. Let the set of all available buyers s_k is qualified for be $B' \subseteq B$. At each information set where s_k is called to play, s_k faces some $B' \subseteq B$. His actions are: $\forall b_j \in B'$, he can either pick b_j or remain unmatched. So, a strategy σ_{s_k} maps each $B' \subseteq B$ to $B' \cup \{s_k\}$, where s_k denotes being unmatched. Playing σ_{s_k} that maximizes R_{s_k} is an ODS for s_k, because s_k picks his most preferable buyer from each $B' \subseteq B$ with the regard to his R_{s_k} and gets it. Any other σ'_{s_k} either yields the same buyer at every $B' \subseteq B$ or worse. Since s_k is only called to play at most once, σ_{s_k} obviously dominates any such σ'_{s_k}. □

Next, consider OESP-implementability of the SD mechanism with cutoffs. G^P OESP-implements SD mechanism with cutoffs if for each $e = (R_S, l_S, l_B) \in \mathcal{E}$, the ODS equilibrium outcome of G^P coincides with SD mechanism with cutoffs.

Proposition 12. There is no G^P that OESP-implements SD with cutoffs.

Proof: Without loss of generality, I prove this for the setting with one buyer and two sellers. The argument generalizes to the setting with more than one buyer and two sellers.

Suppose G^P OESP-implements SD with cutoffs. Consider an information set $I_{s_1} \in \mathcal{I}$ such that in any previous information sets nature plays. Let $s_1 \in S$ be the seller who moves in I_{s_1}. Let $s_2 \in S$ be the other seller. Let $b_1 \in B$ be the buyer.

Let $R_S \in \mathcal{R}^{\mathcal{S}}$ be the profile of preferences of the sellers such that both sellers prefer to be matched to the buyer, than to be unmatched. Let $l_{b_1} = 0$.

Let $\Gamma = G^P(R_S)$. G^P OESP-implements SD with cutoffs implies that $\forall l_S$, there is an ODS equilibrium outcome of Γ, μ^{ODS} such that $\mu^{ODS} = \phi(R_S, l_S, l_B)$.

Without loss of generality, by the pruning principle for obvious dominance (Li (2017), Bade and Gonczarowski (2017), Pycia and Troyan (2019), Troyan (2019)), I assume that Γ is pruned; meaning $\forall I_{s_i} \in \mathcal{I}_{s_i}$ every $a \in A(I_{s_i})$ is OD for some type l_{s_i}, $\forall l_{s_i} \in l_S$.

There are only two possible efficient matchings, since b_1 is acceptable for both sellers. Let

μ^1 be such that s_1 is matched to b_1 and s_2 is unmatched: $\mu^1(s_1) = b_1$ and $\mu^1(s_2) = s_2$. Let μ^2 be such that s_1 is unmatched and s_2 is matched to b_1: $\mu^2(s_1) = s_1$ and $\mu^2(s_2) = b_1$. If $l_{s_1} > l_{s_2}$, then $\phi(R_S, l_S, l_B) = \mu^1$; and if $l_{s_1} < l_{s_2}$, then $\phi(R_S, l_S, l_B) = \mu^2$.

For each information set I can map each action of the seller playing in it to the set of all matchings at some terminal node following that action. Let $A(I_{s_1}) \subseteq \mathcal{A}_{s_1}$ denote the set of all actions available to s_1 at $I_{s_1} \in \mathcal{I}$. For all $a \in A(I_{s_1}) \in \mathcal{A}_{s_1}$, let $\mathcal{M}(a)$ be the set of all matchings at some terminal node following a. Since Γ OESP-implements the SD mechanism with cutoffs, $\exists a^* \in A(I_{s_1})$ such that $\forall a \in A(I_{s_1})$, a^* obviously dominates a. Then $\forall \mu \in \mathcal{M}(a^*)$ and $\forall \mu' \in \mathcal{M}(a)$, $\mu(s_1)\ R_{s_1}\ \mu'(s_1)$ and $\phi(R_S, l_S, l_B) \in \mathcal{M}(a^*)$.

Let $l_{s_1} = 2$ and $l_{s_2} = 3$. Let $\gamma = \phi(R_S, l_S, l_B)$. Then we have $\gamma = \mu^2$, and $\gamma \in \mathcal{M}(a^*)$. So, $\mu^2 \in \mathcal{M}(a^*)$. Let us look at any $a \in A(I_{s_1}) \backslash \{a^*\}$. Suppose $l^a_{s_1}$ is the type that player $s_1 \in S$ would have to have for a to be his obviously dominant strategy. It is possible that s_1 has type $l^a_{s_1}$ when s_2 has a lower type $l'_{s_2} < l^a_{s_1}$. Then s_1 is matched and $\phi(R_S, l^a_{s_1}, l'_{s_2}, l_B) = \mu^1$. So, $\mu^1 \in \mathcal{M}(a)$; contradicting a^* is ODS.

So, there is no $I_{s_i} \in \mathcal{I}$ where any player has more, than one action a, $\forall a \in A(I_{s_i}) \in \mathcal{A}_{s_i}$. Therefore, it has to be that the game form Γ is such that at every $I_{s_i} \in \mathcal{I}$ every $s_i \in S$ has just one acton $a \in A(I_{s_i})$. Regardless of l_S, the same outcome occurs, which means G^P can not implement SD with cutoffs. Therefore, $\nexists G^P$ that OESP-implements SD with cutoffs. □

Corollary 1. SD mechanism with cutoffs is not OSP-implementable.

Proof: Recall that OSP-implementability requires both OPSP- and OESP-implementability. Since SD mechanism with cutoffs is not OESP-implementable, by Proposition 11, it is not OSP-implementable. □

1.8 Conclusion

Hidden private information of some agents can be crucial for the others in establishing their preferences. This motivated my study of the information revelation problem where agents reveal more than just preferences and the others' preferences depend on such private information. I use a simple one-to-one matching model with sellers on the one side of the market and buyers on the other side. In my model, buyers' preferences depend on private information of the sellers, in particular, their levels of experience and qualifications.

I show the existence of the strategy-proof mechanism that returns stable and, therefore, individually rational and Pareto efficient matching in this setting. The Serial Dictatorship algorithm with cutoffs defines this mechanism. I prove that the SD mechanism with cutoffs is not OSP-implementable, a solution concept introduced recently by Li (2017). I model what it means for a mechanism to be obviously experience and preference strategy-proof (OESP and OPSP). It appears that despite SD with cutoffs being OPSP, it is not OESP and, therefore, not OSP.

This research mainly focuses on introducing private information other than the preferences—levels of experience—into the two-sided matching model and making one sides' preferences completely dependent on this information of the other side in order to analyze agent's incentives to misreport their private information. A natural extension from this work would be to expand from one-dimensional to multidimensional private information model. In other words, to introduce additional types of private information into the model, such as: proximity to work, tastes and personal views, availability during the workday, etc. It can be challenging to incorporate expanded private information into this model and still get the same results for the SD with cutoffs mechanism. One option is to find and propose another mechanism that can satisfy the listed properties above or at least strategy-proofness (obvious strategy-proofness). Another one is to introduce some general utility function, sellers' cost of projects, buyers' budget constraints, etc. For example, Klumpp (2009) considers the problem of assigning sellers and buyers into stable matches, considering their distance in some attribute space. They locate agents along a line and decrease the match surplus function in the distance between partners in order to investigate the structure of stable assignments under both non-transferable utility and transferable utility. One can try to apply their recursive algorithm to my model with levels of experience as private information or additionally introduced geographical proximity.

It is also possible to expand the model by making both sides' preferences dependent on the private information of the other side; meaning not only buyers' preferences should depend on the levels of experience of the sellers, but also sellers' preferences should depend on some relevant private information of the buyers. For instance, geographical location, clients' availability and rating, etc. This can bring significant changes to the model itself and can also require to remodel the design of preference structuring, etc. This kind of possible extensions would allow the model and the economic setting to become even more realistic.

1.9 Appendices

1.9.A Example showing that stability does not imply Pareto efficiency.

Let $S = (s_1, s_2)$ and $B = (b_1, b_2)$. Preferences and levels of experience are such that:

$l_{b_1} = 1$	$l_{b_2} = 1$	$l_{s_1} = 2$	$l_{s_2} = 2$
R_{b_1}	R_{b_2}	R_{s_1}	R_{s_2}
$\{s_1, s_2\}$	$\{s_1, s_2\}$	b_1	b_2
b_1	b_2	b_2	b_1
		s_1	s_2

There are two possible matchings (considering only those where all agents are matched):

$\mu_1 : (s_1, b_1), (s_2, b_2)$ is stable, individually rational, and Pareto efficient;

$\mu_2 : (s_1, b_2), (s_2, b_1)$ is stable, individually rational, but not Pareto efficient, because b_1 and b_2 can be better off (μ_1) without making anyone worse off. Therefore, μ_1 Pareto dominates μ_2.

So, stability implies only Individual Rationality, but does not imply Pareto Efficiency, if there are ties. □

1.9.B Incompatibility of stability and strategy-proofness when sellers can overreport their experience levels.

Proposition 13. If sellers can overreport their experience levels, then there is no stable and strategy-proof mechanism.

Proof: Without loss of generality, I prove this for the setting with one buyer and two sellers. The argument generalizes to the setting with more than one buyer and two sellers.

Let $S = (s_1, s_2)$ with $l_{s_1} < l_{s_2}$. Let $B = (b_1)$ with l_{b_1}. Let both sellers be qualified for b_1, $l_{b_1} \leq l_{s_1} < l_{s_2}$. Let $b_1 \; P_{s_1} \; s_1$ and $b_1 \; P_{s_2} \; s_2$. By Proposition 1, there is at most one (unique) stable matching, $\mu : \mu(s_1) = s_1$ and $\mu(s_2) = b_1$.

Suppose s_1 can misreport: $l'_{s_1} > l_{s_2}$. In this case the new economy has a unique stable matching as well, $\mu' : \mu'(s_1) = b_1$ and $\mu'(s_2) = s_2$. Since I do not assume that sellers have an upper bound in the levels of experience corresponding to the true level of experience, sellers can not be caught in a lie ex post and can not be fired for it (e.g. for falsifying documents).

As a result, s_1 is better off by misreporting l'_{s_1}. Moreover, both s_2 and b_1 strictly prefer to be matched to each other under the truthful information revelation; meaning (s_2, b_1) block μ'. Furthermore, since s_1 is qualified for b_1, s_1 would not be fired ex post even if the true l_{s_i} is verifiable ex post. So, there is no stable and strategy-proof mechanism if reported levels for s_i are not restricted to be no greater, than the truth: $\overset{F}{l_{s_i}} < \overset{T}{l_{s_i}}$. □

1.9.C Cases when SD with cutoffs is seller experience monotonic.

