GRAPE Group for Research in APplied Economics

GRAPE Working Paper # 42

Are simple mechanisms optimal when agents are unsophisticated?

Jiangtao Li ® Piotr Dworczak

FAME | GRAPE, 2020

Foundation of Admirers and Mavens of Economics
Group for Research in Applied Economics

Are simple mechanisms optimal when agents are unsophisticated?

Jiangtao Li ®
School of Economics,
Singapore Management University

Piotr Dworczak®
Northwestern University,
FAME | GRAPE

Abstract
We study the design of mechanisms involving agents that have limited strategic sophistication. The literature has identified several notions of simple mechanisms in which agents can determine their optimal strategy even if they lack cognitive skills such as predicting other agents' strategies (strategy-proof mechanisms), contingent reasoning (obviously strategy-proof mechanisms), or foresight (strongly obviously strategy-proof mechanisms). We examine whether it is optimal for the mechanism designer who faces strategically unsophisticated agents to offer a mechanism from the corresponding class of simple mechanisms. We show that when the designer uses a mechanism that is not simple, while she loses the ability to predict play, she may nevertheless be better off no matter how agents resolve their strategic confusion.

Keywords:
simple mechanisms, complex mechanisms, robust mechanism design, dominant-strategy mechanisms, obviously strategy-proof mechanisms, strongly obviously strategy-proof mechanisms

JEL Classification
D71, D82, D86

Corresponding author
Piotr Dworczak, piotr.dworczak@northwestern.edu.

Acknowledgements
The authors' names are in certified random order, as described by Ray ® Robson (2018). The authors thank Mohammad Akbarpour, Tilman Börgers, Gabriel Carroll, Jeff Ely, Scott Duke Kominers, and Shengwu Li for helpful discussions. Li thanks the Department of Economics at Northwestern for hospitality and support during the academic year 2018-2019, when some of the work on this paper was undertaken.

Published by: FAME | GRAPE
ISSN: 2544-2473

Foundation of Admirers and Mavens of Economics
ull. Mazowiecka 11/14
00-052 Warszaw
Poland

W | grape.org.pl
E | grape@grape.org.pl
TT | GRAPE_ORG
FB | GRAPE.ORG
PH | +48 799 012 202

1 Introduction

It is widely accepted that "real-life" economic agents are not as rational as their counterparts in economic models. When agents have limited strategic sophistication, economists naturally lose confidence in the performance of mechanisms that require the participants to engage in complicated mental tasks. For example, achieving a Bayesian Nash equilibrium in a mechanism requires the agents to know the distribution of other agents' private information and correctly forecast other agents' play; this is why strategy-proof (SP) mechanisms are generally perceived as being superior for practical purposes.[1] Following the mounting evidence that even dominant strategies are difficult to identify for real-life agents, several recent papers have identified mechanisms in which agents can determine their optimal strategy under even weaker assumptions about their strategic sophistication. Li (2017) proposes the notion of obviously strategy-proof (OSP) mechanisms in which agents can determine their optimal strategy even if they cannot engage in contingent reasoning. Pycia and Troyan (2019) strengthen the notion of simplicity even further by relaxing the assumption that agents can predict their own future moves, and define (among other intermediate concepts) strongly obviously strategy-proof (SOSP) mechanisms.

For the purpose of this paper, we call a mechanism *simple* if, given the assumed level of strategic sophistication, agent can determine their optimal strategy. For example, if agents cannot predict other agents' play, we may be only comfortable assuming that they will not play a weakly dominated strategy. A SP mechanism is simple because a dominant strategy can be identified as the unique strategy (up to payoff equivalence) that is not weakly dominated for such agents. If the designer instead offers a mechanism that is not SP, she can no longer predict how agents will behave. More generally, we call a mechanism *complex* if it creates *strategic confusion* for the agents, understood as the inability to determine their optimal strategy in the mechanism.

The key observation of this paper is that the inability of the designer to predict the outcome of a complex mechanism need not be a sufficient reason for

[1] In the mechanism design context, it is more common to use the terminology "dominant-strategy mechanisms." We use the term "strategy-proof mechanisms" in this paper to draw a parallel to the notions of obviously strategy-proof mechanisms and strongly obviously strategy-proof mechanisms.

the use of simple mechanisms. As long as the designer is ultimately concerned with maximizing her payoff —which is typically assumed in mechanism design—in many cases complex mechanisms are unambiguously preferred by the designer to simple ones.

Since a complex mechanism, by definition, leads to a set of possible outcomes, we need to specify what we mean by the designer preferring a complex mechanism to a simple one. We analyze two notions. Under weak dominance, the complex mechanism generates a weakly higher expected payoff to the designer, no matter how agents resolve their strategic confusion (that is, no matter which strategy they choose from the set of strategies that they can identify as potentially optimal), and generates a strictly higher expected payoff in some cases. Under strong dominance, the designer obtains a strictly higher expected payoff in all cases, regardless of how agents behave when they are "confused" by the mechanism. To understand these notions, it is instructive to consider some examples. In these examples, we use the notion of OSP for illustration; however, the insights are robust to the particular choice of the notion of simplicity. In the paper, we demonstrate robustness to the notion of simplicity by also analyzing SP—the least demanding notion—and SOSP—the most demanding notion.

Example 1. Consider a single-unit auction with N bidders, with independent private values distributed according to a regular distribution. Suppose that the bidders cannot engage in contingent reasoning. As formalized and shown by Li (2017), bidders will not play an obviously dominated strategy, and thus agents' behavior (up to payoff equivalence) is pinned down if and only if the designer uses an OSP mechanism.[2] The revenue-maximizing mechanism—which requires implementing the outcome in which the object is allocated to the highest type as long as that type is above a distribution-dependent reserve price—can be OSP-implemented by the ascending clock auction. In the ascending clock auction, active bidders choose whether to exit as the clock price increases, and bidders who exit remain inactive thereafter. The auction stops when all but one bidder exit. The remaining bidder wins the object and pays the clock price.

[2] A strategy is obviously dominant if, for any deviation, at any information set where both strategies first diverge, the best outcome under the deviation is no better than the worst outcome under the dominant strategy. A mechanism is OSP if it has an equilibrium in obviously dominant strategies. See Section 2 for a formal definition.

However, consider the following modified mechanism, which we call the ascending clock auction with jump bidding, that differs from the ascending clock auction only in one aspect: Each bidder is allowed to speed up the clock by jump bidding, that is, making a higher bid than the current clock price.[3] This mechanism is not OSP because making a jump bid (to a bid b) is not obviously dominated for a bidder with value $v > b$ at the clock price $p < b$. Indeed, making a jump bid to b yields the best-case payoff of $v - b$ to the bidder, while following the default strategy (of exiting when the clock price reaches v) yields a payoff of 0 in the worst case.[4] Agents are strategically confused in that they now have multiple strategies that are not obviously dominated—the mechanism is complex.

Given the assumed strategic sophistication of the bidders, the designer could not predict whether jump bidding will occur or not. Still, a revenue-maximizing auctioneer might prefer the ascending clock auction with jump bidding to the ascending clock auction: If none of the bidders jump bids, then the performance of the ascending clock auction with jump bidding is the same as that of the ascending clock auction; in the event that some bidder (say bidder i) jump bids, the expected revenue of the ascending clock auction with jump bidding is strictly higher than that of the ascending clock auction. This is because there is positive probability that the highest valuation among the other bidders is between the current clock price and the jump bid of bidder i.[5] □

The key idea behind the above example is that agents are offered an additional option in the mechanism that—if taken—benefits the designer. This additional option would be recognized as unattractive by an agent with sufficient strategic sophistication; yet, it may be chosen by a less sophisticated agent. We generalize this insight to a wide variety of settings, and show that it applies even when there is a single agent.

When a complex mechanism weakly dominates the best simple mechanism,

[3] Jump bidding is a prevalent feature of real-world ascending auctions. Avery (1998) studies a common-value auction and shows that jump bidding may be employed to intimidate one's opponents. This signaling incentive is irrelevant in the private-value setting.

[4] Of course, bidders would not jump bid if they could engage in contingent reasoning.

[5] Although we consider the single-unit auction in this example, the logic immediately carries over to the class of binary allocation problems. Li (2017) shows that any OSP mechanism in the binary allocation problem is essentially a personal-clock auction. If the designer is maximizing revenue, then for any personal-clock auction, by allowing jump bidding, we obtain a complex mechanism that weakly dominates it.

the designer has a reason to purposefully introduce strategic confusion. However, since weak dominance only requires that the designer obtain a strictly higher payoff in some but not all cases, for a sufficiently pessimistic designer who believes that agents choose strategies that are the worst possible for her whenever they are confused, the difference between a simple mechanism and the complex mechanism that weakly dominates it may seem irrelevant.[6] The pessimism of the designer may be well-founded if the contingency in which her payoff is strictly higher under the complex mechanism seems implausible—even if not formally ruled out by the assumed rationality of agents.

However, this reservation is mute under the notion of strong dominance. We emphasize that the notion of strong dominance is remarkably strong, as it requires that the superior mechanism generate a strictly higher expected payoff to the designer, regardless of how agents behave when they are confused. The following example illustrates.

Example 2. Consider the problem of a trading platform that intermediates trade between two dealers and maximizes intermediation profits. Each dealer starts with no inventory, and can buy or (short) sell one unit of the asset. The platform cannot hold any inventory (ex-post market-clearing is imposed). In contrast to the single-unit auction example above, each dealer may become either a buyer or a seller, depending on the realization of the privately observed valuation and the choice of the mechanism.[7] Dealer A's valuation for the asset is either 0 or $2/3$. Dealer B's valuation for the asset is either $1/3$ or 1. The designer of the trading platform believes individual types to be equally likely but correlated across dealers: $\pi((0,\, 1/3)) = \pi((2/3,\, 1)) = \kappa > 1/4$.

The optimal OSP mechanism for $\kappa = 2/5$ yields an expected profit of $1/5$ for the platform and can be implemented by the following extensive-form game, which

[6] It may be useful to draw an analogy to the notion of weak dominance between two strategies in game theory. The arguments for and against the weakly dominated strategy carry over to the weakly dominated mechanism. In particular, the designer could still perceive a weakly dominated mechanism to be optimal if she *believes with certainty* that agents choose strategies that are worst possible for her whenever they are confused, just as a weakly dominated strategy could be perceived to be optimal as it could be a best response to a *degenerate* belief about her opponent's strategies.

[7] Intuitively, because the agent's role as the buyer or the seller is endogenously determined, the binding incentive constraints cannot be pinned down ex-ante. For related models with this feature, see for example Cramton et al. (1987), Lu and Robert (2001), Chen and Li (2018), and Loertscher and Marx (2020).

can be seen as a version of a personal-clock auction (see Li (2017)):

1. Dealer A is asked whether she would like to sell the asset at the price 0; if she says "yes," then that trade is implemented; if she says "no," then:

2. Dealer B is asked whether she would like to sell the asset at the price $1/3$; if she says "yes," then that trade is implemented; if she says "no," then there is no trade.

In all cases in which trade takes place, the platform charges a fee of $1/3$ to the buyer, that is, the buyer pays the buyer price plus $1/3$.[8]

It is obviously dominant for type 0 of dealer A to accept the initial offer, and for type $2/3$ to reject it. Similarly, it is obviously dominant for type $1/3$ of dealer B to accept the final offer, and for type 1 to reject it. It follows that the platform's profit is $1/3$ except when the type profile is $(2/3, 1)$. Intuitively, the inefficient no-trade outcome is implemented so that type 0 has an obviously dominant strategy: If dealer B were buying the asset from dealer A conditional on the profile $(2/3, 1)$, then the best possible outcome for type 0 from rejecting the initial offer yields a strictly positive payoff, while her equilibrium strategy yields a payoff of 0. For comparison, this inefficiency would be avoided if the solution concept were relaxed to the standard strategy-proofness: As we show in Appendix A.1, the optimal SP mechanism implements efficient trade and generates an expected profit of $4/15$ for the platform.

However, consider an alternative mechanism, which can be seen as a personal descending clock auction:

1. Dealer B is asked whether she would like to buy the asset at the price 1; if she says "yes," then that trade is implemented; if she says "no," then:

2. Dealer A is asked whether she would like to buy the asset at the price $2/3$; if she says "yes," then that trade is implemented; if she says "no," then:

3. Dealer B buys the asset at the price $1/3$.

[8]We show the optimality of this mechanism in Appendix A.1. The specific value of κ is not important for the result. For example, if $\kappa = 1/3$ (resp. $\kappa = 1/2$ (perfect correlation)), then the same OSP mechanism is optimal and generates an expected profit of $2/9$ (resp. $1/6$), leading to the same conclusion.

