CAHIER DE RECHERCHE #2104E
DÉPARTEMENT DE SCIENCE ÉCONOMIQUE
FACULTÉ DES SCIENCES SOCIALES
UNIVERSITÉ D'OTTAWA

WORKING PAPER #2104E
DEPARTMENT OF ECONOMICS
FACULTY OF SOCIAL SCIENCES
UNIVERSITY OF OTTAWA

Obviously Strategy-proof Mechanism Design With Rich Private Information

Mariya Halushka[1]

April 2021

[1] Department of Economics, University of Ottawa mhalu055@uottawa.ca

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.

JEL Classification: D82, D90

Keywords: Obvious strategy-proofness, Mechanism design

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., 2019). 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.

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

[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.

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 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 this case the ordering

[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.

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 behaviour.

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.

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 pa-

[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. This is highly implausible behaviour.

per contributes to (Mackenzie, 2020; Li and Dworczak, 2020). 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 5.2). 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.

I contribute to the mechanism design literature (Zhang and Levin, 2017; 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. 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. I define the properties of social choice rules in Section 3. I provide the results in Section 4. I present the applications in Section 5. I conclude in Section 6.

2 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 $y \in Y$.

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.[7]

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 $\theta_i \in \Theta_i$. 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 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

[7] The case where feasible set is fixed can be accommodated by setting $Z(\theta_I) = Y$ for each $\theta_I \in \Theta_I$.

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.

A type-strategy $\sigma_i : \Theta_i \rightarrow S_i$ 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 $i \in I$.

3 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[8] under any terminal history following s_i is at least as good as an $R_i(\theta_i)$-best outcome[9] 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.

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

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 ($\Gamma$, σ_I^{ODS}) implements f in obviously dominant strategies. If there is (Γ, σ_I^{ODS}) that implements f in obviously dominant strategies, then f is OSP.[10]

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

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 $f(\theta_i, \theta_{-i})\ I_i(\theta_i)\ f(\theta_i', \theta_{-i})$.

4 Results

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

Theorem 1. Every OSP social choice rule f is invariant.

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 $f(\theta_i, \theta_{-i}) \neq f(\theta_i', \theta_{-i})$.[11]

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

[10] 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).

[11] Note, if $f(\theta_i, \theta_{-i}) = f(\theta_i', \theta_{-i})$ agent i is trivially indifferent between 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 $R_i(\theta_i)$-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 $R_i(\theta)$ and a is an obviously dominant strategy at $R_i(\theta')$.

There are 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. □

5 Applications

5.1 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 changes.

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.[12] 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. This social choice rule is strategy-proof, but not OSP.

5.2 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

[12] 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.

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). $Z(R_S, l^T_S) = \{y \in Y$: for each $s_i \in S$, if $(s_i, p_j, l_{s_i}) \in y$, then $l_{s_i} \leq l^T_{s_i}\}$ 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 agents are punished (fired).

Halushka (2020) 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.[13] Furthermore, each such rule depends on private information of the agents (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.[14] These results are corollaries of Theorem 1. Note that I have assumed that projects are not strategic. This strengthens the impossibility result.

[13]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.

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

5.3 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, 2000; Abdulkadiroglu and Sönmez, 2003; Roth et al., 2004; Bade, 2015). Papai (2001) 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 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.[15] 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,[16] rather than just define it as a function of what

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

[16] 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.

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.

6 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 environments, like in Section 5.2, where a strategy-proof way of solving the problem exists, nonetheless, obvious strategy-proofness can not be satisfied. The only information agents have to report in that case 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.

References

Abdulkadiroglu, Atila and Tayfun Sönmez (1999) "House Allocation with Existing Tenants," *Journal of Economic Theory*, Vol. 88, No. 2, pp. 233–260.

——— (2003) "School Choice: A Mechanism Design Approach," *American Economic Review*, Vol. 93, No. 3, pp. 729–747.

Ashlagi, Itai and Yannai A. Gonczarowski (2018) "Stable Matching Mechanisms are Not Obviously Strategy-Proof," *Journal of Economic Theory*, Vol. 177, pp. 405–425.