Proposition 14. When there is a single buyer and $|S| \geq 1$, SD with cutoffs is seller experience monotonic.

Proof: Let $S = (s_1)$ with l_{s_1}. Let $B = (b_1)$ with $l_{b_1} \leq l_{s_1}$. Let $b_1 \; P_{s_1} \; s_1$. By applying SD with cutoffs, $\mu : \mu(s_1) = b_1$.

Suppose l_{s_1} increases, $l'_{s_1} > l_{s_1}$, then $\mu' : \mu'(s_1) = b_1$. So, regardless of any increase in s_1's experience level, b_1 is never worse off.

Let $S = (s_1, s_2)$ with $l_{s_1} < l_{s_2}$. Let $B = (b_1)$ with l_{b_1}. Let both sellers be qualified for the buyer, $l_{b_1} \leq l_{s_1} < l_{s_2}$. Let $b_1 \; P_{s_1} \; s_1$ and $b_1 \; P_{s_2} \; s_2$. By applying SD with cutoffs, $\mu : \mu(s_2) = b_1$ and $\mu(s_1) = s_1$.

Suppose l_{s_1} increases, $l'_{s_1} \geq l_{s_2}$, s_1 chooses first, then s_2 chooses and $\mu' : \mu'(s_1) = b_1$ and $\mu'(s_2) = s_2$. So, regardless of any increase in sellers experience levels, b_1 is never worse off. Furthermore, any other seller is never better off. The argument generalizes to the case with one buyer and $|S| > 2$.

□

1.9.D Cases when SD with cutoffs is not seller experience monotonic.

Case 1: Let $S = (s_1, s_2)$ and $B = (b_1, b_2)$. Preferences and levels of experience are such that:

$l_{b_1} = 2$	$l_{b_2} = 2$	$l_{s_1} = 1$	$l_{s_2} = 2$
R_{b_1}	R_{b_2}	R_{s_1}	R_{s_2}
s_2	s_2	b_1	b_1
s_1	s_1	b_2	s_2
b_1	b_2	s_1	

At (R_S, l_S, l_B) the SD algorithm with cutoffs yields the matching μ:(s_2, b_1), (s_1, b_2).

Suppose now $l'_{s_1} > l_{s_2}$. In this cases s_1 choses first. Since s_1 is qualified for b_1 now and b_1 is available, at $(R_S, l'_{s_1}, l_{-s_1}, l_B)$ the SD algorithm with cutoffs yields $\mu'(s_1) = b_1$.

In the next step s_2 chooses, but his top choice is unavailable now, so s_2 prefers to stay unmatched.

So, the SD algorithm with cutoffs yields the matching μ': (s_1, b_1), s_2, b_2. Therefore, increasing the experience level for one of the sellers (s_1) can make some buyer worse off (b_2).

Case 2: Let $S = (s_1, s_2, s_3)$ and $B = (b_1, b_2)$. Preferences and levels of experience are such that:

$l_{b_1} = 1$	$l_{b_2} = 2$	$l_{s_1} = 2$	$l_{s_2} = 3$	$l_{s_3} = 4$
R_{b_1}	R_{b_2}	R_{s_1}	R_{s_2}	R_{s_3}
s_3	s_3	b_2	b_2	b_2
s_2	s_2	b_1	b_1	s_3
s_1	s_1	s_1	s_2	
b_1	b_2			

At (R_S, l_S, l_B) the SD algorithm with cutoffs yields the matching μ: (s_3, b_2), (s_2, b_1), s_1.

Suppose now $l'_{s_2} > l_{s_3}$. In this cases s_2 choses first. Since s_2 is qualified for b_2 and b_2 is available, at $(R_S, l'_{s_2}, l_{-s_2}, l_B)$ the SD algorithm with cutoffs yields $\mu'(s_2) = b_2$.

In the next step s_3 chooses, but his top choice is unavailable now, so s_3 prefers to stay unmatched. Finally s_1 chooses. Since b_1 is available now and s_1 is qualified for b_1, $\mu'(s_1) = b_1$.

So, the SD algorithm with cutoffs yields the matching μ': (s_2, b_2), (s_1, b_1), s_3. Therefore, increasing the experience level for one of the sellers (s_2) can make some sellers better off (s_1), and some buyers worse off (b_1).

The argument generalizes to the case with $|S| \geq 4$ and $|B| \geq 2$. □

Chapter 2

Obviously Strategy-proof Mechanism Design With Rich Private Information

Abstract

I consider settings with rich private information—an agent's type may include private information other than just his preferences. In such settings, I identify a necessary condition for obviously strategy-proof implementation of social choice rules. I consider applications to strict preferences, matching and object allocation.

Keywords: Obvious strategy-proofness, Mechanism design.

2.1 Introduction

If a social choice rule elicits private information in a way that no agent ever has any incentive to misreport, then it is strategy-proof.[1] However, in practice it is not enough to make agents understand the rule for them to be truthful (Kagel et al., 1987; Charness and Levin, 2009; Martínez-Marquina et al., 2019a). Therefore, I focus on "obvious dominance", which is a much stronger property than weak dominance. An obviously dominant strategy is one that can be identified as dominant even by agents who can not engage in contingent reasoning. This means that the worst outcome from the obviously dominant strategy is at least as good as the best outcome from any deviation (Li, 2017). A social choice rule is obviously strategy-proof (OSP) if there exists some extensive game form such that, for every profile of preferences, there is an equilibrium in obviously dominant strategies and the outcome coincides with the outcome of the rule. One of the general messages of this paper is that obvious strategy-proofness is excessively restrictive in some situations where some private information beyond preferences needs to be revealed.

In my model, there is a finite set of agents and each agent has a type, which is his private information. I present a setting with "rich" type information, where each agent's type includes non-preference information in addition to his preferences over outcomes. A social choice rule maps each profile of types to an outcome.

My goal is to identify a necessary condition for a social choice rule to be OSP in such "rich" type environments. I define a social choice rule as being *invariant (to own non-preference information)* if whenever it responds to an agent's non-preference information, while his preferences are the same, the agent is indifferent between the outcomes. In other words, the agent is not affected by his own non-preference information that the social choice rule elicits. I prove that every OSP social choice rule is invariant.

Consider the following example of the extensive form game that is not OSP, but does implement a strategy-proof rule. I describe a natural setting where agents can only underreport their seniority levels. In this example, for a specific modification of serial dictatorship, such underreporting does not seem to be a plausible behaviour. Yet, OSP requires a rule to protect against it. Suppose there is a set of workers and each worker has a type. The type includes preferences over tasks as well as a seniority level (a worker can be "senior" or "junior"). Furthermore, suppose that the seniority level is ex-post verifiable and overreporting carries a penalty. In other words, if, after the fact, the reported type of some worker is higher than his true type, the mechanism designer is able to impose a large penalty (e.g. firing this

[1] A social choice rule is strategy-proof if it is a weakly dominant strategy for each agent to reveal his private information truthfully in the normal form game induced by the rule.

worker).[2] This captures the situation that the worker can not falsify his seniority level, but can withhold (Hurwicz et al., 1995).[3,4]

Consider a modified serial dictatorship, where seniors pick first and then juniors. Within each type group, ties are broken the same way. First, each worker reports his type. Second, break ties, so the senior worker with better tie breaker moves first and picks his most preferred task. Next, the senior worker with the second-best tie breaker picks his most preferred task from those that are left and so on. Juniors pick in the same fashion after the last senior picks. In this case, no worker has an incentive to misreport non-preference private information—seniority level—in addition to the preferences. As a result, this social choice rule is strategy-proof; meaning it is a weakly dominant strategy for workers to report their true type (Halushka, 2020).

However, according to the Theorem 1, this social choice rule can not be implemented in the obviously strategy-proof way. The main reason is that the seniority level affects the task the worker gets, so the rule is not invariant. To see this more concretely, consider an extensive game form, where workers first, simultaneously state their seniority and then pick tasks sequentially. Under truthful reporting, this extensive form selects the same outcome (allocation of tasks) as the above modified serial dictatorship. By existing results (Li, 2017), a standard serial dictatorship is OSP, but in my case the ordering depends on the reports made. So, this game does not have an equilibrium in obviously dominant strategies.[5] In fact, I show that *there is no extensive game form that implements this rule in obviously dominant strategies*. However, it is a very strange behaviour for an agent not to play truthfully; meaning it would be highly implausible for an agent to undermine and move himself lower in the queue by lying and reporting a lower seniority level.[6] So, obvious strategy-proofness is arguably too strong a criterion as it requires a rule to be immune to even such implausible

[2]Overreporting is obviously dominated in this case, since the worker is guaranteed to be fired, while reporting at or under the true seniority level, the worker will not be fired. So, the worst outcome that can happen if the worker reports the seniority less than or equal to his true one is better than being fired, which is the best outcome that can happen if he overreports.

[3]Ex-post verifiability is similar to the assumption of Hurwicz et al. (1995) that endowments can be manipulated by agents downwards, but not upwards, as it may be required to "place the claimed endowments on the table". Therefore, an agent's strategy domain is limited by his (true) endowment.

[4]In this case, for example, if you are senior you can report being senior or junior, but if you are junior you can report only being junior.

[5]For example, consider the second worker in the tie breaker, when he is a senior (or any other agent after him). The worst outcome he gets if he reports truthfully is his second-best choice (in case the first worker reports that he is a senior and has the same top choice). The best outcome he gets if he reports that he is a junior is his top choice (in case the first worker has different top choice). So, it is not an obviously dominant strategy for the second worker to report truthfully (Halushka, 2020).

[6]Suppose there are three workers. If the second worker is a senior, he can not possibly gain by reporting himself to be a junior. However, he can harm himself, since it could move him below the third worker. In this setting, this is highly implausible behaviour.

behaviour.[7]

I consider a few applications, starting with the case of agents having strict preferences over the outcomes. I show that it is not possible to elicit any private information other than preferences in this setting. Furthermore, invariance becomes a simpler property, preference-onlyness: a social choice rule ignores private information other than preferences. Next, I present an application to matching, which is a special case of Halushka (2020). Lastly, I consider a class of sequential dictatorship rules for object allocation.