In all cases, the platform charges a fee of $1/3$ to the seller, that is, the seller receives the buyer price but pays a $1/3$ fee to the platform. This mechanism is not OSP: Type 1 of dealer B is confused because by accepting the initial offer she gets a payoff of 0; by rejecting, she receives a payoff of $-2/3$ or $2/3$, depending on the behavior of dealer A. However, regardless of how dealer B resolves her confusion, trade always happens, and the platform is guaranteed a profit of $1/3$ (hence it also receives $1/3$ in expectation, in fact, for any prior distribution of types). Finally, for each type, non-participation is obviously dominated.[9] Thus, as long as dealers do not play strategies that are obviously dominated, by adopting a complex mechanism, the platform beats not only the optimal OSP mechanism but also the optimal SP mechanism. □

While being dominated may be seen as an argument against simple mechanisms, it is not our point to criticize simplicity in general. For one thing, we do not take into account additional (mental or financial) costs that agents may incur when taking part in a complex mechanism relative to the costs of participation in a simple mechanism. Moreover, the flip side of our analysis is that we identify environments in which simple mechanisms are *not* weakly or strongly dominated. This may be seen as an optimality foundation for the use of simple mechanisms, complementary to existing reasons. For example, we show that, under some additional conditions, single-agent posted price mechanisms are not weakly dominated for the notions of SP and OSP (interestingly, this is not true under SOSP). We also show that the optimal simple mechanism is not strongly dominated when the designer's maximized objective function (under the chosen solution concept) is the same as the value of a relaxed problem in which the incentives constraints are only imposed on pairs of types that form a tree. A notable instance of such a setting for the notion of SP is when the the uniform shortest-path tree condition holds;[10] such environments include cases in which types are one-dimensional and single-crossing holds, e.g., single-unit auctions with private values.

Overall, the message of the paper is that one should not take the use of simple

[9]The platform could ensure that non-participation is obviously *strictly* dominated by lowering the fee by an arbitrarily small $\epsilon > 0$.

[10]See Chen and Li (2018). This result generalizes the insight of Yamashita (2015).

mechanisms for granted when agents are strategically unsophisticated; rather, their optimality should be carefully established for each environment in question.

1.1 Related literature

Simple mechanisms: This paper contributes to the design of mechanisms involving strategically unsophisticated agents. While Li (2017), Börgers and Li (2019), and Pycia and Troyan (2019) provide notions of simplicity and the characterizations of simple mechanisms according to these notions, our focus is on the tradeoff between simplicity and optimality, and we examine whether there is a foundation for the use of simple mechanisms from an optimality perspective.[11] Indeed, we show that in many cases, the designer might prefer a mechanism that is not simple.

Robust mechanisms design: Traditional models in mechanism design make strong assumptions about the detailed knowledge of the designer about the inputs to the mechanism design model. The literature of robust mechanism design seeks to relax these assumptions; see Carroll (2019) for a recent survey. While the leading interpretation of our exercise is that agents have limited strategic sophistication, an alternative interpretation—tying our work to the burgeoning literature of robust mechanism design—is that we relax the assumption about the designer's knowledge of the strategic reasoning process of the agents (beyond some minimal rationality assumptions).

SP mechanisms: Chung and Ely (2007) perform an exercise similar to ours in that, in the single-unit auction setting, they examine whether there is a foundation for the use of dominant-strategy mechanisms from an optimality perspective. The agents are sophisticated and play a Bayesian Nash equilibrium, but the auctioneer has uncertainty about the agents' hierarchies of beliefs about each other. They establish conditions under which the use of dominant-strategy mechanisms has a maxmin foundation. Börgers (2017) raises a criticism of the notion of the maxmin foundation: The optimal dominant-strategy mechanism—even with the maxmin

[11] Li (2017) and Pycia and Troyan (2019) formulate solution concepts that are stronger than the standard strategy proofness. In contrast, motivated by the observation that for many mechanism design problems, the class of SP mechanisms is quite small and only includes mechanisms that are rather unattractive for the designer, Börgers and Li (2019) propose a class of mechanisms—strategically simple mechanisms—that includes, but is strictly larger than, the set of dominant strategy mechanisms. There are, of course, many dimensions to simplicity. Our paper and the papers cited here primarily concern the strategic dimension.

foundation—could still be weakly dominated.

Implementation in undominated strategies: Our analysis for the solution concept of SP is closely related to the strand of the mechanism design literature that studies implementation in undominated strategies; see, for example, Börgers (1991), Jackson (1992), Börgers and Smith (2012), Carroll (2014), and Yamashita (2015). It is known from these papers that the optimal SP mechanism could be weakly or strongly dominated by complex mechanisms; we provide new examples and additional structural insights. Börgers (1991) shows that there are non-dictatorial procedures for collective decision making which ensure that collection decisions are Pareto efficient if all agents choose strategies that are not weakly dominated. Börgers and Smith (2012) show that, for bilateral trade and voting, there exist mechanisms that weakly dominate the optimal SP mechanism if all agents choose strategies that are not weakly dominated.[12] Yamashita (2015) develops a novel methodology of establishing an upper bound of the highest worst-case payoff for the designer, and applies the methodology to several settings, including the private-value auction and bilateral trade, and interdependent-value auction. Our results in Section 4.3 extend the insight of Yamashita (2015) to a larger class of environments and other solution concepts such as OSP and SOSP.

Complex mechanisms: In terms of understanding complex mechanisms with cognitively limited agents, Jakobsen (2020) is similar in spirit to our paper. He studies a mechanism design problem involving a principal and a boundedly rational agent, and shows that by expressing a mechanism as a complex contract, the principle could manipulate the agent into believing that truthful reporting is optimal. This paper differs from ours in that it imposes specific assumptions about how the agent resolves uncertainty for any given complex contract. In particular, the agent is modeled as a maxmin decision maker. In contrast, we take a robust approach, and do not impose any assumption about how agents resolve their strategic confusion for complex mechanisms.

[12] We are aware of two other papers that study questions different from ours but contain examples that could be used to show instances of strong dominance of an SP mechanism: Bergemann and Morris (2005)[Example 2] who study robust implementation; and Carroll (2014) who proves a complexity result for undominated-strategy implementation.

2 Preliminaries

Environment. There is a finite set of N agents, $\mathcal{N} = \{1, 2, \ldots, N\}$, and an arbitrary (possibly infinite) set of alternatives $\mathcal{X}$. Each agent i has payoff-relevant information indexed by $\theta_i \in \Theta_i$, where Θ_i is finite. We refer to θ_i as agent i's type. The set of possible type profiles is $\Theta = \times_{i \in \mathcal{N}} \Theta_i$ with representative element $\theta = (\theta_1, \theta_2, \ldots, \theta_N)$. The type profile is distributed according to a prior probability distribution $\pi \in \Delta\Theta$. As is standard, we write $\theta_{-i} \in \Theta_{-i} = \times_{j \neq i} \Theta_j$ for a type profile of agents other than agent i.

Each agent i is endowed with a utility function $u_i : \mathcal{X} \times \Theta_i \to \mathbb{R}$ (we assume private values). The designer has a utility function $v : \mathcal{X} \times \Theta \to \mathbb{R}$. That is, $u_i(x, \theta_i)$ and $v(x, \theta)$ denote type θ_i's utility and the designer's utility, respectively, when the type profile is θ and the implemented alternative is x. We assume that the designer is an expected utility maximizer with respect to the distribution π of types. We make no such assumption about the agents because how they form beliefs about other agents' types is irrelevant given the solution concepts we consider.

The designer may wish to use randomization in the mechanism. To incorporate this possibility, we introduce a "dummy" agent $i = 0$ with no preferences, called Nature. The game form is allowed to feature nodes at which Nature is called to play, and the distribution over strategies is picked by the designer as part of the mechanism. We use "bar" to denote the extended profiles that include Nature, for example, $\bar{\mathcal{N}} = \mathcal{N} \cup \{0\}$.

Mechanisms. We consider finite mechanisms that are imperfect-information, extensive-form games with perfect recall and consequences in $\mathcal{X}$.[13] The definition is standard; to shorten the exposition, we only introduce notation associated with a generic game Γ that we are going to use:

(1) $\mathcal{H}$ is a (finite) set of histories, with representative element h;

(2) $h_\emptyset$ is the initial (empty) history;

(2) $\subset$ is the precedence relation over histories;

[13] Finiteness is an important assumption: Jackson (1992) shows that inifinite mechanisms may implement virtually any decision rule in undominated strategies, relying on an infinite hierarchy of weakly dominated strategies with no dominant strategy "at the top."

(3) $\mathcal{Z}$ is the set of terminal histories, with representative element z;

(4) $g(z) \in \mathcal{X}$ is the outcome resulting from z;

(5) I_i denotes an information set of agent i;

(6) $A(I_i)$ is the set of actions available at information set I_i;

(7) A (pure) strategy S_i chooses an action $a \in A(I_i)$ at every information set I_i of agent $i \in \bar{\mathcal{N}}$, and $\mathcal{S}_i$ is the collection of all (pure) strategies for player i;[14]

(8) A strategy profile $\bar{S} = (S_0, S_1, ..., S_N)$ specifies a strategy for each player;

(9) $z(h, \bar{S})$ denotes the terminal history that results when we start at history h and play proceeds according to the strategy profile $\bar{S}$;

(10) $\pi_0 \in \Delta(\mathcal{S}_0)$ is a full-support probability distribution over Nature's strategies.[15]

We say that the information set I_i is on the path of play of strategy S_i if there exists $\bar{S}_{-i}$ and $h \in I_i$ such that $h \subset z(h_\emptyset, S_i, \bar{S}_{-i})$. Given two strategies S_i and S_i', we define $\beta(S_i, S_i')$ to be the set of information sets that are on the path of play of both S_i and S_i'. Under perfect recall, $I_i \in \beta(S_i, S_i')$ implies that S_i and S_i' choose the same actions at all information sets preceding I_i. If $I_i \in \beta(S_i, S_i')$ but S_i and S_i' choose different actions at I_i, then we call I_i an earliest point of departure for these two strategies. We let $\alpha(S_i, S_i')$ denote the set of all earliest points of departure for these two strategies. Finally, we let $\Phi(S_i, I_i) = \{S_i' : I_i \in \beta(S_i, S_i'), S_i(I_i) = S_i'(I_i)\}$ be the set of player i's strategies that agree with S_i at I_i and all information sets that precede it.

Solution concepts. We now define three solution concepts that we use throughout this paper.

Definition 1 (SP)**.** *S_i is dominated (or SP-dominated) for type θ_i of agent i if there exists another strategy S_i' such that for all $\bar{S}_{-i}$,*

$$u_i(g(z(h_\emptyset, S_i, \bar{S}_{-i})), \theta_i) \leq u_i(g(z(h_\emptyset, S_i', \bar{S}_{-i})), \theta_i),$$

[14]It suffices for our purposes to focus on pure strategies; mixed strategies could be incorporated without changing any of our results.

[15]That π_0 has full support over Nature's strategies is equivalent to assuming that agents ignore zero-probability events as possible contingencies; in the opposite case, certain paradoxical results can be obtained (e.g., a strategy that gives an agent the worst possible outcome with probability one is weakly undominated if it offers the best possible outcome following a zero-probability event).

with the inequality being strict for some $\bar{S}_{-i}$.

Definition 2 (OSP)**.** *S_i is obviously dominated (or OSP-dominated) for type θ_i of agent i if there exists another strategy S'_i such that for all $I_i \in \alpha(S_i, S'_i)$,*

$$\sup_{h \in I_i, \bar{S}_{-i}} u_i(g(z(h, S_i, \bar{S}_{-i})), \theta_i) \leq \inf_{h \in I_i, \bar{S}_{-i}} u_i(g(z(h, S'_i, \bar{S}_{-i})), \theta_i),$$

with the inequality being strict for some $I_i \in \alpha(S_i, S'_i)$.

Definition 3 (SOSP)**.** *S_i is strongly obviously dominated (or SOSP-dominated) for type θ_i of player i if there exists another strategy S'_i such that for all $I_i \in \alpha(S_i, S'_i)$,*

$$\sup_{h \in I_i, \bar{S}_{-i}, R_i \in \Phi(S_i, I_i)} u_i(g(z(h, R_i, \bar{S}_{-i})), \theta_i) \leq \inf_{h \in I_i, \bar{S}_{-i}, R'_i \in \Phi(S'_i, I_i)} u_i(g(z(h, R'_i, \bar{S}_{-i})), \theta_i),$$

with the inequality being strict for some $I_i \in \alpha(S_i, S'_i)$.