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

Bade, Sophie and Yannai A. Gonczarowski (2017) "Gibbard-Satterthwaite Success Stories and Obvious Strategyproofness," *EC '17: Proceedings of the 2017 ACM Conference on Economics and Computation*, p. 565.

Charness, Gary and Dan Levin (2009) "The Origin of the Winner's Curse: A Laboratory Study," *American Economic Journal: Microeconomics*, Vol. 1, No. 1, pp. 207–236.

Ferraioli, Diodato, Adrian Meier, Paolo Penna, and Carmine Ventre (2020) "Novel Paradigm for the design of Obviously Strategyproof Mechanisms," *CEUR Workshop Proceedings*, Vol. 2756, No. 5, pp. 184–189.

Fujinaka, Yuji and Toshiji Miyakawa (2017) "Ex-post incentive compatible and individually rational assignments in housing markets with interdependent values," *SSRN*, p. 22.

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

Hassidim, Avinatan, Deborah Marciano, Assaf Romm, and Ran Shorrer (2017) "The mechanism is truthful, why aren't you?" *American Economic Review*, Vol. 107, No. 5, pp. 220–224.

Hurwicz, Leonid, Eric Maskin, and Andrew Postlewaite (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*, pp. 367–433.

Kagel, John, Ronald Harstad, and Dan Levin (1987) "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study," *Econometrica*, Vol. 55, No. 6, pp. 1275–304.

Li, Jiangtao and Piotr Dworczak (2020) "Are Simple Mechanisms Optimal when Agents are Unsophisticated?" *Available at SSRN 3632971.*

Li, Shengwu (2017) "Obviously Strategy-Proof Mechanisms," *American Economic Review*, Vol. 107, No. 11, pp. 3257–3287.

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

Martínez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa (2019) "Failures in Contingent Reasoning: The Role of Uncertainty," *American Economic Review*, Vol. 109, No. 10, pp. 3437–3474.

Maskin, Eric (1999) "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, Vol. 66, pp. 23–38.

Munoz-Rodriguez, Edwin (2021) "Incentive Compatible Triaging in Deceased-Donor Transplantation."

Myerson, Roger (1981) "Optimal Auction Design," *Mathematics of Operations Research*, Vol. 6, No. 1, pp. 58–73.

Ortner, Juan (2020) "Bargaining with Evolving Private Information," *The 31st International Conference on Game Theory at Stony Brook*.

Papai, Szilvia (2000) "Strategyproof Assignment by Hierarchical Exchange," *Econometrica*, Vol. 68, No. 5, pp. 1403–1433.

——— (2001) "Strategyproof and Nonbossy Multiple Assignments," *Journal of Public Economic Theory*, Vol. 3, No. 3, pp. 257–271.

Pycia, Marek and Peter Troyan (2019) "A Theory of Simplicity in Games and Mechanism Design," *CEPR Discussion Paper No. DP14043*, pp. 1–74.

Pycia, Marek and Utku Ünver (2017) "Incentive compatible allocation and exchange of discrete resources," *Theoretical Economics*, Vol. 12, No. 1, pp. 287–329.

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

Roth, Alvin E., Tayfun Sönmez, and M. Utku Ünver (2004) "Kidney Exchange," *The Quarterly Journal of Economics*, Vol. 119, No. 2, pp. 457–488.

Saks, Michael and Lan Yu (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, Ran and Sándor Sóvágó (2018) "Obvious Mistakes in a Strategically Simple College Admissions Environment: Causes and Consequences," *Available at SSRN 2993538*.

Troyan, Peter (2019) "Obviously Strategy-proof Implementation of Top Trading Cycles," *International Economic Review*, Vol. 60, No. 3, pp. 1249–1261.

Zhang, Luyao and Dan Levin (2017) "Bounded rationality and robust mechanism design: An axiomatic approach," *American Economic Review*, Vol. 107, No. 5, pp. 235–239.