2.2 Related Literature

There is experimental evidence (Hassidim et al., 2017; Shorrer and Sóvágó, 2018; Rees-Jones and Skowronek, 2018) that some agents are limited in their ability to engage in contingent reasoning. This results in untruthful information revelation, even when a social choice rule is strategy-proof. If a social choice rule is obviously strategy-proof, then it can be implemented in a way that even agents with such cognitive limitations can recognize their strategies as weakly dominant. Li (2017) provides experimental evidence that dynamic implementation of Serial Dictatorship mechanisms leads to much higher truth-telling rates than a static implementation even though the latter is strategy-proof.

There is a robust literature on obvious strategy-proofness, which my paper contributes to (Mackenzie, 2020; Li and Dworczak, 2020; Mandal and Roy, 2022). Ashlagi and Gonczarowski (2018) show that there does not exist an OSP social choice rule that returns a stable matching. In line with their results, I also show that no stable social choice rule is OSP even for very restricted preferences when there is non-preference private information (Section 2.6.2). Pycia and Troyan (2019) characterize simple social choice rules in general social choice environments both with and without transfers, including obviously strategy-proof rules.

Bade and Gonczarowski (2017) investigate whether some Pareto optimal and strategy-proof social choice rules are OSP in domains such as object assignment, single-peaked preferences, and combinatorial auctions. Troyan (2019) identifies an acyclicity condition that is both necessary and sufficient to implement Top Trading Cycles (TTC) in an obviously strategy-proof way. I also identify a necessary condition for a social choice rule to be OSP and follow up with the examples in matching and object allocation. As an extension of Troyan

[7]Some agents experience cognitive limitations when examining the contingencies (Li, 2017). However, I show that some contingencies are just too clear to be ignored even by cognitively limited agents.

(2019), Mandal and Roy (2022) consider assignment problems where individuals are to be assigned at most one indivisible object and monetary transfers are not allowed. They provide a characterization of assignment rules that are Pareto efficient, nonbossy, and implementable in OSP mechanisms. As corollaries of their result, a characterization of OSP-implementable fixed priority top trading cycles (FPTTC) rules, hierarchical exchange rules, and trading cycles rules are obtained.

I contribute to the mechanism design literature (Zhang and Levin, 2017a; Ferraioli et al., 2020), focusing on private information other than preferences. Fujinaka and Miyakawa (2017) introduce private information on the quality of a house, known only to the initial owner, and present a model of housing markets with interdependent values.[8] Similarly, I introduce private non-preference information included in the agent's type. Ortner (2020) studies how the revelation of new private information affects bargaining outcomes, he considers a dynamic environment in contrast to my paper.

My work is also related to Munoz-Rodriguez (2021), where he uses mechanism design tools to study the problem of deceased-donor organ allocation in the presence of dynamic asymmetric information about transplant candidates' medical urgency. He finds that the current prioritization scheme in US is not incentive compatible, since patients and physicians have strong incentives to engage in strategic manipulations regarding the degree of medical urgency. In line with my paper, the non-preference private information (in this case it is patient's medical urgency) affects the outcome. So, even though the proposed optimal prioritization rules are incentive compatible, the question is whether these rules are OSP. This raises the problem also considered by Halushka (2020). I present a strategy-proof rule that is not invariant in a general one-to-one matching model setting. I also illustrate why OSP implementation fails despite of the social choice rule being strategy-proof.

The paper is organized as follows: I formalize the model in Section 2.3. I define the properties of social choice rules in Section 2.4. I provide the results in Section 2.5. I present the applications in Section 2.6. I conclude in Section 2.7.

2.3 Model

Let $I = \{1, ..., n\}$ be a finite set of agents, with a typical element $i \in I$; n is the total number of agents. Let Y be a set of possible outcomes, with a typical element $y \in Y$.

[8] They introduce private information on the quality of a house (i.e., high or low), which is known only to the initial owner. The ex-post preference of an agent depends on the private information of the other agents with regard to the quality of houses.

For each $i \in I$, let $\theta_i \in \Theta_i$ be the agent i's type, where Θ_i is the set of all i's possible types. Agent i's type, θ_i, is his private information. Let θ_{-i} denote the list of other agents' types. Let $\theta_I \in \Theta_I$ be the profile of types for all the agents, where $\Theta_I = \times_{i \in I} \Theta_i$ is the set of all type profiles. I allow the set of feasible outcomes to depend on agents' types. Let $Z(\theta_I) \subseteq Y$ be the set of feasible outcomes when the type profile is θ_I.[9]

Let $\mathcal{R}_i$ be the set of possible preference relations over Y for agent i. Let $R_i : \Theta_i \to \mathcal{R}_i$ be a mapping from i's types to preference relations. For each $i \in I$, I use $R_i(\theta_i)$ to denote i's preference relation over outcomes when i's type is $\theta_i \in \Theta_i$.[10] I use $P_i(\theta_i)$ to denote strict preference, and $I_i(\theta_i)$ to denote indifference. Given a type profile $\theta_I \in \Theta_I$, $R_I(\theta_I) = (R_1(\theta_1), ..., R_n(\theta_n))$ is the profile of preferences of the agents.

A social choice rule is a mapping from each profile of types to an outcome, $f : \Theta_I \to Y$ such that for each $\theta_I \in \Theta_I$, $f(\theta_I) \in Z(\theta_I)$.

Let $\mathcal{G}$ be the set of all extensive game forms with perfect recall and finite depth (Li, 2017; Pycia and Troyan, 2019). Let $\Gamma \in \mathcal{G}$ denote an extensive game form with outcomes in Y; meaning each terminal history $\omega \in \Gamma$ results in some outcome $y(\omega) \in Y$.

For each $i \in I$, let $\mathcal{I}_i$ denote the information sets for i, with representative element $I_i \in \mathcal{I}_i$. Let H be the set of all histories, with representative element h. Let $\succ$ denote the precedence order over h (Li, 2017). For any $h \in I_i$ and $h' \in I'_i$, if $h \succ h'$, then $I_i \succ I'_i$.

For each $i \in I$, let $\mathcal{A}_i$ denote the set of all possible actions for i. Let $A(I_i) \in \mathcal{A}_i$ denote the set of all actions available to i at $I_i \in \mathcal{I}_i$.

A strategy is $s_i : \mathcal{I}_i \to \mathcal{A}_i$, where $\mathcal{A}_i = \cup_{I_i \in \mathcal{I}_i} A(I_i)$, such that for each $I_i \in \mathcal{I}_i$, $s_i(I_i) \in A(I_i)$. In other words, s_i is a complete contingent plan; it chooses, for each information set $I_i \in \mathcal{I}_i$, a particular action $a \in A(I_i)$. Let S_i be the set of all strategies for i. A strategy profile $s_I = (s_i)_{i \in I}$ specifies a strategy for each $i \in I$.

For a pair of strategies $s_i, s'_i \in S_i$, the earliest points of departure between s_i and s'_i, $\alpha(s_i, s'_i)$, are the earliest information sets I_i, with respect to $\succ$, at which s_i and s'_i diverge $(s_i(I_i) \neq s'_i(I_i))$ (Li, 2017). In other words, $\alpha(s_i, s'_i)$ are the earliest information sets, where these two strategies choose different actions. Note that there can be no earlier point of departure under perfect recall. For any $I_i \in \alpha(s_i, s'_i)$ and any $I'_i \in \mathcal{I}_i$ such that I'_i precedes I_i, $s_i(I'_i) = s'_i(I'_i)$; meaning that at every previous information set s_i and s'_i chose the same action.

[9]The case where feasible set is fixed can be accommodated by setting $Z(\theta_I) = Y$ for each $\theta_I \in \Theta_I$.
[10]R_i only depends on θ_i.

A type-strategy $\sigma_i : \Theta_i \to S_i$ specifies a strategy for every type of agent i, where $\sigma_i(\theta_i)$ denotes the strategy that i plays when his type is θ_i. Let Σ_i be the set of all type-strategies for i. A type-strategy profile $\sigma_I = (\sigma_i)_{i \in I}$ specifies a type-strategy for each $i \in I$.

2.4 Properties of Social Choice Rules

Below I define the central concept of obvious dominance, which states that a strategy can be identified as weakly dominant even by an agent who can not engage in contingent reasoning.

Definition 1. Obviously Dominant Strategy. For each $i \in I$ and each $\theta_i \in \Theta_i$, given Γ, $s_i \in S_i$ is an obviously dominant strategy if, for each $s'_i \in \Sigma_i$, at each $I_i \in \alpha(s_i, s'_i)$, an $R_i(\theta_i)$-worst outcome[11] under any terminal history following s_i is at least as good as an $R_i(\theta_i)$-best outcome[12] under any terminal history following s'_i. For each $i \in I$, given Γ, $\sigma_i^{ODS} \in \Sigma_i$ is an obviously dominant type-strategy if for each $\theta_i \in \Theta_i$, $\sigma_i^{ODS}(\theta_i)$ is an obviously dominant strategy.

A social choice rule is OSP if there exists some extensive game form such that, for every profile of types, there is an equilibrium in obviously dominant type-strategies and it coincides with the outcome as if the rule was applied.

Definition 2. Obvious Strategy-proofness (OSP). Suppose Γ is such that every $i \in I$ has an obviously dominant type-strategy $\sigma_i^{ODS}(\theta_i)$. For each $\theta_I \in \Theta_I$, let $y^{ODS}(\theta_I)$ be the outcome resulting from the players following $\sigma_I^{ODS}(\theta_I)$. If for each $\theta_I \in \Theta_I$, $y^{ODS}(\theta_I) = f(\theta_I)$, then $(\Gamma,\ \sigma_I^{ODS})$ implements f in obviously dominant strategies. If there is $(\Gamma,\ \sigma_I^{ODS})$ that implements f in obviously dominant strategies, then f is OSP.[13]

The next property says that if a social choice rule depends on an agent's private non-preference information, then this agent is indifferent.

[11]An $R_i(\theta_i)$-worst outcome in $Y' \subseteq Y$ is $y \in Y'$ such that for each $y' \in Y'$, $y'\ R_i\ (\theta_i)\ y$.

[12]An $R_i(\theta_i)$-best outcome in $Y' \subseteq Y$ is $y \in Y'$ such that for each $y' \in Y'$, $y\ R_i\ (\theta_i)\ y'$.

[13]This is a form of weak implementation in line with Myerson (1981), Saks and Yu (2005), Li (2017) and others. This is in contrast with full implementation as in Maskin (1999).

Definition 3. Invariant (to Own Non-Preference Information). A social choice rule f is invariant if for each $i \in I$, each pair $\theta_i, \theta_i' \in \Theta_i$, and each $\theta_{-i} \in \Theta_{-i}$, whenever $R_i(\theta_i) = R_i(\theta_i')$, we have $f(\theta_i, \theta_{-i})\ I_i(\theta_i)\ f(\theta_i', \theta_{-i})$.