In words, a strategy S_i is dominated for type θ_i if there exists another strategy S'_i that yields a higher payoff for type θ_i for any fixed strategy profile for other players and Nature. A strategy S_i is obviously dominated for type θ_i if there exists another strategy S'_i such that, starting at any earliest point of departure, the worst possible payoff under S'_i for type θ_i across all strategies of other players and Nature is higher than the best possible payoff under S_i for type θ_i across all strategies of other players and Nature. Finally, a strategy S_i is strongly obviously dominated for type θ_i if there exists another strategy S'_i such that, starting at any earliest point of departure, the worst possible payoff under S'_i for type θ_i across all strategies of other players, Nature, and moves of player i at future information sets is higher than the best possible payoff under S_i for type θ_i across all strategies of other players, Nature, and moves of player i at future information sets. The three dominance concepts are nested: Strongly obviously dominated strategies are obviously dominated, and obviously dominated strategies are dominated. From now on, we fix a dominance concept $K \in \{$SP, OSP, SOSP$\}$; unless explicitly stated otherwise, all statements involving a generic dominance concept K apply to any K in that set.

Remark 1 (On randomization)**.** Under the solution concepts analyzed in this paper, one has to take a stance on how agents reason about randomization: When evaluating a strategy, they could either *(i)* take expectations with respect to designer's randomization within the mechanism, or *(ii)* condition on each outcome of the randomization device. Our definitions capture the second possibility. This implies that Definition 1 of dominance for strategy-proof mechanisms is somewhat non-standard.[16] However, this convention keeps the three solution concepts more analogous, since in OSP (see Li (2017)) and SOSP (see Pycia and Troyan (2019)), dominance is defined with respect to Nature's moves. Such an approach to randomization seems natural given the focus on simplicity, since randomization may plausibly be seen as making reasoning in the mechanism more difficult for players; an opposite view would implicitly assume that—in the mind of a player—there is a distinction between the moves of Nature and the moves of a strategic player with a constant utility function. That being said, since we imposed no a priori restrictions on the space of primitive outcomes $\mathcal{X}$, it is possible that $\mathcal{X}$ already contains lotteries over deterministic decisions (and the utilities $u_i(x, \theta_i)$ are expected utilities derived from some primitive utility for deterministic decisions); in that case, randomization via Nature's moves is redundant and our definition of strategy-proof mechanisms essentially reduces to a standard one. Thus, our stronger notion of strategy-proofness only has bite when we explicitly assume that $\mathcal{X}$ does not contain lotteries (which we will do for some applications).

For any type θ_i of agent $i \in \mathcal{N}$, we can think of the dominance relation as defining a partial order $\prec_{\theta_i}^K$ on the set of player i's strategies: $S_i \prec_{\theta_i}^K S_i'$ if S_i is K-dominated by S_i' according to θ_i. We call two strategies S_i and S_i' payoff-equivalent for type θ_i if $u_i(g(z(h_\emptyset, S_i, \bar{S}_{-i})), \theta_i) = u_i(g(z(h_\emptyset, S_i', \bar{S}_{-i})), \theta_i)$ for all $\bar{S}_{-i}$. A strategy S_i is K-dominant (for some type θ_i) if all non-payoff-equivalent strategies S_i' are K-dominated by it.

Strategic confusion, simple and complex mechanisms. Throughout, we adopt the following assumption about the rationality of players.

[16]However, we are not first to consider this extension of the standard notion. A (randomized) mechanism is called a universally-truthful mechanism—a common solution concept in the computer science literature, see for example Nisan and Ronen (2001)—if every mechanism in the support of the randomized mechanism is a dominant-strategy mechanism. The mechanism is truthful even when the realization of the random coins is known.

Assumption 1 (Strategic sophistication)**.** *For any $i \in \mathcal{N}$, no type θ_i of player i plays a K-dominated strategy.*

Assumption 1 is a relatively weak assumption on behavior. In particular, we do not assume any degree of knowledge of rationality—players are not assumed to engage in iterative elimination of dominated strategies.

For mechanism Γ with the corresponding set of possible strategies $\mathcal{S}_i$ of player i, we define the set of strategies that type θ_i can possibly play under Assumption 1:

$$U_i^K(\theta_i) = \{S_i \in \mathcal{S}_i : \nexists S_i' \in \mathcal{S}_i,\ S_i \prec_{\theta_i}^K S_i'\}.$$

That is, $U_i^K(\theta_i)$ is the set of K-undominated strategies for type θ_i. Given the finiteness of the mechanism, that set is non-empty.

If type θ_i has a K-dominant strategy, then the set $U_i^K(\theta_i)$ is not necessarily a singleton; however, all strategies in $U_i^K(\theta_i)$ must be payoff-equivalent for θ_i. In this case, the player is truly indifferent between these dominant strategies, and hence we do not want to label this as "strategic confusion" (otherwise, no solution concept could avoid strategic confusion). However, these dominant strategies need not be payoff-equivalent for the designer. It is customary in mechanism design to let the designer select which dominant strategy the player should play, and we follow the same convention here.[17] Formally, we will treat $U_i^K(\theta_i)$ as consisting of equivalence classes of strategies that are payoff-equivalent for θ_i, and we let the designer pick the representative of each equivalence class (we leave this description verbal in order not to further complicate our notation).

Definition 4 (Strategic confusion and complex mechanism)**.** *Fixing a mechanism Γ, type θ_i of player $i \in \mathcal{N}$ is said to be strategically confused (under solution concept K) if $U_i^K(\theta_i)$ contains at least two (not payoff-equivalent) strategies. In such case, we call mechanism Γ complex (for type θ_i of agent i).*

Strategic confusion means that player i with type θ_i has more than one K-undominated strategy in the mechanism, and therefore, in the absence of further

[17] Various justifications can be offered for this assumption. One is that the mechanism is paired with "recommended" strategies for the players, and the players follow the recommendation as long as it constitutes a dominant strategy. Another one is that—under some permissive conditions—it is possible to perturb the mechanism slightly to guarantee that the dominant strategy preferred by the designer is the unique dominant strategy for each type. A pragmatic justification is that optimal mechanism would often fail to exist otherwise.

assumptions on behavior or strategic reasoning, it is impossible to determine which strategy she will select. Of course, a valid interpretation is that player i "knows" which strategy to choose but the designer is not willing to make an assumption about that choice.

The literature has identified classes of "simple" mechanisms that can be played even if agents are strategically unsophisticated, in the sense of only satisfying Assumption 1. When players cannot forecast their opponents' play, they will nevertheless know how to behave if they have a dominant strategy (a defining property of SP mechanisms). If, moreover, players cannot engage in contingent reasoning, they should be offered an obviously dominant strategy (leading to the notion of OSP mechanisms, see Li (2017)). Finally, if players lack foresight and cannot even predict their own future moves, they should have a strongly obviously dominant strategy (leading to the notion of SOSP mechanisms, see Pycia and Troyan (2019)). This motivates our definition of a simple mechanism.[18]

Definition 5 (Simple mechanism)**.** *A mechanism Γ is simple if for any agent $i \in \mathcal{N}$, no type θ_i is strategically confused.*

Weak and strong dominance of mechanisms. An advantage of a simple mechanism from the point of view of the designer is that she can predict how agents will behave. In contrast, if any agent is strategically confused, the designer—based only on Assumption 1—cannot determine the path of play. This seems to provide a strong argument in favor of simple mechanisms. However, that benefit is diminished if the designer can achieve better outcomes using a mechanism that confuses some types. In this work, we study two notions of "better outcomes" that we introduce below.

Let $\boldsymbol{S}_i$ denote a type-strategy for player i, that is, $\boldsymbol{S}_i(\theta_i)$ is the strategy selected by type θ_i of player $i \in \mathcal{N}$. We let $\boldsymbol{S}_i \subset U_i^K$ mean that $\boldsymbol{S}_i$ is a selection from the correspondence of K-undominated strategies, i.e., $\boldsymbol{S}_i(\theta_i) \in U_i^K(\theta_i)$ for all $\theta_i \in \Theta_i$. Define a correspondence

$$V(\Gamma) = \mathrm{CH}\left(\left\{\mathbb{E}_{\theta\sim\pi,\, S_0\sim\pi_0}\left[v(g(z(h_\emptyset,\, (S_0,\, \boldsymbol{S}_i(\theta_1), \ldots,\, \boldsymbol{S}_i(\theta_N)))),\, \theta)\right] : \{\boldsymbol{S}_i \subset U_i^K\}_{i\in\mathcal{N}}\right\}\right)$$

[18] We emphasize that we only view this definition as being appropriate in the context of the question that we study. There are various other notions of simplicity in mechanism design that are orthogonal to the issue at hand.

which is the range of designer's expected payoffs over all possible, potentially randomized, ways in which confused types can resolve their strategic confusion. Note, however, that when computing the range of designer's payoffs we assume that which undominated strategy a player selects cannot depend on the realization of other players' types. By definition, $V(\Gamma)$ is a singleton when Γ is a simple mechanism but it may be an interval when Γ is complex.

Definition 6 (Weak dominance)**.** *A mechanism* Γ *is weakly dominated if there exists a mechanism* Γ' *such that*

$$\sup V(\Gamma) \leq \inf V(\Gamma') \text{ and } \sup V(\Gamma) < \sup V(\Gamma').$$

Definition 7 (Strong dominance)**.** *A mechanism* Γ *is strongly dominated if there exists a mechanism* Γ' *such that*

$$\sup V(\Gamma) < \inf V(\Gamma').$$

A mechanism Γ is weakly dominated by a mechanism Γ' if the expected payoff for the designer under Γ' is at least as large as the expected payoff under Γ, regardless of how confused types select their strategies; moreover, the expected payoff under Γ' is strictly larger under some selection. A mechanism Γ is strongly dominated by a mechanism Γ' if the expected payoff for the designer under Γ' is strictly larger than the expected payoff under Γ, regardless of how confused types select their strategy.

Although the above definitions are general, we will only apply them to the case when Γ is a simple mechanism. Moreover, we will focus on cases when Γ is an *optimal* simple mechanism (that is, it maximizes the designer's expected payoff among all simple mechanisms), in which case it can only be weakly or strongly dominated by a complex mechanism.

Participation constraints. In many settings (including the ones in which the designer's objective is to maximize her revenue), it is necessary to impose a participation constraint for the problem to be well defined. For a simple mechanism, under the solution concepts in question, it is natural to impose ex-post individual-rationality which requires that for any player i and type $\theta_i \in \Theta_i$, the chosen strategy

$S_i = \boldsymbol{S}_i(\theta_i)$ yields a non-negative payoff in all possible cases:

$$u_i(g(z(h_\emptyset, S_i, \bar{S}_{-i})), \theta_i) \geq 0, \quad \forall \bar{S}_{-i}.^{19} \tag{1}$$

We introduce two extensions of this condition to complex mechanisms. We say that a mechanism Γ provides *partial incentives to participate* if for all $i \in \mathcal{N}$, all $\theta_i \in \Theta_i$, there *exists* $S_i \in U_i^K(\theta_i)$ such that condition (1) holds. We say that a mechanism Γ provides *full incentives to participate* if for all $i \in \mathcal{N}$, all $\theta_i \in \Theta_i$, *for all* $S_i \in U_i^K(\theta_i)$ condition (1) holds. The notion of partial incentives to participate is more appropriate when non-participation is thought of as a strategy: The condition states that the non-participation strategy is (obviously) dominated, and hence is not chosen by a rational player. The notion of full incentives to participate is more appropriate if non-participation is thought of as the option for each player to walk away from the mechanism at any point, including after learning her final payoff.

3 Weak dominance

In this section, we analyze the concept of weak dominance of mechanisms. The overall insight is that simple mechanisms are weakly dominated by complex mechanisms in many environments. We demonstrate this by first formulating and proving an abstract result that shows conditions under which a simple mechanism can be turned into a complex mechanism that weakly dominates it by giving one of the players an additional strategic option that—if taken—benefits the designer. Then, we apply this abstract result in various settings. Throughout, we fix a solution concept $K \in \{\text{SP, OSP, SOSP}\}$. Unless stated otherwise, our results hold for any such K.

Let Γ be a simple mechanism and let $\mathcal{Y} \subseteq \mathcal{X}$. Define

$$\Theta_i^{\mathcal{Y}} = \left\{\theta_i \in \Theta_i : \max_{x \in \mathcal{Y}} u_i(x, \theta_i) > \min_{\bar{S}_{-i}} u_i(g(z(h_\emptyset, \boldsymbol{S}_i(\theta_i), \bar{S}_{-i})), \theta_i)\right\}.$$

[19] This implies that strategy S_i obviously dominates non-participation, but not that it strongly obviously dominates non-participation. For SOSP, one may wish to require this condition also for all strategies $R_i \in \Phi(S_i, I_i)$, where I_i is the first information set of player i, for each possible history (this would not change our results in any substantial way).

That is, $\Theta_i^{\mathcal{Y}}$ is the set of types of agent i that strictly prefer some outcome in $\mathcal{Y}$ to the worst possible outcome in the simple mechanism Γ.