2.5 Results

I show that every obviously strategy-proof social choice rule is invariant. So, when a social choice rule is OSP and responds to i's non-preference information, then i is indifferent between the outcomes. In other words, i is not affected by i's own non-preference information.

Theorem 1. Every OSP social choice rule f is invariant.[14]

Proof: Suppose for agent $i \in I$, there is a pair $\theta_i, \theta_i' \in \Theta_i$ and $\theta_{-i} \in \Theta_{-i}$, such that $R_i(\theta_i) = R_i(\theta_i')$ and $f(\theta_i, \theta_{-i}) \neq f(\theta_i', \theta_{-i})$.[15]

Let Γ be a game form and σ_I be a type-strategy profile that OSP-implement f. Let $s = \sigma(\theta)$ and $s' = \sigma(\theta')$ be the obviously dominant strategy equilibria under (θ_i, θ_{-i}) and (θ_i', θ_{-i}), respectively. Let ω and ω' be the terminal histories under s and s', respectively. So, i plays s_i when i's type is θ_i and s_i' when i's type is θ_i'. All other agents play s_{-i} when their types are θ_{-i}, so for each $j \in I$ such that $j \neq i$, $s_j' = s_j$.

Since (Γ, σ_I) OSP-implements f and $f(\theta_i, \theta_{-i}) \neq f(\theta_i', \theta_{-i})$, the outcomes under s and s' are different, $y(\omega) \neq y'(\omega')$. Furthermore, the terminal histories under s and s' are also different, $\omega \neq \omega'$. Therefore, since $s_{-i} = s_{-i}'$, $s_i \neq s_i'$.

There is an earliest point of departure between s_i and s_i', $I_i \in \alpha(s_i, s_i')$, at which s_i and s_i' diverge ($s_i(I_i) \neq s_i'(I_i)$).

For each information set I map each action to the set of all outcomes at some terminal node following that action. Let $A(I_i) \subseteq \mathcal{A}_i$ denote the set of all actions available to i at $I_i \in \alpha(s_i, s_i')$.

For all $a \in A(I_i)$, let $Y(a)$ be the set of all outcomes at some terminal node following a. Since Γ OSP-implements f, there exists $a^* \in A(I_i)$, such that for each $a \in A(I_i) \backslash \{a^*\}$, a^* obviously dominates a; meaning the $R_i(\theta_i)$-best outcome following a is not better than the

[14]Note, the converse does not hold: if f is invariant, then it does not mean f is OSP. For example, condition the classic top trading cycles (TTC) rule (that determines a core allocation) on any non-preference private information of the agents so that TTC is invariant. So, TTC is not OSP (Li, 2017) and remains not OSP after conditioning.

[15]Note, if $f(\theta_i, \theta_{-i}) = f(\theta_i', \theta_{-i})$ agent i is trivially indifferent between the outcomes.

$R_i(\theta_i)$-worst outcome following a^*.

Furthermore, $a = s_i(I_i)$ is an obviously dominant strategy. So, for each $a'' \in A(I_i)\backslash\{a\}$, a obviously dominates a''; meaning the $R_i(\theta_i)$-best outcome following a'' is not better than the $R_i(\theta_i)$-worst outcome following a. Also, $a' = s_i'(I_i)$ is an obviously dominant strategy. So, for each $a'' \in A(I_i)\backslash\{a'\}$, a' obviously dominates a''; meaning the $R_i(\theta_i')$-best outcome following a'' is not better than the $R_i(\theta_i')$-worst outcome following a'. Since $R_i(\theta_i) = R_i(\theta_i')$, a' is an obviously dominant strategy at $R_i(\theta_i)$ and a is an obviously dominant strategy at $R_i(\theta_i')$.

There is more than one obviously dominant strategy—a and a', so i is indifferent between all of the outcomes resulting from following a and a', $Y(a)$ and $Y(a')$ respectively. If a obviously dominates a', then the $R_i(\theta_i)$-best outcome following a' is not better than the $R_i(\theta_i)$-worst outcome following a. Also, if a' obviously dominates a, then the $R_i(\theta_i)$-best outcome following a is not better than the $R_i(\theta_i)$-worst outcome following a'. So, for each $y \in Y(a)$ and each $y' \in Y(a')$, $y\ R_i\ (\theta_i)\ y'$ and $y'\ R_i\ (\theta_i)\ y$, as well as $y\ R_i\ (\theta_i')\ y'$ and $y'\ R_i\ (\theta_i')\ y$. Therefore, for each $y \in Y(a)$ and each $y' \in Y(a')$, $y\ I_i\ (\theta_i)\ y'$ and $y'\ I_i\ (\theta_i')\ y$. Then, since $f(\theta_i, \theta_{-i}) \in Y(a)$ and $f(\theta_i', \theta_{-i}) \in Y(a')$, $f(\theta_i, \theta_{-i})\ I_i(\theta_i)\ f(\theta_i', \theta_{-i})$ and $f(\theta_i, \theta_{-i})\ I_i(\theta_i')\ f(\theta_i', \theta_{-i})$.

So, every OSP f is invariant. □

2.6 Applications

2.6.1 Strict Preferences

As I show above, if a social choice rule depends on rich type information (when type includes non-preference private information), then there are necessarily indifferences in agents' preferences. However, when agents have strict preferences over the outcomes, it is not possible for the rule to act on any private information other than preferences.

The following property of the social choice rule says that it ignores private information other than preferences.

Definition 4. Preference-only. A social choice rule f is preference-only if for each pair $\theta_I, \theta_I' \in \Theta_I$, such that $R_I(\theta_I) = R_I(\theta_I')$, we have $f(\theta_I) = f(\theta_I')$.

There is an important difference between being invariant and preference-only. The latter applies to changes to the whole profile; meaning each agents' type can be changed at the same time as long no agent's preferences change.

When agents have strict preferences over the outcomes, the necessary condition for obvious strategy-proofness becomes a simpler one—preference-onlyness.

Proposition 1. If there are no indifferences in agents' preferences over outcomes, then every OSP social choice rule f is preference-only.

Proof: Let the pair $\theta_I, \theta'_I \in \Theta_I$ be such that $R_I(\theta_I) = R_I(\theta'_I)$. We change the profile of types $\theta_I \in \Theta_I$ to $\theta'_I \in \Theta_I$ by changing the type of one agent $i \in I$ at a time, from $\theta_i \in \theta_I$ to $\theta'_i \in \theta'_I$.

First, change the type of the agent 1 from $\theta_1 \in \Theta_1$ to $\theta'_1 \in \Theta'_1$, and, for each $i \in I \backslash \{1\}$, keep $\theta_{-i} \in \theta_I$. By applying Theorem 1, agent 1 is indifferent between the outcomes under $s = \sigma(\theta_1)$ and $s' = \sigma(\theta'_1)$, y and y' respectively. So, for each $y \in Y(a)$ and each $y' \in Y(a')$, $y \; I_1 \; (\theta_1) \; y'$ and $y' \; I_1 \; (\theta'_1) \; y$. Then, since $f(\theta_1, \theta_{-i}) \in Y(a)$ and $f(\theta'_1, \theta_{-i}) \in Y(a')$, $f(\theta_1, \theta_{-i}) \; I_1(\theta_1) \; f(\theta'_1, \theta_{-i})$ and $f(\theta_1, \theta_{-i}) \; I_1(\theta'_1) \; f(\theta'_1, \theta_{-i})$. However, if there are no indifferences in agent 1's preferences, then $f(\theta_1, \theta_{-i}) = f(\theta'_1, \theta_{-i})$.

Next, keep changing the type of each agent $i \in I \backslash \{1\}$ from $\theta_i \in \theta_I$ to $\theta'_i \in \theta'_I$ one agent at a time until, for each $i \in I$, $\theta'_i \in \theta'_I$. By applying Theorem 1 and the same reasoning, we have $f(\theta_1, \theta_2, \theta_i, ..., \theta_n) = f(\theta'_1, \theta'_2, \theta'_i, ..., \theta'_n)$, that is $f(\theta_I) = f(\theta'_I)$.

Therefore, for each pair $\theta_I, \theta'_I \in \Theta_I$, such that $R_I(\theta_I) = R_I(\theta'_I)$, we have $f(\theta_I) = f(\theta'_I)$, so f is preference-only. □

Proposition 1 does not hold if one weakens OSP to strategy-proofness. Consider the following example. Let $\mathcal{P}$ be the set of strict preferences over Y and for each $i \in I$, $\theta_i = \mathcal{P} \times \{0, 1\}$ be i's type.[16] For each $\theta_i \in \Theta_i$, denote by $s(\theta_i) \in \{0, 1\}$ i's non-preference private information at type θ_i. Define a social choice rule f such that $f(\theta)$ is the most preferred outcome for $j = argmin\{i \in I : s(\theta_i) = 1\}$ if there is $i \in I$ such that $s(\theta_i) = 1$ or the most preferred outcome of agent 1 otherwise.[17] This social choice rule is strategy-proof, but not OSP.

[16] I assume ex-post verifiability by the same reasoning as before; meaning one can withhold and report type 0 instead of 1, but one can not overreport 1 if his true type is 0. The main reason is that it is verifiable after the fact and an untruthful agent is penalized.

[17] In other words, look at all those agents $i \in I$ for whom $s(\theta_i) = 1$ and look at the i with the lowest index (that is j), then pick the most preferred outcome for j if there is such an i. If not, then pick the most preferred outcome of 1 otherwise.

2.6.2 Matching

Consider the following application in matching theory, which is a special case of Halushka (2020). On one side of the market, there are sellers with different levels of experience. An individual seller is denoted by $s_i \in S$ for each $i \in \{1, ..., n\}$, where S is a finite set of sellers and n is the total number of sellers. On the other side of the market, there are buyers that have projects needing different levels of experience. An individual project is denoted by $p_j \in P$ for each $j \in \{1, ..., m\}$, where P is a finite set of projects and m is the total number of projects. For each $s_i \in S$, $l_{s_i} \in \mathbb{R}_+$ is the level of experience of s_i. $l_S = (l_{s_1}, ..., l_{s_n})$ is the profile of levels of experience of the sellers. $R_S = (R_{s_1}, ..., R_{s_n})$ is the profile of preferences of the sellers, where R_{s_i} is a strict preference relation of the seller s_i over the projects and the option of being unmatched. $\mathcal{R}_{s_i}$ is the set of all preference relations for s_i. For each $s_i \in S$, let $\Theta_{s_i} = \mathcal{R}_{s_i} \times \mathbb{R}_+$ be the set of all possible types of s_i. For each $s_i \in S$, $\theta_{s_i} \in \Theta_{s_i}$ is s_i's type that holds private information about R_{s_i} and l_{s_i}. Let $\theta_S \in \Theta_S$ be the profile of types of all the sellers, where Θ_S is the set of all type profiles of the sellers. $R_P = (R_{p_1}, ..., R_{p_m})$ is the profile of preferences of the projects, where R_{p_j} is a preference relation of the project p_j over the sellers and the option of being unmatched (all projects prefer sellers with the higher experience to the ones with lower experience).

An outcome (a matching), $y \subseteq S \times P \times \mathbb{R}_+$, is such that for each $s_i \in S$, if $(s_i, p_j, l_{s_i}) \in y$, then there is no $p'_j \neq p_j$ such that $(s_i, p'_j, l_{s_i}) \in y$, and for each $p_j \in P$, if $(s_i, p_j, l_{s_i}) \in y$, then there is no $s'_i \neq s_i$ such that $(s'_i, p_j, l'_{s_i}) \in y$. This means that y is a mapping of sellers to projects, such that only one seller can be assigned to a project. Y is the set of all possible outcomes (matchings). $Z(R_S, l_S^T) = \{y \in Y :$ for each $s_i \in S$, if $(s_i, p_j, l_{s_i}) \in y$, then $l_{s_i} \leq l_{s_i}^T\}$ is the set of all feasible outcomes (matchings).

An economy is described by the types of the sellers. A social choice rule maps each profile of sellers' types to an outcome, $f : \Theta_S \to Y$ such that for each $\theta_S \in \Theta_S$, $f(\theta_S) \in Z(\theta_S)$. I assume ex-post verifiability, meaning if after the fact the outcome is not feasible with respect to the true type, then sellers are punished (fired).

Chapter 1 shows the existence of strategy-proof social choice rules that return stable and, therefore, individually rational and Pareto efficient outcomes (matchings). Stability states that there is no blocking project and seller pair, such that the seller prefers the project and has the higher type (experience level), than the seller project is assigned to under the rule (the project always prefers the seller with a higher type). Individual rationality means that each agent finds his assignment at least as good as being unmatched. Pareto efficiency means there is no other matching that makes at least one agent/project better off without making another one worse off. Each of these social choice rules is a modification of the

Serial Dictatorship rule.[18] Furthermore, each such rule depends on private information of the agents—Chapter 1 (Halushka, 2020): R_{s_i} and l_{s_i}. Varying the types (experience levels) of the agents affects the outcome that such a rule induces. So, none of these social choice rules are invariant. In fact, no stable rule is OSP.[19]

Corollary 1. SD with cutoffs is not OSP, since it is not OESP.

Proof: SD with cutoffs is not invariant, then by Theorem 1, it is not OESP. As a result, it's not OSP. □

Note that I have assumed that projects are not strategic. This strengthens the impossibility result.

2.6.3 Object allocation

Incentive compatibility and efficiency are primary concerns when designing social choice rules for the allocation and exchange of discrete resources (objects) (Abdulkadiroglu and Sönmez, 1999; Papai, 2001; Abdulkadiroglu and Sönmez, 2003; Roth et al., 2004; Bade, 2015). Papai (2000) characterizes the class of group strategy-proof, Pareto efficient and "reallocation proof" rules. Pycia and Ünver (2017) drop "reallocation proofness" from this characterization. Bade and Gonczarowski (2017) and Pycia and Troyan (2019) characterize the set of all OSP and Pareto efficient rules.

For the object allocation problem, Bade and Gonczarowski (2017) characterize the class of OSP and Pareto efficient matching rules as sequential barter with lurkers. Each social choice rule in this class establishes matchings in many trading rounds with at most two owners at each round. In each such round, each owner points to his preferred house and sequentially owns each house that is not matched yet. As long as no one is matched, the rule chooses an increasing set of houses and has them point to owners. These choices may be based on the preferences of already matched owners or other information. Each of the owners has the option to leave with a house that he owns or to swap if both agree. Once a cycle forms, the

[18] Consider a modified serial dictatorship. First, each seller reports his experience level. Second, break all ties for the sellers with the same reported experience level in the same way if there are any. Next, starting from the seller with the best tie-breaker of the highest experience level and proceeding until the seller with the worst one of the lowest experience level, assign every seller his most preferred available project and so on.

[19] In line with Li (2017); Ashlagi and Gonczarowski (2018) and others.

owners and houses in that cycle are matched. When a lurker appears, he may ultimately be matched to all but one special house in the set. If he favours this special house the most, he may "lurk" it and no longer be an owner.[20] If no agent who is entitled to be matched with this special house chooses to do so, then the lurker obtains it as a residual claimant. Otherwise, the lurker gets the second-best house in this set.

Sequential barter with lurkers is a rule that designates, for each round, houses and owners who will pick next as a function of the preferences of already matched owners or as a function of who picked what before. This information determines further ordering in a sequential barter with lurkers. Furthermore, to extend the understanding of the application one can condition this sequential barter with lurkers on some other private non-preference information of the owners,[21] rather than just define it as a function of what house was picked before him or the preferences of the already matched owners. In other words, one can extend the type of the owner to contain other non-preference private information in addition to preferences.

Similarly, one can take any OSP and Pareto efficient social choice rule (Bade and Gonczarowski, 2017; Pycia and Troyan, 2019) and extend it in a way that it is invariant and the rule will remain OSP: along with conditioning on the information about what happened before, one can condition on any non-preference private information of earlier agents. However, Theorem 1 says that one can not go beyond these extensions. The only viable additional changes that can be brought to the rule is conditioning for later agents on previous agents' allocations or on previous agents' non-preference information.

2.7 Conclusion

In this article I focus on OSP mechanism design with rich private information. My main contribution is to identify a necessary condition for OSP social choice rules when an agent's type can include private information other than just preferences. I prove that every OSP social choice rule satisfies invariance to own non-preference information. Furthermore, to highlight the importance of the result I present a few applications, among them: strict preferences case, matching and object allocation.

My results demonstrate that there are some contingencies that are too clear to be ignored even by cognitively limited agents. For example, in Section 2.6.2 the only information agents

[20]There are at most two owners at any point, and additionally any number of lurkers, each for a different house

[21]It can be any kind of the information that does not affect the agent's outcome and that the agent has to reveal when picking the house.

have to report is their types (experience levels). So, it would be highly implausible that the agents would undermine themselves by underreporting their types (experience levels) and move themselves lower in the queue. Therefore, my findings suggest that obvious strategy-proofness may be excessively restrictive and demanding.

A natural extension from this work is questioning whether obvious strategy-proofness is the right criterion for some types of market design problems. Obvious strategy-proofness includes guarding against certain kinds of implausible suboptimal behaviour. This opens the question of developing a model that refines these types of suboptimal behaviours in the setting with rich private information and testing it experimentally.

Chapter 3

Almost Obvious Dominance

Abstract

The main assumption behind the obvious dominance is that agents might be cognitively limited and can not engage in contingent reasoning at all. This assumption is unreasonably weak compared to the standard assumption that agents can perfectly distinguish contingencies. I strengthen it slightly by assuming that agents are able to do at least some contingent reasoning. I define what it means for the strategy to be obviously dominant with respect to a partition of the state space. I call such strategies *partition dominant strategies*. A strategy is an *almost obviously dominant* if, for all possible partitions, except coarsest, it can be identified as being partition dominant. My hypothesis is that even though some agents can not do state-by-state reasoning as rational players do, they are able to do at least some partitioning of the other player's actions and regardless of how the partitioning is done, the agents can identify an almost obviously dominant strategy.

Keywords: Obvious dominance, weak dominance, mechanism design.

3.1 Introduction

Obvious strategy-proofness is a solution concept that considers the way cognitively limited agents behave and choose their strategies. An obviously dominant strategy is one that can be identified as dominant even by agents who can not engage in contingent reasoning. Obvious strategy-proofness requires equilibrium in obviously dominant strategies (Li, 2017). Obvious dominance is too strong to make a recommendation in many games, since the assumption that agents can not engage in any contingent reasoning at all is too weak. Properties like strategy-proofness or obvious strategy-proofness say that a mechanism is not susceptible to certain kinds of behaviour (full information and full rationality in the case of strategy-proofness, inability to distinguish any contingencies in the case of obvious strategy-proofness). The greater the range of behaviours a property says a mechanism must accommodate, the stronger the property. That is why obvious strategy-proofness is stronger than strategy-proofness. My hypothesis is that agents are able to do at least some contingent reasoning as long as it is easy to do. Therefore, I propose one (particular) step closer to strategy-proofness from obvious strategy-proofness. That step is to not have to accommodate people who can not do any contingent reasoning at all.

Consider the following game. There are two players and only one desired object. There are only two possible outcomes: first player or second one gets the object.[1] First, each player has to report a natural number, starting from 1. Second, the player with the lowest number reported gets the object. Ties are broken in favour of the first player.

In this game, the first player has an obviously dominant strategy (ODS), which is to report 1. However, the second player does not have an ODS.

Consider the game from the second player's prospective. Suppose he can bipartition the set of all natural numbers with respect to the first player's actions; meaning he can not perform the state-by-state reasoning in the same way as a rational player can, but he can do at least some contingent reasoning. Regardless of what the partitioning is, for the partition component containing 1 and some other natural numbers, the second player can not distinguish between contingencies. He can not make a decision based on obvious dominance.[2] Moreover, for the other component of the partition (the one not containing 1): the second

[1]There is another possible outcome when nobody gets the object. However, I do not consider it, assuming efficient allocation of the object.

[2]If the component of the partition includes only 1: the second player's worst possible outcome from reporting 1 (he does not get the object) is the same as the best possible outcome from reporting any natural number greater than 1 (he does not get the object). So, reporting 1 does not obviously dominate and is not obviously dominated.

player is always better off reporting 1 (he always gets the object), than any other natural number.