Proposition 1. *Fix a simple mechanism Γ. Suppose that for any agent $i \in \mathcal{N}$,*[20] *there exists $\mathcal{Y} \subseteq \mathcal{X}$ and a simple mechanism $\Gamma_{-i}^{\mathcal{Y}}$ played by agents $-i$ with an outcome space $\mathcal{Y}$ such that*

1. *Each outcome $x \in \mathcal{Y}$ occurs at some terminal node in $\Gamma_{-i}^{\mathcal{Y}}$;*
2. *For any $\theta_i \in \Theta_i^{\mathcal{Y}} \neq \emptyset$, the designer prefers (strictly for some $\theta_i \in \Theta_i^{\mathcal{Y}}$) the conditional expected payoff from the mechanism $\Gamma_{-i}^{\mathcal{Y}}$ to the conditional expected payoff from the mechanism Γ.*

Then, the mechanism Γ is weakly dominated.

Proof. For each solution concept $K \in \{\text{SP, OSP, SOSP}\}$, we show that the simple mechanism is weakly dominated by explicitly constructing a complex mechanism that dominates it. The mechanism we construct can be interpreted as a delegation mechanism in which one agent is delegated to choose a simple mechanism (for the agents to play) from two simple mechanisms specified by the designer.[21] We present the proof for the solution concept of SP here; the remaining two cases are proven similarly but require certain adjustments (see Appendix A.2 for details).

Fix a player i. We add a new node for player i from which play begins. Player i chooses from two options: "Choose Γ" or "Choose $\Gamma_{-i}^{\mathcal{Y}}$." If she chooses Γ, the game tree is the one associated with Γ. If she chooses $\Gamma_{-i}^{\mathcal{Y}}$, the game tree is the one associated with the game $\Gamma_{-i}^{\mathcal{Y}}$. We call this new composite game Γ'.

First, we claim that all players other than i have a dominant strategy in Γ'. This is immediate from the fact that each player $-i$ has a dominant strategy in Γ and a dominant strategy in $\Gamma_{-i}^{\mathcal{Y}}$. Second, all types of player i not in $\Theta_i^{\mathcal{Y}}$ also have a dominant strategy which is to choose Γ, and then follow the same strategy $\boldsymbol{S}_i(\theta_i)$ that was dominant for Γ. This is immediate from the fact that for these types, the best that can happen after choosing $\Gamma_{-i}^{\mathcal{Y}}$ cannot be better than the worst possible outcome in Γ when they follow $\boldsymbol{S}_i(\theta_i)$. Third, we claim that for

[20] When $K \in \{\text{SP, OSP}\}$, it suffices that one such i exists; however, most of our example are symmetric so this extension is not important for our purposes.

[21] These complex mechanisms are "type 1 strategically simple" according to the notion in Börgers and Li (2019).

all types $\theta_i \in \Theta_i^{\mathcal{Y}}$, the option to choose $\Gamma_{-i}^{\mathcal{Y}}$ is SP-undominated. Indeed, fix the strategy profile $\bar{S}_{-i}$ that yields the minimum in the definition of $\Theta_i^{\mathcal{Y}}$, and—using condition 1 in the proposition—let $\bar{S}_{-i}^{\mathcal{Y}}$ be the profile that leads to the outcome $x^\star \in \text{argmax}_{x \in \mathcal{Y}} u_i(x, \theta_i)$ in the game $\Gamma_{-i}^{\mathcal{Y}}$. Then, if players $-i$ follow the strategy $(\bar{S}_{-i}, \bar{S}_{-i}^{\mathcal{Y}})$ in Γ', by definition of $\Theta_i^{\mathcal{Y}}$, the best response for type θ_i is to choose the game $\Gamma_{-i}^{\mathcal{Y}}$ at the first decision node.

Overall, it is an undominated strategy for types $\theta_i \in \Theta_i^{\mathcal{Y}}$ to choose the game $\Gamma_{-i}^{\mathcal{Y}}$, and when they do, play among players $-i$ in that subgame proceeds as in the original game $\Gamma_{-i}^{\mathcal{Y}}$. By condition 1 in the proposition, the designer receives a higher (sometimes strictly) conditional expected payoff in that case, compared to the conditional expected payoff she would have received in the game Γ. Therefore, the game Γ is weakly dominated by Γ'. □

The idea behind the proof is simple. Given some initial simple mechanism Γ, agent i is offered the "concession option" that guarantees herself an outcome in $\mathcal{Y}$ but at the cost of being excluded from further play. If this option is not chosen, play proceeds as in the original mechanism Γ. The key difference between a player and the designer when evaluating the concession option is that the player is only assumed to avoid K-dominated strategies while the designer is an expected-payoff maximizer. Therefore, the agent will not rule out a strategy that gives her a very high payoff in some case, no matter how small the probability of that outcome. The designer exploits that and structures the set $\mathcal{Y}$ and the mechanism $\Gamma_{-i}^{\mathcal{Y}}$ in a way that gives her a high payoff on average, while guaranteeing at least one good outcome for player i, with small probability.

Next, we investigate how the proposition can be applied in various settings.

3.1 Efficiency in bilateral trade

We first consider the classical problem of designing an efficient mechanism that allows a seller of an indivisible object to trade with one potential buyer. Both agents have quasi-linear preferences. It is known from Hagerty and Rogerson (1987) that the only SP mechanisms (and hence also OSP and SOSP mechanisms) that satisfy ex-post budget balance and individual rationality are posted price mechanisms. In a posted price mechanism, the designer chooses a (possibly random) price, without taking into account any of the agents' private information. The outcome depends on agents' private information only through their decision to trade, or not to trade, at the price proposed by the designer. Trade comes about only when both agents agree. Thus, full efficiency is not implementable. In what follows, we will argue—based on Proposition 1—that there exists a complex mechanism that weakly dominates the simple mechanism in terms of the expected total surplus.[22]

We will verify the assumptions of Proposition 1 with i being the seller (an analogous argument works for the buyer). Fixing Γ, let p be the posted price.[23] Let the set $\mathcal{Y}$ contain two options: *(1)* No trade and *(2)* Trade at price $p - \epsilon$ for some $\epsilon > 0$. The mechanism $\Gamma^{\mathcal{Y}}_{-i}$ consists in the buyer choosing one option from $\mathcal{Y}$. Clearly, that mechanism is simple for the buyer. Condition 1 in Proposition 1 is trivially satisfied. $\Theta^{\mathcal{Y}}_i$ contains all types of the seller strictly below $p - \epsilon$. Finally, the second condition in Proposition 1 is satisfied as long as the buyer chooses option (2) with positive probability which is when her type is above $p - \epsilon$. Hence, as long as the type spaces for the buyer and the seller are not degenerate, we can find ϵ for which the simple mechanism Γ is weakly dominated.

Intuitively, the complex mechanism constructed in the proof of Proposition 1 can be viewed as a "price cap mechanism." The designer sets a price cap of p but the seller has the option to lower the price to $p - \epsilon$. This leads to strategic confusion of some seller types but no matter how these types behave, total surplus can only improve. Note that this mechanism provides full incentives to participate for both the seller and the buyer.

3.2 Revenue maximization

Next, we argue that Proposition 1 applies to any standard revenue-maximization problem with quasi-linear utilities. Let X be the space of possible allocations (which could be random as well), and define $\mathcal{X} = X \times \mathbb{R}^N$, with $(y, t_1, ..., t_N)$ interpreted as an outcome in which allocation y is implemented, and player i pays the designer t_i. We have $u_i((y, t_1, ..., t_N), \theta_i) = \tilde{u}_i(y, \theta_i) - t_i$, for some arbitrary $\tilde{u}_i(y, \theta_i)$ assumed non-negative and non-constant in θ_i. The designer maximizes

[22] Börgers and Smith (2012) show this result for the solution concept of SP.

[23] If that price is random, then we fix p to be the realization that yields the highest conditional expected payoff for the designer—this can only increase the expected surplus in the mechanism Γ.

expected revenue.

We will consider two cases: In the first, the designer is satisfied with a mechanism that provides partial incentives to participate; we will show that this leads to the best simple mechanism being weakly dominated in a particularly blatant way, with revenue that is unbounded in the best case for the designer. In the second, the superior mechanism will provide full incentives to participate.

Weak dominance under partial incentives to participate. Fix a simple mechanism Γ, with some expected revenue R. Suppose first that $N \geq 2$. For any i, define $\mathcal{Y} \subset \mathcal{X}$ to contain two outcomes: *(1)* Player i pays M to the designer while some player j receives $\epsilon > 0$ from the designer, and *(2)* Player i receives M from the designer while player j receives 0 from the designer. The mechanism $\Gamma^{\mathcal{Y}}_{-i}$ is defined as having just one information node for player j who chooses between the two options in $\mathcal{Y}$. This satisfies condition 1 of Proposition 1. Moreover, this mechanism is simple because player j has a K-dominant strategy to select option *(1)*. When M is large enough, $\Theta_i^{\mathcal{Y}} = \Theta_i$, and when additionally ϵ is small enough, condition 2 of Proposition 1 holds, since the conditional expected payoff from the mechanism $\Gamma^{\mathcal{Y}}_{-i}$ is unbounded in M. Thus, by Proposition 1, the simple mechanism Γ is weakly dominated. It remains to check that the weakly-dominating complex mechanism provides partial incentives to participate. This is immediate from the proof of Proposition 1 that explicitly constructs the weakly-dominating complex mechanism Γ': Intuitively, in Γ', player i chooses between the games Γ and $\Gamma^{\mathcal{Y}}_{-i}$; thus, since each type of player i had a K-dominant strategy satisfying (1) in Γ, each type continues to have at least one K-undominated strategy satisfying (1) in Γ'.

This construction can also be adapted to the case $N = 1$. Let $\mathcal{Y} \subset \mathcal{X}$ contain two outcomes: *(1)* Player i pays M to the designer, and *(2)* Player i receives M from the designer. The mechanism $\Gamma^{\mathcal{Y}}_{-i}$ is defined as having just one information node for Nature that chooses option *(1)* with probability $1 - \epsilon$ and option *(2)* with probability ϵ. For ϵ sufficiently small, by Proposition 1 and the same arguments as before, the mechanism Γ is weakly dominated.

It is known (see for example Ashlagi and Gonczarowski (2018) and Pycia and Troyan (2019)) that randomization cannot increase the designer's payoff within the class of simple mechanisms.[24] Interestingly, the above example shows that this does not imply that the designer should never randomize when facing unsophisticated agents. On the contrary, randomization can be used to purposefully confuse the agents in order to obtain a superior outcome, at least in some cases.

The above weakly-dominating mechanisms may seem stark and unlikely to "work" in practice. The construction for the case $N \geq 1$ relies on the fact that agents are not assumed to engage in iterative elimination of dominated strategies; the construction for the case $N = 1$ relies on the fact that agents are not assumed to evaluate lotteries by computing their expected payoffs. Under the alternative interpretation of robust mechanism design, the designer is not willing to make such assumptions. We offer two comments: First, similarly to how many results are interpreted in mechanism design, we view the value of these examples as illustrating the possibility that a simple mechanism may be dominated. Our construction is optimized for making the mathematical argument concise; there may exist more "subtle" ways to weakly dominate a simple mechanism. Second, anecdotally,[25] sellers frequently offer seemingly unattractive options to customers hoping to exploit their potential inability to rank these options as inferior.

Weak dominance under full incentives to participate. To simplify exposition, we consider the classical problem of allocating a single object to one of N ex-ante identical bidders. Note that the ascending clock auction with jump-bidding in Example 1 in the introduction provides full incentives to participate under OSP. This is because if type $v \geq b$ makes a jump bid to b at price $p < b$ (jump-bidding is dominated for all other types), then she gets a non-negative payoff after every possible history. Obviously, her payoff is also non-negative if she does not jump-bid. We generalize this example by applying Proposition 1.

Claim 1. *Suppose that $N \geq 2$, and let $\bar{u}$ be the highest possible valuation for the object. If it is not a revenue-maximizing simple mechanism to sequentially offer the object at a price of $\bar{u}$ to all the players, then the optimal simple mechanism is weakly dominated.*

The proof can be found in Appendix A.3. The idea is that the designer

[24]For SP mechanisms, this is a consequence of our Definition 1, see Remark 1.

[25]See, for example, Chernev et al. (2015) for a survey of results on choice overload.

can weakly dominate the payoff from a simple mechanism by approaching one player at the beginning of the game and offering her the best possible allocation at two potential prices: *(i)* the highest possible valuation for that allocation, or *(ii)* the highest possible valuation minus some $\delta > 0$; which of the two prices is implemented depends on the choice of some other player who is given incentives to always choose the higher price (similarly as in the previous construction). Accepting the designer's offer is undominated for the highest type (and only the highest type if δ is small enough) as long as it is possible that the she gets a payoff of 0 in some history in Γ (which is guaranteed by symmetry). At the same time, her payoff is non-negative in all possible histories no matter how she behaves, so full incentives to participate are provided. Finally, the assumption of Claim 1 guarantees that the designer was extracting less than $\bar{u}$ from the highest type, so if the highest type accepts the offer, the designer receives a strictly higher revenue.

The assumption of Claim 1 provides an example of a setting in which the optimal simple mechanism is *not* weakly dominated: If all players have a value of $\bar{u}$ for the object with probability one, then sequentially offering the object at a price of $\bar{u}$ to all the players is an optimal simple mechanism that is not weakly dominated under the full incentives to participate (for it to be dominated, there would have to exist an on-path history in which some player is charged more than $\bar{u}$ but that would be incompatible with full incentives to participate).

Based on the above, one may hope that single-player posted price mechanisms are not weakly dominated under full incentives to participate. This is not always the case. Consider the notion of SOSP and let $p^\star$ be the revenue-maximizing price under π. Suppose that $p^\star$ is not equal to the highest possible type under π. Then, a "one-person descending clock auction" with a minimum price $p^\star$ weakly dominates the posted price mechanism while providing full incentives to participate.[26] Indeed, all types $\theta < p^\star$ have an SOSP-dominant strategy to never buy (this strategy gives a non-negative payoff). For types $\theta \geq p^\star$, it is SOSP-dominant to stay in the auction when the price is weakly above θ, and SOSP-undominated to buy at any price strictly below θ (again, in all cases, the payoff is non-negative). Clearly, the designer can only be weakly better off, and she is strictly better off whenever the

[26]This design is formally a sequential posted price mechanism in which the designer starts from a price equal to the highest type and drops the price by some sufficiently small ϵ in every round, with the last price equal to $p^\star$.

highest type decides to buy at a price strictly between $p^\star$ and her value (which is offered in some round of the auction if the price grid is fine enough).

However, we do get a positive result for the notions of SP and OSP.

Claim 2. *Suppose that $N = 1$, and that there exists a unique optimal simple mechanism (in which case it must be a posted price mechanism). Then, under SP and OSP, that mechanism is not weakly dominated.*

The proof is relatively involved, so we relegate it to Appendix A.4. Moreover, in the Supplemental Material OA.1, we show that the assumption that the optimal simple mechanism is unique (which is satisfied for a generic distribution π) cannot be dispensed with—we construct an example in which the optimal simple mechanism is not unique and we show that it is weakly dominated.

3.3 Voting

While Proposition 1 has the most power in environments with money, we construct one example in a setting with non-transferable utility. There are two agents and three alternatives: $\mathcal{X} = \{a, b, c\}$. Each agent's type space consists of all six possible rankings. Suppose that the designer's payoff is 0 if at least one player gets her worst option or if players get options that they rank differently; it is 1 otherwise (that is, when both players get their first choices, or both players get their second choices). The best optimal mechanism is dictatorship with full range; without loss of generality, player $-i$ chooses her most preferred alternative.[27]

We will show that this mechanism is weakly dominated. The idea is that player i (the non-dictator) has the option to eliminate alternative a from the consideration set of the dictator. Formally, let $\mathcal{Y} = \{b, c\}$ and the mechanism $\Gamma^{\mathcal{Y}}_{-i}$ is that player $-i$ chooses the best alternative in $\mathcal{Y}$. We have $\Theta^{\mathcal{Y}}_i = \Theta_i \setminus \{abc, acb\}$. Condition 1 of Proposition 1 holds, so it is enough to verify condition 2. The outcome of $\Gamma^{\mathcal{Y}}_{-i}$ is different only when player $-i$ would have chosen a in Γ. But for types of i in $\Theta^{\mathcal{Y}}_i$, this means that—whenever the two outcomes potentially differ—the designer's payoff is 0 in Γ. Thus, she gets weakly more in $\Gamma^{\mathcal{Y}}_{-i}$ conditional

[27] Again, by symmetry, we can assume that the mechanism is symmetric, and then it is enough to check the assumptions of Proposition 1 for one player.

on any $\theta_i \in \Theta_i^{\mathcal{Y}}$. To see that she gets sometimes strictly more, note that when $\theta_i = bca$ and $\theta_{-i} = acb$, the designer gets 1 in $\Gamma_{-i}^{\mathcal{Y}}$.

4 Strong dominance

This section considers strong dominance of simple mechanisms. Section 4.1 and Section 4.2, respectively, show that if *(i)* a property that we call accommodation of additional types is violated or *(ii)* there is richness in the agent's preference, then there exists a payoff function of the designer such that the best simple mechanism is strongly dominated. Section 4.3 provides a sufficient condition such that the optimal simple mechanism is not strongly dominated. Throughout, we fix a solution concept $K \in \{\text{SP, OSP, SOSP}\}$. Unless stated otherwise, our results hold for any such K.

4.1 Accommodation of additional types

This subsection introduces a property that we call the accommodation of additional types (AAT). We show that if the AAT property is violated, there exists a payoff function for the designer such that the best simple mechanism is strongly dominated.

Fixing the type space Θ, distribution π, players' preferences, and $K \in \{\text{SP, OSP, SOSP}\}$ (the "environment"), we say that an outcome $\lambda : \Theta \to \Delta(\mathcal{X})$ is *simply-implementable* if there exists a simple mechanism Γ whose equilibrium leads to the outcome $\lambda(\theta)$ whenever the realized type profile is θ. We let $\Lambda(\Theta)$ denote the set of simply-implementable outcomes on the type space Θ. We denote by $\Theta \setminus \{\theta_i\}$ the type space $\Theta_1 \times \ldots \times \Theta_i \setminus \{\theta_i\} \times \ldots \times \Theta_N$.

Definition 8. *We say that the environment has the accommodation of additional types (AAT) property if for any i, any $\theta_i \in \Theta_i$, and any outcome $\lambda \in \Lambda(\Theta \setminus \{\theta_i\})$, there exists $\bar{\lambda} \in \Lambda(\Theta)$ such that $\bar{\lambda}|_{\Theta \setminus \{\theta_i\}} = \lambda$.*[28]

The AAT property says that given an arbitrary simple mechanism on the type space $\Theta \setminus \{\theta_i\}$, we can always accommodate an additional type θ_i of agent i, that is, assign a K-dominant strategy to θ_i while keeping the outcome of the mechanism for the remaining types unchanged.

[28] We write $\bar{\lambda}|_{\Theta \setminus \{\theta_i\}}$ for the outcome $\bar{\lambda}$ restricted to $\Theta \setminus \{\theta_i\}$.

For example, for SP mechanisms, the AAT property holds in environments with one-dimensional types and single-crossing. Indeed, when a new type is added, the mechanism does not need to be adjusted at all, and the new type will have a dominant strategy among the strategies already offered by the mechanism. Intuitively, this is because the incentive constraint for the new type will only bind against a single existing type, so we can simply offer the strategy of that type to the new type.

Proposition 2. *Suppose that the AAT property fails. Then, there exists a payoff function of the designer such that the best simple mechanism is strongly dominated.*

The intuition behind Proposition 2 is straightforward. Suppose that the AAT property fails when some type θ_i is "added" to the type space $\Theta_i \setminus \{\theta_i\}$. Then, there exist objective functions for the designer that are maximized at some outcome that is simply-implementable on $\Theta \setminus \{\theta_i\}$ only when type θ_i is not present, and do not depend "too much" on the outcome implemented for type θ_i (for example, because θ_i has low probability under π). Under such an objective, the designer can strictly improve upon the best simple mechanism by "ignoring" type θ_i: She offers the same mechanism that she would have offered if type θ_i were not present. The mechanism is complex on Θ because θ_i is strategically confused. However, when the designer's payoff does not depend too much on the outcome for θ_i, it will strongly dominate the optimal simple mechanism on Θ. The proof of Proposition 2—which formalizes this argument—can be found in Appendix A.5.

Proposition 2 points to a certain weakness of simple mechanisms—they can be overly restrictive by requiring that all types are not confused, while the behavior of some of these types could be insignificant for the designer (because these types have low probability, or they do not contribute to the designer's payoff).[29] Because Proposition 2 considers an abstract environment, it only predicts strong dominance of a simple mechanism for *some* objective function. However, a similar reasoning can be applied for any fixed objective function: When the set of simply-implementable outcomes shrinks as the type space grows, there will in general be a trade-off between implementing a mechanism that is simple on a subset of the type space (with remaining types being confused) and making the mechanism simple for all

[29] Note that the solution concept of Bayesian Nash equilibrium does not exhibit such a problem.

types. Sometimes, the trade-off will be resolved in favor of a complex mechanism. The following example illustrates.

Example 3. Consider a voting environment with two agents and three alternatives, $\mathcal{X} = \{a, b, c\}$. Each agent's type can be represented as a ranking of the three alternatives. More specifically, each agent gets utility 1 if her top choice is implemented, 1/2 if her second choice is implemented, and 0 otherwise. The distribution of types π is i.i.d. uniform. The designer would like to maximize welfare but is Rawlsian and risk-averse: If u_i is the ex-post utility of agent i, then the designer's payoff is $v(\min(u_1, u_2))$ for some strictly concave and increasing function v.

The best simple mechanism (for any $K \in \{\text{SP, OSP, SOSP}\}$) is dictatorship with full range.[30] The outcome of that mechanism (with the row player being a dictator) is illustrated in Table 1.

Table 1: The best simple mechanism

	abc	acb	bac	bca	cab	cba
abc	a	a	a	a	a	a
acb	a	a	a	a	a	a
bac	b	b	b	b	b	b
bca	b	b	b	b	b	b
cab	c	c	c	c	c	c
cba	c	c	c	c	c	c

Table 2: A complex mechanism

	abc	acb	bac	bca	cab	cba
abc	a	a	a	a	a	a
acb	a	a	a	a	a	a
bac	b	b	b	b	b	b
bca	b	b	b	b	b	b
cab	a	a	a	c	c	c
$cba(1)$	a	a	a	c	c	c
$cba(2)$	b	b	b	b	b	b

Consider instead a delegation mechanism in which the designer delegates the choice of the menu to the row player, who chooses a menu for the column player to choose from (from a given set of menus that the designer has specified): The row player can *(i)* choose alternative a—menu $\{a\}$, *(ii)* choose alternative b—menu $\{b\}$, or *(iii)* eliminate alternative b and leave the choice between a and c to the column player—menu $\{a, c\}$.

Types abc, acb, bac, bca of the row player retain their K-dominant strategies. Type cab has a K-dominant strategy to choose the menu $\{a, c\}$. However, type

[30] It suffices to show this for the concept of SP since it is the most permissive one (and because dictatorship is a simple mechanism under all notions). Due to our Definition 1 of SP, it is without loss of optimality to look at deterministic mechanisms, and thus only ordinal preferences of players matter. Therefore, dictatorship is the only feasible SP mechanism by the GibbardâĂŞSatterthwaite theorem. Finally, we can directly verify that giving the dictator full range is optimal for our objective function.

cba does not have a K-dominant strategy: Both menu $\{b\}$ and menu $\{a, c\}$ are not K-dominated—the mechanism is complex. This is illustrated in Table 2.

However, whichever strategy type *cba* chooses, the expected payoff to the designer is strictly higher than in the best simple mechanism. Indeed, consider first the case in which type *cba* chooses menu $\{a, c\}$. Then, the difference in expected payoffs to the designer between the complex mechanism and the best simple mechanism (which can be calculated by comparing the cells of the two tables with different outcomes) is

$$\frac{1}{36} \times \left[v\left(\frac{1}{2}\right) - v(0)\right] > 0.$$

Now, consider the case in which type *cba* chooses menu b. The difference in the expected payoffs of the designer is

$$\frac{1}{36} \times [-2v(1) + 4v(1/2) - 2v(0)] > 0,$$

by the strict concavity of v. Thus, the complex mechanism is guaranteed to yield a strictly higher expected payoff to the designer regardless of how agents behave.