Consider the following motivating example. The second player partitions the first player's actions into two sets (components of the partition): $A^1 \cup A^2 = \mathbb{A}$, where $\mathbb{A} = \{1, 2, 3\}$. Case 1, partition A_1: $A_1^1 = \{1\}$, $A_1^2 = \{2, 3\}$. In this case, 1 is a PDS: there is at least one component of the partition, A_1^2, over which 1 obviously dominates all other actions and there are no other components over which 1 is obviously dominated. If the second player thinks the first player's action is in A_1^2, then he reports 1 and guarantees himself an object; 1 obviously dominates all other actions over A_1^2. If the second player thinks the first player's action is in A_1^1, then regardless of the second player's actions, he never gets the object - reporting 1 does not obviously dominate and is not obviously dominated. So, reporting 1 is a PDS for the second player with respect to A_1. Case 2, partition A_2: $A_2^1 = \{2\}$, $A_2^2 = \{1, 3\}$. Following the same reasoning, 1 is a PDS: if the first player's action is in A_2^1, then 1 obviously dominates all other actions over A_2^1; if in A_2^2, then 1 does not obviously dominate and is not obviously dominated. So, reporting 1 is a PDS for the second player with respect to A_2. Case 3, partition A_3: $A_3^1 = \{3\}$, $A_3^2 = \{1, 2\}$. Following the same reasoning, 1 is a PDS: if the first player's action is in A_3^1, then 1 obviously dominates all other actions over A_3^1; if in A_3^2, then 1 does not obviously dominate and is not obviously dominated. So, reporting 1 is a PDS for the second player with respect to A_3. If the second player can do state-by-state reasoning, partition A_4: $A_4^1 = \{1\}, A_4^2 = \{2\}, A_4^3 = \{3\}$. This is the finest partition and reporting 1 is a weakly dominant strategy for the second player. Following the same reasoning, 1 is a PDS: 1 obviously dominates all other actions over A_4^2 and A_4^3 and there are no other components over which 1 is obviously dominated, so 1 is a PDS for the second player with respect to A_4. Note, following Li (2017), if the second player can not perform contingent reasoning at all: $A_5 = \{1, 2, 3\}$. This is the coarsest partition and there is no ODS for the second player. So, 1 is a PDS for the second player with respect to all possible partitions (A_1, A_2, A_3, A_4), but the coarsest (A_5).

So, the second player can identify that reporting 1 is an *almost obviously dominant strategy (AODS)* for him, even though it is not an ODS, as long as he can perform at least some contingent reasoning: not the finest partitioning as rational player does, but also not the coarsest as Li (2017) suggests. So, the strategy is an *almost obviously dominant strategy (AODS)* if, for all possible partitions, except coarsest, it can be identified as being *partition dominant strategy (PDS)*; meaning, for all possible partitions, but the coarsest, there exists at least one component of the partition, over which it obviously dominates all other strategies (the preference-worst outcome following this strategy is at least as good as a preference-best outcome following any deviation) and there is no other component of the partition over which it is obviously dominated. Furthermore, the second player splits all the first player's possible actions into two sets and regardless of how he does this partitioning, as long as both sets are not empty, he can identify a PDS. I show that considering finer partitions does not change anything.

3.2 Related Literature

There is experimental evidence (Hassidim et al., 2017; Shorrer and Sóvágó, 2018; Rees-Jones and Skowronek, 2018) that some agents are limited in their ability to engage in contingent reasoning. Li (2017) suggests that if a social choice rule is obviously strategy-proof, it can be implemented in a way that even agents with such cognitive limitations can recognize their strategies as weakly dominant. He introduces obvious dominance to characterize the behaviour of such agents and introduces OSP. I weaken the concept of obvious dominance and introduce almost obvious dominance to characterize the choices of the agents that, as I assume, are able to engage at least in some contingent reasoning. One can then use this definition to similarly define the notion of almost OSP mechanisms.

I contribute to the mechanism design literature involving unsophisticated agents (Chen and Li, 2018; Börgers and Li, 2019). Li and Dworczak (2020) study the design of simple mechanisms where even cognitively limited agents can determine their optimal strategy. Pycia and Troyan (2021) relax the assumption that agents can predict their own future moves and define strongly obviously strategy-proof mechanisms, strengthening the notion of simplicity. I expand the theoretical horizon even further by assuming that even unsophisticated agents can partition all other agents' actions (states) into at least two sets, and define almost obviously strategy-proof mechanisms.

There is a robust literature on obvious strategy-proofness, which my paper contributes to, such as: Bade and Gonczarowski (2017); Mackenzie (2020); Ashlagi and Gonczarowski (2018); Pycia and Troyan (2019); Martínez-Marquina et al. (2019b); Troyan (2019); Ferraioli et al. (2020); Troyan and Morrill (2020) and many others. Zhang and Levin (2017b,c) offer an axiomatic approach when direct mechanisms, where agents are unable to form the finest partition of their universe, might rely on coarser partitions of events containing more than one state. Due to obvious dominance being too strong in some cases, I also focus on partitioning of the opponents' actions (states) by the agents, as opposed to Zhang and Levin partitioning the agents' own universe, and define partition dominant strategy. My main hypothesis is in line with Zhang and Levin (2017b,c), the agents don't need to have much strategic sophistication to determine their optimal strategies, they should be able to engage at least in some contingent reasoning.

The paper is organized as follows: I formalize the model in Section 3.3. I give an example of the game in Section 3.4. I present the assumption on agents' reasoning abilities in Section 3.5. I define almost obvious dominance in Section 3.6. I relate it to obvious dominance and weak dominance in Section 3.7. I conclude in Section 3.8.

3.3 The Model

Let $N = \{1, ..., n\}$ be a finite set of agents, with typical element $k \in N$; where n is the total number of agents. Let Y be the set of all possible outcomes, with a typical element $y_k \in Y$; y_k denotes the outcome of agent k.

For each $k \in N$, let A_k denote the set of all possible actions for k, with typical element $a_k \in A_k$. Let $\mathcal{A}_N$ be the set of all profiles of actions for the agents.

Let S be the states of the world. For each k, let $\Omega_k = S \times A_{-k}$ be agent k's state space with a typical element $\omega_k \in \Omega_k$. Let Ω_N be the set of all states for the agents.

Let $\mathcal{P}_k$ be the set of all possible partitions of Ω_k. Let $\bar{\mathcal{P}}_k$ be the set of all non-trivial partitions of Ω_k.[3] For each $P_{m,k} \in \mathcal{P}_k$, $\{P^1_{m,k}, P^2_{m,k}, ..., P^{|P_{m,k}|}_{m,k}\}$ are the components of $P_{m,k}$, where m is the index of the partition.

Let $P_{m,k} \in \mathcal{P}_k$ be a bipartition if it has exactly two components.

A rule, $\varphi : \Omega_N \to Y$, selects an outcome for each state.

3.4 An Example of a Game

I give an example of a game expressed using the abstract model, which is a serial dictatorship (SD) with an endogenously determined order.

First, each agent reports a natural number, starting from 1. Second, each agent is randomly assigned a lottery number. Then the agent with the lowest number reported wins. Ties are broken according to the lottery number.[4]

[3] By non-trivial partitions I mean all but the coarsest partition ($\{\Omega_{-k}\}$). I exclude singleton partition, which is the coarsest (as in Li (2017)), but not the partition of singletons (the finest one).

[4] I assume that the lottery numbers are never tied.

Suppose $N = \{1, 2\}$. For 1, the state $\omega_1 = (S, a_2)$ consists of the lottery outcome $S \in \{l_1, l_2\}$ and the action of the other agent $a_2 \in \mathbb{N}$. The lottery determines who gets the priority if there are ties.[5] $Y = \{y_1, y_2\}$ is the set of possible outcomes, where y_1 and y_2 are the outcomes when agent 1 or 2 wins, respectively.

$\mathcal{P}_1$ is the set of all partitions of $\{l_1, l_2\} \times \mathbb{N}$.

3.5 An Assumption on Reasoning Ability

Let us consider agent 1 and his actions in the above game. If he is a fully rational agent, 1 should always report the lowest number possible, which is 1. Let $a_1^r \in A_1$ denote the action that agent 1 would play, if he was able to reason state-by-state. So, $a_1^r = 1$ is a weakly dominant strategy in this game for 1.[6]

Suppose 1 can not perform contingent reasoning at all—can not think hypothetically (Li, 2017). Then, 1 can not understand that he should always play a_1^r, regardless of what 2 does. Therefore, he looks for an obviously dominant strategy (ODS): the worst 1 can get from a_1^r is $y_2 \in Y$, the best 1 can get from deviating is y_1. So, choosing a_1^r for 1 is not an ODS.

Obvious dominance is too strong to make a recommendation in this game. The main goal is to identify what is "too strong" about obvious dominance in this case and offer the alternative theory.

On the other hand, 1 can "easily" identify a_1^r as a weakly dominant strategy. However, what does it mean that this strategy is *easy* to identify for an agent who can not engage in fully contingent reasoning?

Strategies that are like these, even though identifying them involves some contingent reasoning, are easy to identify; meaning the type of contingent reasoning is very easy to do and understand.

Li (2017) focuses on one of the two extremes - on agents that can not engage in contingent reasoning at all. Oppositely, there is another extreme - fully rational agents that can do state-by-state reasoning. I assume that agents have an ability to do at least some contingent reasoning. I look at the agents who are in between and have at least some partition in their

[5] l_1 and l_2 are the lottery outcomes for players when 1 wins or 2 wins respectively (one can use a coin-flip to break ties). This is just an example, one can have a larger state space. In this particular example I consider only two states: agent 1 has a tie-breaker and agent 2 has a tie-breaker.

[6] Choosing a_1^r always gives agent 1 at least as good outcome as choosing any deviation, no matter what 2 does, and there is at least one set of 2's actions for which a_2^r gives a better outcome than any deviation.

mind. In other words, I focus on the agents who can partition their opponents' actions into *at least* two contingencies or two sets, and can contingently reason across them, but not within.

3.6 Contingent Reasoning

I formally discuss agents' ability to reason in a contingent matter.

First, I specify when the strategy is considered to be dominant with respect to the partition—partition dominance.

Definition 1. Partition Dominant Strategy (PDS). For each $k \in N$, each $a_k \in A_k$ and each $P_{m,k} \in \mathcal{P}_k$, $a_k \in A_k$ is a partition dominant strategy with respect to $P_{m,k}$ if there is $c = 1, ..., |P_{m,k}|$ such that for each $a'_k \in A_k$ the preference-worst[7] outcome following a_k is at least as good as the preference-best[8] outcome following a'_k when the state is in $P^c_{m,k}$; and there is no $c' = 1, ..., |P_{m,k}|$ such that the preference-worst outcome following a'_k is at least as good as the preference-best outcome following a_k when the state is in $P^{c'}_{m,k}$.

Next, consider fully rational agents that can distinguish all other players' actions and do state-by-state reasoning. I can write down all the agent 1's partitions of the state space as a lattice.[9] So, if the agent 1 is fully rational, then he can do the comparison within the whole set Ω_{-1} and see that choosing a^r_1 is his weakly dominant strategy.