Note that the AAT property is violated. To wit, the simple mechanism defined by the first 5 rows of Table 2 by excluding type *cba* cannot be extended to a simple mechanism with type *cba* added back. Type *cba* is payoff-relevant for the designer; nevertheless, the designer prefers to "ignore" that type by implementing a mechanism that is simple for the remaining types and that makes *cba* strategically confused. □

More generally, the AAT property fails in many classical social choice environments. It is well known that the only SP mechanisms whose range contains at least three alternatives are dictatorships; however, there are nontrivial strategy-proof social choice functions on restricted domains. Indeed, much of the research on strategy-proof social choice can be seen as establishing possibility results for (various) restricted domains.[31] Thus, the AAT property fails. Proposition 2 then implies that in the social choice environment, it might be beneficial to "ignore" some types and employ a social choice rule that is SP-implementable on a smaller

[31] We refer interested readers to Barberà (2010) for a survey on strategy-proof social choice.

domain, rather than using a SP mechanism on a larger domain. In particular, this will hold if the "problematic" types occur with sufficiently low probability. For a concrete example, it is known that a social choice function on profiles of single-peaked preferences over a totally ordered set is strategy-proof if and only if it is a generalized median voter scheme. If the designer finds the outcome of some generalized median voter scheme more desirable than that of a dictatorship, and the probability that the true type profile is contained in the set of single-peaked preferences is high enough, then the generalized median voter scheme strongly dominates the best simple mechanism on the full domain. Similar examples can be found for the notions of OSP and SOSP.[32]

4.2 Richness in preferences

In this subsection, we show that if there is sufficient richness in the agents' preferences, that is, all preference orderings over alternatives are plausible, then there exists some payoff function of the designer such that the best simple mechanism is strongly dominated. This is true even in the case of a single agent.

Proposition 3. *Suppose that* $N \geq 1$, $|\mathcal{X}| \geq 6$, *and for at least one* $i \in \mathcal{N}$, *all possible rankings of alternatives appear in* Θ_i. *Then, for some objective function of the designer, the optimal simple mechanism is strongly dominated.*

We emphasize that Proposition 3 only predicts that the optimal simple mechanism is strongly dominated for *some* (carefully constructed) objective function of the designer. As a result, the proof of the proposition is essentially an example. The proposition does not imply that there cannot be a foundation for simple mechanisms even on rich preference domains if the objective function belongs to a smaller class. However, it does point out that rich preference domains allow the designer to use complex mechanism to implement outcomes that are not implementable using simple mechanisms.

Proof of Proposition 3. We explicitly construct an example such that the best simple mechanism is strongly dominated for all $K \in \{\text{SP, OSP, SOSP}\}$. Since we can select an objective function for the designer, we can without loss of generality

[32] See Bade and Gonczarowski (2017) who characterize the class of OSP-implementable and unanimous social choice functions for single-peaked preferences, among other applications.

assume that there is only one player with three equally-likely types, $\Theta = \{u, m, d\}$, $\mathcal{X} = \{U, U', M, M', D, D'\}$, with the following preferences:

1. type u: $M > U > D' > D > M' > U'$;
2. type m: $D > M > U' > U > D' > M'$;
3. type d: $U > D > M' > M > U' > D'$.

The designer gets a utility of 1 if the type is $j \in \{u, m, d\}$ and she implements outcome J or J'; she gets -1 otherwise.

Claim 3. *The best simple mechanism is to implement any fixed (possibly random) outcome; the expected payoff for the designer is* $-1/3$.

It suffices to argue that this is true with SP as the solution concept. Recall that the simple mechanism must provide a dominant strategy also with respect to Nature's moves (randomization), and thus it is without loss of optimality for the designer to consider deterministic mechanisms when optimizing in the class of simple mechanisms. A deterministic SP mechanism for a single agent can be represented as a direct assignment of alternatives to types such that no type strictly prefers another type's assignment to her own. Suppose that there exists such an assignment that gives the designer an expected payoff strictly above $-1/3$. Then, at least two types $j \in \{u, m, d\}$ must be assigned either J or J'. If any type j is assigned J', then no other alternative can be offered by the mechanism since j ranks J' last. Thus, the mechanism must offer J to two distinct types j. However, that's a contradiction because at least one of these types would prefer the allocation of the other one, no matter which two types $j \in \{u, m, d\}$ we choose.

To finish the proof, we construct a superior complex mechanism Γ.

Claim 4. *There exists a complex mechanism* Γ *that guarantees the designer an expected payoff of* 0.

In the mechanism Γ, the designer uses a randomization device that has equally likely outcomes H and T, and offers three possible strategies to the agent, as represented by the following normal form:

	H	T
S_U	U	D'
S_M	M	U'
S_D	D	M'

For any solution concept $K \in \{\text{SP}, \text{OSP}, \text{SOSP}\}$, type $j \in \{u, m, d\}$ is confused between the two strategies offering J and J', respectively, with the former one (denoted S_J) leading to the 2nd or 3rd alternative, and the latter to the 1st or 6th alternative. However, the remaining strategy is K-dominated for j by the strategy S_J as it leads to the 4th or 5th alternative. Regardless of how j resolves her strategic confusion, either J or J' is implemented with probability 1/2, and thus the designer obtains 0 in expectation. □

While the example in Proposition 3 is carefully constructed and has no practical interpretation, it illustrates a rather surprising possibility: A simple mechanism can be strongly dominated even when there is a single agent, as long as the agent is not assumed to evaluate lotteries using the notion of expected utility, as captured by Definitions 1-3.[33] Effectively, the designer is using randomization to confuse the agent in order to achieve a better outcome regardless of how the agent resolves the confusion.

For the solution concept of SP, Proposition 3 remains true even if there are only 4 alternatives—we present an example in the Supplemental Material OA.2.1. We conjecture that more than 4 alternatives are needed to build an example of strong dominance of a simple one-player mechanism for the notions of OSP and SOSP.

One may be concerned that Proposition 3 relies crucially on our assumption about how agents reason about randomization devices in the mechanism. However, the conclusion of the proposition remains robust to alternative assumptions as long as there are at least two players with rich enough (cardinal) preferences over pure outcomes. Formally, we can exploit the observation from Remark 1: When $\mathcal{X}$ is assumed to contain all possible lotteries over pure outcomes, it is as if agents evaluated all lotteries in the mechanism by their expected utility. In that case, we

[33] Formally, our example relies on the fact that $\mathcal{X}$ does not contain lotteries over deterministic outcomes; see Remark 1.

can still construct an example—similar to the one above—in which the optimal simple mechanism is strongly dominated. The construction is formally presented in the Supplemental Material OA.2.2. Intuitively, we introduce a second player who has two equally-likely types, H or T, and add an additional component to the designer's utility which makes it suboptimal to implement outcomes J' when the second player's type is H, and suboptimal to implement deterministic outcomes J when the second player's type is T, with $J \in \{U, M, D\}$. Then, we show that under a particular cardinal representation of player 1's ordinal preferences described above, the optimal simple mechanism yields an expected payoff of approximately $-1/6$. The complex mechanism presented above still guarantees an expected payoff of 0, and thus it strongly dominates the optimal simple mechanism.

4.3 An optimality foundation for simple mechanisms

In this subsection, we show that in a class of environments that we define below, the best simple mechanism is not strongly dominated. The key observation is that for any finite mechanism Γ (simple or complex), $\inf V(\Gamma)$ is weakly less than the maximum expected payoff the designer could obtain in a mechanism that only satisfies a subset of incentive constraints that correspond to the edges of an *arbitrary* tree in the type space. Thus, if the designer's expected payoff from the best simple mechanism is the same as that in the relaxed problem with incentive constraints that correspond to the edges of *some* tree in the type space, then the best simple mechanism is not strongly dominated.

Let $G = (V, E)$ be a directed graph with vertex set V and edge set $E \subseteq V \times V$. A graph is called connected if there exists a path between any two vertices. A graph is called a tree if it is connected and has no cycles. For each agent i, consider a tree $T_i = (\Theta_i, E_i)$ where $E_i \subseteq \Theta_i \times \Theta_i$.[34] Each directed edge $(\theta_i, \theta_i') \in E$, also denoted $\theta_i \to \theta_i'$, corresponds to the incentive constraint that type θ_i does not want to adopt the strategy of type θ_i'. For each agent i, fix a tree T_i. The collection of trees $\{T_i\}_{i \in \mathcal{N}}$ then defines a relaxed optimization problem in which the only incentive constraints are the ones that correspond to the edges on the trees. Formally, the

[34]If the agents have the option of opting out, we let θ_0 denote the dummy type that corresponds to not participating in the mechanism. Let $\hat{\Theta}_i = \Theta_i \cup \{\theta_0\}$, and a tree for agent i is $T_i = (\hat{\Theta}_i, E_i)$ where $E_i \subseteq \hat{\Theta}_i \times \Theta_i$

relaxed problem is to find a mechanism Γ that assigns a strategy $\boldsymbol{S}_i(\theta_i)$ to each type θ_i of each player $i \in \mathcal{N}$ and maximizes the designer's objective functions among all such mechanisms with the property that if $\theta_i \to \theta_i'$, then $\boldsymbol{S}_i(\theta_i') \prec_{\theta_i}^K \boldsymbol{S}_i(\theta_i)$ (note that if T_i were not a tree but the complete graph, this would correspond to finding the optimal simple mechanism).

Proposition 4. *Suppose that there exists a collection of trees $\{T_i\}_{i\in\mathcal{N}}$ such that the designer's expected payoff from the relaxed optimization problem corresponding to $\{T_i\}_{i\in\mathcal{N}}$ is the same as from using the optimal simple mechanism. Then, the optimal simple mechanism is not strongly dominated.*

Proposition 4 is relatively abstract. Moreover, it does not offer any guidance on how to look for the "right" tree in the type space (for most choices of the trees, the relaxed problem will have a maximum that is not achievable by a simple mechanism, even if for some tree the maxima are the same). For illustration, we apply Proposition 4 to show that the optimal SP mechanism is not strongly dominated in the setting of a revenue-maximizing designer. Since for any mechanism, a strategy that is not dominated is also not OSP-dominated, whenever the optimal SP mechanism can be implemented via an OSP mechanism, our analysis below implies that the optimal OSP mechanism is not strongly dominated.[35]

In a revenue-maximization problem, each agent can opt out, and we let θ_0 denote the dummy type that corresponds to not participating in the mechanism.[36] In what follows, we show that the assumption of Proposition 4 is satisfied when the uniform shortest-path tree condition holds and the distribution π is regular, as defined by Chen and Li (2018). Loosely speaking, these conditions ensure that, in the optimal SP mechanism, the set of binding constraints for agent i is independent of the types of agents other than i. To be rigorous and self contained, we formally define these terms in the Supplemental Material OA.3. Here, we note that the uniform shortest-path tree condition is of interest because a number of resource allocation problems satisfy it. First, it is satisfied in environments with one-dimensional types and single crossing. This fits many classical applications of

[35] Of course, the same holds for the solution concept of SOSP.

[36] This way of modeling participation implies the partial incentives to participate in the mechanism, thus making the result stronger: Not being strongly dominated by a mechanism with partial incentives to participate implies not being strongly dominated by a mechanism with full incentives to participate.

mechanism design, including single-unit auctions (e.g., Myerson (1981)), public goods (e.g., Mailath and Postlewaite (1990)), and standard bilateral trade (e.g., Myerson and Satterthwaite (1983)). This covers Example 1 in the Introduction, showing that—under the assumptions we imposed there—the ascending clock auction cannot be *strongly* dominated. The uniform shortest-path tree condition also holds in multi-unit auctions with homogeneous or heterogeneous goods, as long as the agents' private values are one-dimensional. Second, the uniform shortest-path tree condition can also be satisfied in some multi-dimensional environments, such as the auction for capacity constrained bidders (see Malakhov and Vohra (2009)).

By the definition of the uniform shortest-path tree, for each agent, there exists a tree T_i such that the designer's expected payoff from using the best SP mechanism is the same as the relaxed optimization problem corresponding to $\{T_i\}_{i\in\mathcal{N}}$. Thus, Proposition 4 applies and we obtain the following corollary.

Corollary 1. *In environments in which the uniform shortest-path tree condition holds and π is regular, the optimal SP mechanism is not strongly dominated.*

Example 2 in the Introduction shows that when the uniform shortest-path tree condition is violated, the conclusion of Corollary 1 can also fail.

5 Conclusion

In mechanism design, it seems useful to distinguish (simple) mechanisms in which agents face a straightforward choice problem from (complex) mechanisms that require agents to engage in complex thinking if they want to determine their optimal strategy. The literature has made a great deal of progress in terms of formulating different notions of simplicity and characterizing mechanisms that are simple according to these notions. However, the understanding of the design of mechanisms with unsophisticated agents, as we argued in this paper, is far from complete. Indeed, in many cases, the designer might prefer mechanisms that are not simple, even under the adversarial selection that agents choose strategies that are the worst possible for the designer whenever agents cannot pin down their optimal strategy.

We suggest some directions for further research. An important avenue is the optimal design of mechanisms when agents are strategically unsophisticated. Our analysis indicates that the optimal design of mechanisms with unsophisticated agents could be challenging: One should not simply optimize over the class of simple mechanisms, at least not without a careful examination to establish their optimality; searching over all mechanisms present new challenges, as the designer has to take into account the worst-case scenario whenever agents are confused.

Relatedly, while we focused primarily on negative results throughout the paper, we expect that establishing optimality foundations for simple mechanisms—not being weakly or strongly dominated—might be a particularly promising research direction. This is because such a foundation can often be found by first solving an easier relaxed problem, and then showing that the upper bound is achieved by a simple mechanism. In this paper, we extended the relaxation proposed first by Yamashita (2015) for SP mechanisms to show some instances in which such a strategy works.

While we focused on the strategic dimension and worked with SP, OSP, and SOSP mechanisms in the private-value setting, the same exercise could be done for other notions of simplicity along the strategic dimension (such as the notion of "strategically simple mechanisms" defined by Börgers and Li (2019)), other dimensions of simplicity such as computational complexity (for example, the sheer size of the strategy space may make a mechanism difficult to understand), and interdependent-value settings.[37]

Finally, it would be interesting to conduct experimental tests of the best simple mechanism and the complex mechanism that weakly dominates it, such as the ascending-clock auction and the ascending-clock auction with jump bidding (in the private-value setting). Since the notion of weak dominance requires that the complex mechanism generates a strictly higher payoff to the designer in some but not all cases, it is useful to find out whether and how often (and in which settings) these complex mechanisms do generate a strictly higher payoff. These findings could then be used to support or invalidate the superiority of these complex

[37] Jehiel et al. (2006)[Example 5.1] illustrates that—in our language—the designer might prefer a complex mechanism to the best simple mechanism in the interdependent-value setting, where the solution concept is ex-post incentive compatibility. Also see Yamashita (2015) for the analysis of interdependent-value auctions.

mechanisms.

References

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

Avery, C. (1998): "Strategic Jump Bidding in English Auctions," *Review of Economic Studies*, 65, 185–210.

Bade, S. and Y. A. Gonczarowski (2017): "Gibbard-Satterthwaite Success Stories and Obvious Strategyproofness," in *Proceedings of the 2017 ACM Conference on Economics and Computation*, 565–565.

Barberà, S. (2010): "Strategy-Proof Social Choice," in *Handbook of Social Choice and Welfare*, ed. by K. Arrow, A. Sen, and K. Suzumura, North-Holland, vol. 2, 731–831.

Bergemann, D. and S. Morris (2005): "Robust Mechanism Design," *Econometrica*, 73, 1771–1813.

Börgers, T. (1991): "Undominated Strategies and Coordination in Normalform Games," *Social Choice and Welfare*, 8, 65–78.

——— (2017): "(No) Foundations of dominant-strategy mechanisms: a comment on Chung and Ely (2007)," *Review of Economic Design*, 21, 73–82.

Börgers, T. and J. Li (2019): "Strategically Simple Mechanisms," *Econometrica*, 87, 2003–2035.

Börgers, T. and D. Smith (2012): "Robustly Ranking Mechanisms," *American Economic Review*, 102, 325–329.

Carroll, G. (2014): "A Complexity Result for Undominated-Strategy Implementation," *Working Paper.*

——— (2019): "Robustness in Mechanism Design and Contracting," *Annual Review of Economics*, 11, 139–166.

Chen, Y.-C. and J. Li (2018): "Revisiting the Foundations of Dominant-Strategy Mechanisms," *Journal of Economic Theory*, 178, 294–317.

Chernev, A., U. BÃűckenholt, and J. Goodman (2015): "Choice overload: A conceptual review and meta-analysis," *Journal of Consumer Psychology*, 25, 333–358.

Chung, K.-S. and J. C. Ely (2007): "Foundations of Dominant-Strategy Mechanisms," *Review of Economic Studies*, 74, 447–476.

Cramton, P., R. Gibbons, and P. Klemperer (1987): "Dissolving a Partnership Efficiently," *Econometrica*, 55, 615–632.

Hagerty, K. M. and W. P. Rogerson (1987): "Robust Trading Mechanisms," *Journal of Economic Theory*, 42, 94–107.

Jackson, M. O. (1992): "Implementation in Undominated Strategies: A Look at Bounded Mechanisms," *Review of Economic Studies*, 59, 757–775.

Jakobsen, A. M. (2020): "A Model of Complex Contracts," *American Economic Review*, 110, 1243–1273.

Jehiel, P., M. Meyer-ter Vehn, B. Moldovanu, and W. R. Zame (2006): "The Limits of Ex Post Implementation," *Econometrica*, 74, 585–610.

Li, S. (2017): "Obviously Strategy-Proof Mechanisms," *American Economic Review*, 107, 3257–87.

Loertscher, S. and L. M. Marx (2020): "A Dominant-Strategy Asset Market Mechanism," *Games and Economic Behavior*, 120, 1–15.

Lu, H. and J. Robert (2001): "Optimal Trading Mechanisms with Ex Ante Unidentified Traders," *Journal of Economic Theory*, 97, 50–80.

Mackenzie, A. (2019): "A Revelation Principle for Obviously Strategy-Proof Implementation," *Working Paper*.

Mailath, G. and A. Postlewaite (1990): "Asymmetric Information Bargaining Problems with Many Agents," *Review of Economic Studies*, 57, 351–367.

Malakhov, A. and R. V. Vohra (2009): "An Optimal Auction for Capacity Constrained Bidders: A Network Perspective," *Economic Theory*, 39, 113–128.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Myerson, R. and M. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–281.

Nisan, N. and A. Ronen (2001): "Algorithmic Mechanism Design," *Games and Economic behavior*, 35, 166–196.

Pycia, M. and P. Troyan (2019): "A Theory of Simplicity in Games and Mechanism Design," *Working Paper*.

Ray, D. ⓡ A. Robson (2018): "Certified Random: A New Order for Coauthorship," *American Economic Review*, 108, 489–520.

Rochet, J.-C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191–200.

Yamashita, T. (2015): "Implementation in Weakly Undominated Strategies: Optimality of Second-Price Auction and Posted-Price Mechanism," *Review of Economic Studies*, 82, 1223–1246.

A Appendix

A.1 The optimal OSP mechanism for Example 2

Here, we show how to derive the optimal OSP mechanism in Example 2. It follows from Theorem 3.1 of Bade and Gonczarowski (2017) that it is without loss of optimality for OSP implementation to look at gradual revelation mechanisms (see Ashlagi and Gonczarowski (2018), Pycia and Troyan (2019) and Mackenzie (2019) for related revelation principles for OSP mechanisms). In our simple example with two players and two types, this means that we can assume that in the best OSP mechanism, at the first decision node, one of the players ("leader") makes a binary decision (with the two types choosing different actions—potentially leading to the same outcome—as part of their obviously dominant strategies), and then in each of the two possible histories, having observed the choice of the leader, the other player ("follower") makes a binary decision.

Therefore, an upper bound on the profit in the optimal OSP mechanism can be derived in the following way. When the follower chooses her optimal action, she already knows the action chosen by the leader. OSP requires that for any choice of the leader, each type of the follower must weakly prefer her equilibrium strategy (action) to choosing the alternative action. Thus, each type of the follower must have a standard *dominant strategy* in the normal-form representation of the game. In contrast, when the leader chooses her action at the initial decision node, it must be that the worst possible payoff from choosing her equilibrium action (over the two possible actions that can be selected by the follower in the subgame) is weakly higher than the best possible payoff from choosing the alternative action. In the normal-form representation of the game, this can be captured by requiring that the payoff from the equilibrium strategy of each type of the leader under any choice of the strategy S_f for the follower is weakly higher than the payoff from following the alternative strategy under any choice of the strategy S'_f (where, importantly, S_f could be different from S'_f).

Summarizing, an upper bound can be derived by using a normal-form game with the usual strategy-proof constraints, except that for the leader the constraints are strengthened in the way described above. Crucially, all these constraints are linear in the allocation and transfers, and so is the objective function of the designer (intuitively, we avoid taking the min and max in the definition of obvious dominance by iterating over all possible *pairs* of strategies (S_f, S'_f) for the follower when comparing two strategies of the leader). Thus, we obtain a linear program that can be numerically solved using a standard linear programming solver. For $\kappa = 2/5$ (resp. 1/3 and 1/2), we obtain an upper bound of 1/5 (resp. 2/9 and 1/6), respectively. This upper bound is achieved by the OSP mechanism described in the example, proving its optimality.

The optimal SP mechanism can be derived by solving the standard linear-programming formulation numerically. For $\kappa = 2/5$, we obtain an optimal SP profit of 4/15. It can be achieved with the following direct mechanism, where the first entry in each cell indicates the direction of trade, the second entry is the seller price, and the third entry is the buyer price.

	$\theta_B = 1/3$	$\theta_B = 1$
$\theta_A = 0$	$A \to B,\, 0,\, \frac{1}{3}$	$A \to B,\, \frac{2}{3},\, \frac{1}{3}$
$\theta_A = 2/3$	$B \to A,\, \frac{1}{3},\, \frac{2}{3}$	$A \to B,\, \frac{2}{3},\, 1$

The platform's profit is 1/3 conditional on all type profiles other than $(\theta_A = 0,\, \theta_B = 1)$ for which the platform actually pays 1/3 to dealer B. Intuitively, this rebate to dealer B is an information rent for type $\theta_B = 1$ that she must receive under strategy-proofness in order not to deviate to reporting $\theta'_B = 1/3$ when dealer A's type is $\theta_A = 0$.

A.2 Proof of Proposition 1

We prove the remaining two cases of Proposition 1.

OSP: The game Γ' is defined just as in the SP case. It suffices to prove that the three claims made in the second paragraph for the case of SP dominance continue to hold for OSP dominance. The first claim extends trivially. As for the second claim, $\theta_i \notin \Theta_i^{\mathcal{Y}}$ is equivalent to saying that choosing the game $\Gamma_{-i}^{\mathcal{Y}}$ is obviously dominated by the strategy "choose Γ" and then play according to $\boldsymbol{S}_i(\theta_i)$, since the unique point of departure for these two strategies is the first node for player i. Finally, the last claim follows from the fact that if a strategy is not SP-dominated for a type, then it is also not OSP-dominated.

SOSP: For this case, we modify the construction of the game Γ'. Pycia and Troyan (2019) prove that it is without loss of generality for the designer to choose a game form in which each player only moves once in every history, in the sense that for any SOSP mechanism, there exists a payoff equivalent SOSP mechanism (for all players and the designer) with that property. Thus, we can replace both Γ and $\Gamma_{-i}^{\mathcal{Y}}$ with their payoff-equivalent versions in which each player moves at most once in every history. Moreover, Pycia and Troyan (2019) show that, without loss of generality, Nature's moves take place at the beginning of each history, if at all.

Let i be the player that moves first in the game Γ (if Nature moves first, i can depend on Nature's choice, and we apply the argument case-by-case, for every realization of Nature's choice). For that player, add a new potential action from the initial node which is to choose the game $\Gamma^{\mathcal{Y}}_{-i}$. If that action is taken by player i, play proceeds to the game tree $\Gamma^{\mathcal{Y}}_{-i}$. This defines a new game Γ'.

Note that in Γ' all players still move at most once in every history; player i has no moves in the game $\Gamma^{\mathcal{Y}}_{-i}$, while players $-i$ either play the game Γ (in which they have at most one move) or the game $\Gamma^{\mathcal{Y}}_{-i}$ (in which they have at most one move). Therefore, all the conclusions from the OSP case exceed to this case (since OSP and SOSP coincide if a player only moves once).

A.3 Proof of Claim 1

Fix Γ that is a revenue-maximizing simple mechanism. Because of symmetry, we can assume without loss of optimality for the designer that Γ is symmetric as well, so we only have to check the assumptions of Proposition 1 for some player i. Define $\mathcal{Y}$ to contain two options: *(1)* Player i wins the object and pays $\bar{u}$ while some player j receives $\epsilon > 0$, and *(2)* Player i wins the object and pays $\bar{u} - \delta$ while player j receives 0. In the game $\Gamma^{\mathcal{Y}}_{-i}$, player j selects one option from $\mathcal{Y}$. $\Gamma^{\mathcal{Y}}_{-i}$ is thus simple, with player j always selecting option *(1)*. Take $\delta > 0$ small enough so that $\Theta_i \cap (\bar{u} - \delta, \bar{u}) = \emptyset$ (using finiteness of the type space). Then, we have $\Theta^{\mathcal{Y}}_i = \{\bar{u}\}$, because for all other types the payoff from $\mathcal{Y}$ is strictly negative (while the original strategy in Γ guarantees a non-negative payoff).[38] Moreover, type $\bar{u}$ always gets a non-negative payoff in $\Gamma^{\mathcal{Y}}_{-i}$ so she has a full incentive to participate. Condition 1 of Proposition 1 holds, while condition 2 is satisfied as long as the mechanism Γ extracts less than $\bar{u}$ from type $\bar{u}$ of player i. Thus, by Proposition 1, Γ is weakly dominated as long as it is not a mechanism that offers a price $\bar{u}$ to player i. By symmetry, the only case not covered by the argument is when Γ is payoff-equivalent to a mechanism that offers a price $\bar{u}$ to all players in random order. However, such a mechanism is suboptimal by assumption.