Definition 2. Weakly Dominant Strategy (WDS). For each $k \in N$ and each $a_k \in A_k$, $a_k \in A_k$ is a weakly dominant strategy if for the finest partition it is a PDS.

Consider agents who can not engage in contingent reasoning at all and are looking for an ODS (Li, 2017). They can not distinguish between other players' actions. In the previously discussed game, agent 1 looks at the set of all outcomes that the state space leads to, 1 can not perform any finer partitioning than $P_{1,1} = \{\Omega_{-1}\}$.

Consider two actions 1 can play $a, b \in A_1$.

[7]An R_k-worst outcome in $Y' \subseteq Y$ is $y \in Y'$ such that for each $y' \in Y'$, $y'\ R_k\ y$, where R_k is k's preference relation.

[8]An R_k-best outcome in $Y' \subseteq Y$ is $y \in Y'$ such that for each $y' \in Y'$, $y\ R_k\ y'$, where R_k is k's preference relation.

[9]In general, if there are t states that the others can play - there are t components of the partition (contingencies) that a rational player can make comparisons across; meaning compare over the finest possible partition, which is the partition of singletons.

Since 1 can not engage in any contingent reasoning, 1 can compare only two sets of outcomes that result from 1 playing one of his actions. For each action $a, b \in A_1$, $Y^a \subseteq Y$ is the set of outcomes 1 can get following $a \in A_1$ and $Y^b \subseteq Y$ is the set of outcomes 1 can get following $b \in A_1$.

So, for $1 \in N$, $a \in A_1$ is an ODS if, for each possible partition, which is $P_{1,1} = \Omega_{-1}$, and for $b \in A_1$, the preference-worst outcome in Y^a following $a \in A_1$ is at least as good as the preference-best outcome in Y^b following $b \in A_1$.

So, if agent 1 can not do contingent reasoning at all, then he can not do any comparisons within the whole set of states and see that choosing a_1^r is his weakly dominant strategy. Agent 1 can only search for ODS, however, there is no ODS in this game.

Following definition is consistent with Li (2017).

Definition 3. Obviously Dominant Strategy (ODS). For each $k \in N$ and each $a_k \in A_k$, $a_k \in A_k$ is an obviously dominant strategy if for the coarsest partition it is a PDS.

Finally, consider the game when 1 can do at least some contingent reasoning and partitions 2's actions (states) into some sets. So, 1 reasons only across these sets, but not within.

Definition 4. Almost Obviously Dominant Strategy (AODS). For each $k \in N$ and each $a_k \in A_k$, a_k is an almost obviously dominant strategy if, for all possible non-trivial partitions, a_k is a PDS.

3.7 Relationships between the dominance concepts

The following results establish the place of an AODS as a concept between a WDS and an ODS. I show that an AODS is weaker than an ODS and stronger than a WDS.

First, I consider the relation between an AODS and a WDS. If a strategy is an AODS, then it is a WDS. However, the reverse is not true, a WDS is not necessarily an AODS.

Proposition 1. For each $k \in N$ and each $a_k \in A_k$, if $a_k \in A_k$ is an AODS, then it is a WDS, but the opposite is not true.

Proof: By definition, if $a_k \in A_k$ is an AODS, then a_k dominates at least over one $P^c_{m,k} \in P_{m,k}$, where $c = 1, ..., |P_{m,k}|$, of every possible partition $P_{m,k} \in \mathcal{P}_k$, but the coarsest. Meaning a_k is a PDS with respect to the partition of singletons, too: set of components of the finest partition, where each component includes only one action (state).[10] Therefore, if $a_k \in A_k$ is an AODS, then a_k is at least as good as any $a'_k \in A_k$ or better in every state. So, if $a_k \in A_k$ is AODS, it is WDS.

However, a WDS does not imply an AODS. Consider the following example. There are two players, 1 and 2. Let $A_1 = (a'_1, a'_2)$ and let $A_2 = (a_1, a_2, a_3, a_4)$ be the set of possible actions for 1 and 2, respectively. I only show player 1's outcomes, they are such that:

		2			
		a_1	a_2	a_3	a_4
1	a'_1	10	8	10	8
	a'_2	9	7	9	7

If 1 does the finest partitioning of 2's actions, $P_{1,1} = \{\{a_1\}, \{a_2\}, \{a_3\}, \{a_4\}\}$, then it is clear that a'_1 weakly dominates a'_2, this is the case if 1 can reason state-by-state. Let 1 be able to do only some contingent reasoning and partition 2's actions in a following way: $P_{2,1} = \{\{a_1, a_2\}, \{a_3, a_4\}\}$. In this case, a'_1 does not partition dominate a'_2.[11] So, a'_1 is not an AODS for P_1. However, a'_1 is a WDS for 1. Therefore, a WDS does not imply an AODS. □

Next, I consider the relation between an AODS and an ODS. If a strategy is an ODS, then it is an AODS, but not the opposite.

Proposition 2. For each $k \in N$ and each $a_k \in A_k$, if $a_k \in A_k$ is an ODS, then it is an AODS, but the opposite is not true.

Proof: By definition, if $a_k \in A_k$ is ODS, it is AODS (Definition 3). If a_k is a PDS for the coarsest partition, it is a PDS for all other finer possible partitions. So, a_k is an AODS.

However, an AODS is not necessarily an ODS. Consider the following example. There are two players, $N = \{1, 2\}$, and only one desired object. There are two outcomes: player 1 or 2 gets the object, y_1 and y_2, respectively. First, each player reports a number in $\{1, 2, 3\}$. Second, the player with the lowest number reported gets the object. Ties are broken in favour of player 1.

[10] This is same partitioning as a rational player would do - state-by-state reasoning.

[11] For $P^1_{2,1}$ and $P^2_{2,1}$ the worst outcome from playing a'_1 is 8 and the best outcome from playing a'_2 is 9. So, since $8 < 9$, a'_1 is not a PDS with respect to $P_{2,1}$.

In this game, player 1 has an ODS, which is to report 1. Player 2 does not have an ODS, however, he has an AODS. Let's first consider how player 2 partitions player 1's actions: $\bar{\mathcal{P}}_2 = \{P_{1,2}, P_{2,2}, P_{3,2}, P_{4,2}, P_{5,2}\}$.[12] In this case $\Omega_2 = A_1 = \{1, 2, 3\}$.

Case 1, partition $P_{1,2}$: $P^1_{1,2} = \{1\}$, $P^2_{1,2} = \{2, 3\}$. In this case, 1 is a PDS: there is at least one component of the partition, $P^2_{1,2}$, over which 1 obviously dominates all other actions and there are no other components over which 1 is obviously dominated. If player 2 thinks player 1's action is in $P^2_{1,2}$, then he reports 1 and guarantees himself an object; 1 obviously dominates all other actions over $P^2_{1,2}$. If player 2 thinks the player 1's action is in $P^1_{1,2}$, then regardless of the player 2's actions, he never gets the object - reporting 1 does not obviously dominate and is not obviously dominated. So, reporting 1 is an AODS for the player 2 with respect to $P_{1,2}$.

Case 2, partition $P_{2,2}$: $P^1_{2,2} = \{2\}$, $P^2_{2,2} = \{1, 3\}$. Following the same reasoning, 1 is a PDS: if player 1's action is in $P^1_{2,2}$, then 1 obviously dominates all other actions over $P^1_{2,2}$; if in $P^2_{2,2}$, then 1 does not obviously dominate and is not obviously dominated. So, reporting 1 is an AODS for player 2 with respect to $P_{2,2}$.

Case 3, partition $P_{3,2}$: $P^1_{3,2} = \{3\}$, $P^2_{3,2} = \{1, 2\}$. Following the same reasoning, 1 is a PDS: if player 1's action is in $P^1_{3,2}$, then 1 obviously dominates all other actions over $P^1_{3,2}$; if in $P^2_{3,2}$, then 1 does not obviously dominate and is not obviously dominated. So, reporting 1 is an AODS for player 2 with respect to $P_{3,2}$.

Case 4, if the second player can do state-by-state reasoning, partition $P_{4,2}$: $P^1_{4,2} = \{1\}, P^2_{4,2} = \{2\}, P^3_{4,2} = \{3\}$. This is the finest partition and reporting 1 is a weakly dominant strategy for player 2. Following the same reasoning, 1 is a PDS: 1 obviously dominates all other actions over $P^2_{4,2}$ and $P^3_{4,2}$ and there are no other components over which 1 is obviously dominated, so 1 is an AODS for player 2 with respect to $P_{4,2}$.

Case 5, following Li (2017), if player 2 can not perform contingent reasoning at all: $P_{5,2} = \{1, 2, 3\}$. This is the coarsest partition and there is no ODS for player 2. So, 1 is a PDS for player 2 with respect to all possible partitions $(P_{1,2}, P_{2,2}, P_{3,2}, P_{4,2})$, but the coarsest $(P_{5,2})$. As a result, 1 is an AODS for player 2.

So, an AODS is not necessarily an ODS. □

The next proposition shows that if a strategy is a PDS of all partitions of size l, then it is also a PDS of all partitions of size greater than l.

Proposition 3. Partition dominance for a given partition implies partition dominance for any finer partition.

[12] I consider all possible partitions of player 1's actions for player 2.

Proof: By contradiction. Suppose $a_k \in A_k$ is a PDS for all possible partitions with l non-empty components. Take any finer partition $P_{m,k}$ with $l+1$ components in the partition. If a_k is not a PDS for $P_{m,k}$, then there must be $a'_k \in A_k$ that obviously dominates a_k over at least one component of $P_{m,k}$. Without loss of generality, suppose this is $P^1_{m,k}$. Now consider some finer partition $P_{m',k} = \{P^1_{m,k}, P^2_{m,k} \cup P^3_{m,k}\ , ..., P^{l+1}_{m,k}\}$. Since a_k is dominated over $P^1_{m,k}$, a_k is not a PDS for $P_{m',k}$, which is a contradiction. □

The next corollary states that to determine if the strategy is an AODS, it is sufficient to look only at all possible bipartitions.

Corollary 1. For each $k \in N$ and each $a_k \in A_k$, in order to determine if the a_k is an AODS, without loss of generality, it is sufficient to consider only all possible partitions $P_{m,k} \in \mathcal{P}_k$, such that $|P_{m,k}| = 2$.

Proof: By Proposition 3. □

Finally, I can look at the game introduced earlier. I prove that a^r_1 is an AODS, even though it is not an ODS.

For the game in Section 3.4 there is an AODS, but no ODS for both players. Consider the game introduced earlier (Section 3.4) from 1's prospective.[13]

Let $P_{m,1}$ be a bipartition of Ω_1. Let $(l_2, 1) \notin P^1_{m,1}$. Then $a_1 = 1$ dominates over $P^1_{m,1}$ and $a_1 \neq 1$ is not dominated over $P^2_{m,1}$. Thus, $a_1 = 1$ is a PDS for $P^1_{m,1}$.

Since this is true for any bipartition, by Proposition 3, $a_1 = 1$ is an AODS.

3.8 Conclusion

A natural extension from Halushka (2020, 2021) is questioning whether obvious dominance is excessively restrictive and is the right criterion for some types of market design problems. In this article I focus on the idea that the agents don't need to have much strategic sophistication to determine their optimal strategies, they should be able to engage at least in some

[13]Since there is a tie-breaker and it is symmetric for agent 2, all the conclusions similarly apply to the symmetric case as if the game was considered from 2's prospective.

contingent reasoning. There are two extremes: rational agent that can do state-by-state reasoning or unsophisticated agent with cognitive limitations that can not do any contingent reasoning at all. I strengthen obvious dominance slightly, compared to the standard assumption of agents perfectly distinguishing contingencies, by assuming that agents are able to do at least some contingent reasoning. Therefore, I account for agents that are in between those two extremes, but still very near an ODS boundary. This is why one can call it an "Almost" ODS.

My main contribution is stating when the strategy is obviously dominant with respect to the partition, defining the partition dominance, and introducing the notion of the almost obvious dominance. I outline the conceptual place of an AODS in between an ODS and a WDS. The lessons drawn from this article can be very useful if one applies the idea of AODS to find solutions to market design problems, where ODS is too demanding. I have demonstrated that one can get more out of the AODS than of ODS in some cases.

An interesting avenue for future research would be to see what mechanisms are almost obviously strategy-proof when one can not find obviously strategy-proof mechanisms. Revisit the problem from Chapter 1 to see if SD with cutoffs is implementable in AODS. One can be interested in iterative elimination of AODS strategies, whether subjects can do it or not, and to what extend. Another way to take the ideas in this paper further is to require a partition dominant strategy only for "relevant" or "essential" partitions (including the partition of singletons), which depend on the specific application. For this, it is crucial how these relevant partitions are determined, which should ideally be based on a theoretically identifiable feature of the specific model.

I believe that the model presented here has the potential to be implemented in the laboratory. For example, it would be interesting to find out what percentage of individuals are at the extremes of sophistication and how many have an intermediate level (Level-k classification of types). The matching model presented in Chapter 1 can be tested in the laboratory, too. The experiment could have two stages: in the first part, the experimenter assesses the sophistication level of the subjects in terms of contingent reasoning; in the second stage, the experimenter lets agents participate in a matching game from Chapter 1. Then, the experimenter may compare the stability of the matching economies as a function of the sophistication level of the subjects.

Bibliography

Abdulkadiroglu, A. and Sönmez, T. (1999). House allocation with existing tenants. *Journal of Economic Theory*, 88(2):233–260.

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

Ashlagi, I. and Gonczarowski, Y. A. (2018). Stable matching mechanisms are not obviously strategy-proof. *Journal of Economic Theory*, 177:405–425.

Bade, S. (2015). Serial dictatorship: The unique optimal allocation rule when information is endogenous. *Theoretical Economics*, 10:385–410.

Bade, S. (2020). Random serial dictatorship: The one and only. *Mathematics of Operations Research*, 45(1):353–368.

Bade, S. and Gonczarowski, Y. A. (2017). Gibbard-satterthwaite success stories and obvious strategyproofness. *EC '17: Proceedings of the 2017 ACM Conference on Economics and Computation*, page 565.

Bikhchandani, S. (2006). Ex post implementation in environments with private goods. *Theoretical Economics*, 1:369–393.

Bogomolnaia, A., Deb, R., and Ehlers, L. (2005). Strategy-proof assignment on the full preference domain. *Journal of Economic Theory*, 123(2):161–186.

Bogomolnaia, A. and Moulin, H. (2001). A new solution to the random assignment problem. *Journal of Economic Theory*, 100(3):295–328.

Börgers, T. and Li, J. (2019). Strategically simple mechanisms. *Econometrica*, 87(6):2003–2035.

Chakraborty, A., Citanna, A., and Ostrovsky, M. (2010). Two-sided matching with interdependent values. *Journal of Economic Theory*, 145:85–105.

Chakraborty, A., Citanna, A., and Ostrovsky, M. (2015). Group stability in matching with interdependent values. *Review of Economic Design, Springer; Society for Economic Design*, 19(1):22.

Charness, G. and Levin, D. (2009). The origin of the winner's curse: A laboratory study. *American Economic Journal: Microeconomics*, 1(1):207–236.

Che, Y.-K., Kim, J., and Kojima, F. (2015). Ex post implementation in environments with private goods. *Journal of Economic Theory*, 158:54–86.

Chen, Y.-C. and Li, J. (2018). Revisiting the foundations of dominant-strategy mechanisms. *Economic Theory, Elsevier*, 178(C):294–317.

Dutta, B., Peters, H., and Sen, A. (2007). Strategy-proof cardinal decision schemes. *Social Choice and Welfare*, 30(1):163–179.

Ferraioli, D., Meier, A., Penna, P., and Ventre, C. (2020). Novel paradigm for the design of obviously strategyproof mechanisms. *CEUR Workshop Proceedings*, 2756(5):184–189.

Fujinaka, Y. and Miyakawa, T. (2017). Ex-post incentive compatible and individually rational assignments in housing markets with interdependent values. *SSRN*, page 22.

Gibbard, A. (1977). Manipulation of schemes that mix voting with chance. *Econometrica*, 45(3):665–81.

Halushka, M. (2020). Revealing more than just preferences: one-to-one matching when one sides' preferences depend on private information of the other side.

Halushka, M. (2021). Obviously strategy-proof mechanism design with rich private information.

Hassidim, A., Marciano, D., Romm, A., and Shorrer, R. (2017). The mechanism is truthful, why aren't you? *American Economic Review*, 107(5):220–224.

Hurwicz, L., Maskin, E., and Postlewaite, A. (1995). Feasible nash implementation of social choice rules when the designer does not know endowment or production sets. *The Economics of Informational Decentralization: Complexity, Efficiency, and Stability*, pages 367–433.

Jehiel, P. and Moldovanu, B. (2001). Efficient design with interdependent valuations. *Econometrica*, 69:1237–1259.

Kagel, J., Harstad, R., and Levin, D. (1987). Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica*, 55(6):1275–304.

Katta, A.-K. and Sethuraman, J. (2006). A solution to the random assignment problem on the full preference domain. *Journal of Economic Theory*, 131(1):231–250.

Klumpp, T. (2009). Two-sided matching with spatially differentiated agents. *Journal of Mathematical Economics*, 45(5-6):376–390.

Li, J. and Dworczak, P. (2020). Are simple mechanisms optimal when agents are unsophisticated? *Available at SSRN 3632971.*

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, 107(11):3257–3287.

Mackenzie, A. (2020). A revelation principle for obviously strategy-proof implementation. *Games and Economic Behavior*, 124:512–533.

Mandal, P. and Roy, S. (2022). Obviously strategy-proof implementation of assignment rules: A new characterization. *International Economic Review*, 63(1):261–290.

Martínez-Marquina, A., Niederle, M., and Vespa, E. (2019a). Failures in contingent reasoning: The role of uncertainty. *American Economic Review*, 109(10):3437–3474.

Martínez-Marquina, A., Niederle, M., and Vespa, E. (2019b). Failures in contingent reasoning: The role of uncertainty. *American Economic Review, 109*, 109(10):3437–74.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66:23–38.

Munoz-Rodriguez, E. (2021). Incentive compatible triaging in deceased-donor transplantation.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

Nandeibam, S. (2013). The structure of decision schemes with cardinal preferences. *Review of Economic Design*, 17(3):163–179.

Nesterov, A. (2017). Fairness and efficiency in strategy-proof object allocation mechanisms. *Journal of Economic Theory*, 170:145–168.

Ortner, J. (2020). Bargaining with evolving private information. *The 31st International Conference on Game Theory at Stony Brook*.

Papai, S. (2000). Strategyproof assignment by hierarchical exchange. *Econometrica*, 68(5):1403–1433.

Papai, S. (2001). Strategyproof and nonbossy multiple assignments. *Journal of Public Economic Theory*, 3(3):257–271.

Pycia, M. and Troyan, P. (2019). A theory of simplicity in games and mechanism design. *CEPR Discussion Paper No. DP14043*, pages 1–74.

Pycia, M. and Troyan, P. (2021). A theory of simplicity in games and mechanism design. *Working Paper No. 393, Available at SSRN 3881072*.

Pycia, M. and Ünver, U. (2017). Incentive compatible allocation and exchange of discrete resources. *Theoretical Economics*, 12(1):287–329.

Rees-Jones, A. and Skowronek, S. (2018). Suboptimal behaviour in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behaviour*, 108:317–330.

Roth, A. E., Sönmez, T., and Ünver, M. U. (2004). Kidney exchange. *The Quarterly Journal of Economics*, 119(2):457–488.

Saks, M. and Yu, L. (2005). Weak monotonicity suffices for truthfulness on convex domains. *Paper presented at Proceedings of the 6th Association for Computing Machinery Conference on Electronic Commerce, Vancouver, BC.*

Shorrer, R. and Sóvágó, S. (2018). Obvious mistakes in a strategically simple college admissions environment: Causes and consequences. *Available at SSRN 2993538*.

Sönmez, T. (1999). Strategy-proofness and essentially single-valued cores. *Econometrica*, 67(3):677–689.

Troyan, P. (2019). Obviously strategy-proof implementation of top trading cycles. *International Economic Review*, 60(3):1249–1261.

Troyan, P. and Morrill, T. (2020). Obvious manipulations. *Journal of Economic Theory*, 185:104970.

Zhang, L. and Levin, D. (2017a). Bounded rationality and robust mechanism design: An axiomatic approach. *American Economic Review*, 107(5):235–239.

Zhang, L. and Levin, D. (2017b). Bounded rationality and robust mechanism design: an axiomatic approach. *American Economic Review: Papers and Proceedings*, 107:235–239.

Zhang, L. and Levin, D. (2017c). Partition obvious preference and mechanism design: Theory and experiment. *Microeconomics: Intertemporal Consumer Choice and Savings eJournal*.