[38] $\bar{u} \in \Theta^{\mathcal{Y}}_i$ as long as there is positive probability that she does not receive the object in Γ – that is guaranteed by the assumption that the players and Γ are symmetric.

A.4 Proof of Claim 2

By assumption, the unique optimal simple mechanism is to offer some price $p^\star$; let $\bar{v}$ denote the corresponding optimal expected revenue. Note that, by optimality, there must exist a type $\theta = p^\star$ in Θ.

Towards a contradiction, suppose that there exists a weakly dominating complex mechanism Γ. In particular, $\inf V(\Gamma) \geq \bar{v}$. We claim that there exists a selection from the set of K-undominated strategies $U_i^K(\cdot)$ in Γ such that *(i)* if every type plays the assigned strategy, the expected revenue for the designer is equal to at least $\inf V(\Gamma)$, and *(ii)* local downward incentive constraints hold, that is, the strategy assigned to type θ_i K-dominates (for θ_i) the strategy assigned to the highest type lower than θ_i. Instead of proving this property directly, we refer the reader to Lemma 1 from Appendix A.6 that establishes a more general result. Moreover, it is well known that for revenue maximization with a single player under SP and OSP, only local downward incentive constraints bind in the optimal mechanism; hence, the expected revenue under the selection of undominated strategies described above is upper-bounded by $\bar{v}$. Since $\bar{v} \leq \inf V(\Gamma)$, we conclude that $\inf V(\Gamma) = \bar{v}$ and there exists a selection from the set of undominated strategies which satisfies local downward incentive constraints and yields an expected revenue of $\bar{v}$.

By the assumption that the optimal simple mechanism is unique, the only way to generate the expected revenue of $\bar{v}$ while satisfying local downward incentive constraints is for all types weakly above $p^\star$ to buy for sure at the expected price of $p^\star$. By the full incentive to participate, type $p^\star$ must have a strategy under which she buys for sure at a price of $p^\star$ in every history. In particular, Γ must offer such a strategy. Moreover, again by the full incentive to participate, this strategy must be K-dominant for type $p^\star$.[39]

By the assumption that Γ weakly dominates the simple mechanism, there must exist a strategy S_1 that is K-undominated for some type θ, and a strategy for Nature S_0 such that the agent pays $q > p^\star$ in the outcome $g(z(h_\emptyset, S_0, S_1))$. Because the strategy "buy for sure at a price of $p^\star$" is available, for S_1 to be K-undominated, it must sometimes (for some strategy of Nature) generate an outcome "buy with probability x at a price of r" that is strictly preferred by θ to the outcome "buy for sure at a price of $p^\star$." At the same time, type $p^\star$ cannot derive positive utility from the outcome "buy with probability x at a price of r" as otherwise the strategy S_1 would be undominated for type $p^\star$, violating the full incentive to participate (since S_1 sometimes leads to the agent paying $q > p^\star$).

[39] Up to this point, the proof also works for SOSP. However, the next step does not.

Moreover, note that $\theta > p^\star$, by the full incentive to participate. We are ready to obtain a contradiction: By the above reasoning, we have $\theta x - r > \theta - p^\star$ and $p^\star x - r \leq 0$. This implies

$$\theta - p^\star < \theta x - r \leq \theta x - p^\star x = x(\theta - p^\star) \leq \theta - p^\star,$$

a contradiction.

A.5 Proof of Proposition 2

Fix a solution concept K. If the AAT property fails, then there exists a type θ_i of agent i and an outcome $\lambda \in \Lambda(\Theta \setminus \{\theta_i\})$ such that for all $\bar{\lambda} \in \Lambda(\Theta)$, $\bar{\lambda}|_{\Theta\setminus\{\theta_i\}} \neq \lambda$. We can assume without loss of generality that there exists such λ that is deterministic for any θ, since if the AAT property holds for all deterministic outcomes, then it also must hold for all stochastic outcomes. Define an objective function v of the designer by

$$v(x, \theta) = \mathbf{1}_{\{x=\lambda(\theta),\, \theta\in\Theta\setminus\{\theta_i\}\}}.$$

Clearly, the best simple mechanism on the type space $\Theta \setminus \{\theta_i\}$ yields a payoff of 1 to the designer. With the objective function v on Θ, the best simple mechanism on the type space Θ must yield an optimal payoff strictly lower than $1 - p$, where p is the unconditional probability of type θ_i under π, since it is not possible, by assumption, to implement the outcome λ on $\Theta \setminus \{\theta_i\}$. We denote the optimal payoff by $b < 1 - p$.

We will show that there exists a complex mechanism that generates a strictly higher expected payoff on Θ than b, regardless of how agents behave when they are confused. Let Γ be a mechanism that implements the outcome λ on $\Theta \setminus \{\theta_i\}$. Consider the same mechanism Γ on Θ. By assumption, Γ is not simple on Θ because type θ_i is strategically confused. But because it is simple on $\Theta \setminus \{\theta_i\}$, all types in $\Theta \setminus \{\theta_i\}$ still play the same (K-dominant) strategy. Therefore, no matter how type θ_i behaves, the designer gets an expected payoff of at least $1 - p$ times 1, which is strictly higher than b, thereby strongly dominating the optimal simple mechanism.

A.6 Proof of Proposition 4

We first state a lemma that builds on the insight in Yamashita (2015, Theorem 1), and can be viewed as its generalization to a larger class of environments and the solution concepts of OSP and SOSP.

Lemma 1. *For any mechanism* Γ*,* $\inf V(\Gamma)$ *is upper bounded by the value of the relaxed problem corresponding to any fixed collection of trees* $\{T_i\}_{i\in\mathcal{N}}$ *(as defined in Section 4.3).*

Proof. Fix an arbitrary finite mechanism Γ, an agent i, and a tree T_i. Let $T_i^+(\theta_i) = \{\theta_i' : \theta_i' \to \theta_i\}$ be the set of types who point towards type θ_i in the tree T_i. Consider the following procedure. Starting at the root of the tree T_i—which is some type θ_i^0 with no edges coming out of it[40]—select any K-undominated strategy for θ_i^0, $S_i^0 \in U_i^K(\theta_i^0)$. Next, for any type $\theta_i' \in T_i^+(\theta_i^0)$, we can find an undominated strategy $S_i' \in U_i^K(\theta_i')$ that either K-dominates S_i^0 or is equal to S_i^0 (this step uses finiteness of the mechanism; if S_i^0 is not in $U_i^K(\theta_i')$, then there must exist a strategy in $U_i^K(\theta_i')$ that K-dominates it). We proceed inductively. Once some type θ_i is assigned a strategy, unless it's a leaf in T_i, we assign undominated strategies to all types $T_i^+(\theta_i)$ that either equal or K-dominate the strategy assigned to θ_i. Because T_i is a finite tree, this procedure must stop at some point, with every type being assigned a strategy. The same procedure is carried out for all other agents.

Because each type is assigned a strategy from $U_i^K(\cdot)$ in the procedure, when all types execute their assigned strategies, the expected payoff $\bar{v}$ to the designer must weakly exceed $\inf V(\Gamma)$ (which is the outcome of the designer-adversarial selection from $U_i^K(\cdot)$). Moreover, the procedure guarantees that the mechanism—along with the assignment of strategies—is feasible for the relaxed problem corresponding to the collection of trees $\{T_i\}_{i\in\mathcal{N}}$. Therefore, $\bar{v}$, and hence also $\inf V(\Gamma)$, is weakly below the value of the relaxed problem. □

The proposition follows immediately: By assumption, there exists a collection of trees $\{T_i\}_{i\in\mathcal{N}}$ such that the value of the relaxed problem corresponding to that collection is the same as the designer's expected payoff from the optimal simple mechanism. Hence, by applying Lemma 1 for the collection $\{T_i\}_{i\in\mathcal{N}}$, we conclude that there cannot exists a mechanism Γ with $\inf V(\Gamma)$ strictly above the expected payoff of the optimal simple mechanism, and hence the optimal simple mechanism is not strongly dominated.

[40]Such a type exists and is unique, by the definition of a tree.

Supplemental Material

OA.1 Supplemental Material for Subsection 3.2

In this appendix, we show that the uniqueness assumption in Claim 2 is needed. We construct an example in which the optimal SP and OSP mechanism is not unique, and are weakly dominated.

The agent has value 1 or 2, with equal probabilities. The optimal SP and OSP mechanisms generate an expected revenue of 1 (which can be obtained by charging a price of 1, or a price of 2). The weakly dominating mechanism features Nature that plays H with some small probability $\epsilon > 0$, and T otherwise. The mechanism offers three strategies to the agent, where the first number in every cell is the probability of trade, and the second number is the payment to the designer.

	H	T
S_1	$(1, 3/2)$	$(1, 3/2)$
S_2	$(1, 1)$	$(1, 2)$
S_3	$(1/2, 1/2)$	$(1/2, 1/2)$

For type 2, S_1 and S_2 are undominated (playing S_3 can be ruled out under the assumption that the designer can choose between strategies that are payoff equivalent, in this case S_1 and S_3). Type 1 has a dominant strategy S_3. Thus, Γ provides full incentives to participate. In the worst case, type 2 plays S_1, and the designer obtains an expected revenue of 1. In the best case, type 2 plays S_2, and the designer obtains an expected revenue of $1/2 + 1/2 \cdot (2 - \epsilon)$ which can be arbitrarily close to $3/2$.

OA.2 Supplemental Material for Subsection 4.2

OA.2.1 An example of strong dominance for SP when $|\mathcal{X}| = 4$

We present here an example to demonstrate that the best SP mechanism can be strongly dominated (for some objective function of the designer) as long as there are at least 4 possible deterministic outcomes. There is a single agent with two types, u and d. Let π_u and $\pi_d = 1 - \pi_u$ denote the respective probability of these two types. There are 4 alternatives; $\mathcal{X} = \{U, U', D, D'\}$. The preferences of the types are given as follows:

1. type u: $U' > D > U > D'$;
2. type d: $D' > U' > D > U$.

The designer receives a payoff of 1 if the type is $j \in \{u, d\}$ and the outcome is J; she receives 1/2 when the type is j and the outcome is J'; in all other cases, she receives 0.

Consider the following mechanism Γ in which the designer uses a randomization device that has two equally likely outcomes "Heads" and "Tails":

	"Heads"	"Tails"
s_u^H	U	U'
s_u^T	U'	U
s_d^H	D	D'
s_d^T	D'	D

This mechanism has a simple interpretation: The player is asked to choose between a $50-50$ lottery over U and U', and a $50-50$ lottery over D and D'. However, the player can additionally choose the side of the coin leading to her preferred outcome.

Claim OA.1. *The mechanism Γ is not SP, and for all values of $\pi_d \in (0,1)$ generates an expected payoff of $3/4$ for the designer.*

Proof. For the solution concept SP, type u is confused between strategies s_u^H and s_u^T, and type d is confused between strategies s_d^H and s_d^T. It is straightforward to verify that this mechanism generates an expected payoff of 3/4 to the designer, no matter which of the two possible strategies these types choose. □

Recall from Definition 1 that our notion of SP requires dominance also with respect to randomization (Nature). That is why each type—when faced with the choice of the side of the coin to "bet on"—is strategically confused. Interestingly, it is not possible to implement the same outcome without creating strategic confusion (while s_u^T dominates s_d^H for type u, s_d^H does not dominate s_u^T for type d; while s_d^T dominates s_u^T for type d, s_u^T does not dominate s_d^T for type u). This implies that the optimal SP mechanism may sometimes be strongly dominated.

Claim OA.2. *If $\pi_d = 2/3$, then there is no SP mechanism that generates the expected payoff of $3/4$ for the designer (the best SP mechanism is to always implement D which yields an expected payoff of $2/3$).*

Proof. We show that the designer could not achieve the expected payoff of 3/4 using a SP mechanism. Note that randomization does not help our designer in a SP mechanism. Indeed, for each SP mechanism in which the designer uses a randomization device, the designer could do weakly better simply by always selecting the same outcome of the randomization device; namely the one associated with the highest expected payoff. Thus, without loss of generality, we restrict attention to deterministic SP mechanisms. Since there is a single agent, each deterministic SP mechanism is equivalent to a menu of deterministic outcomes for the agent to choose from.

To get an expected payoff of at least 3/4, the designer must offer D in the mechanism, and D must be chosen by type d (indeed, in the opposite case, the designer's expected payoff is upper bounded by $1/3 + 2/3 \cdot 1/2 = 2/3$). However, in this case, neither D' nor U' can be offered, as then d would not choose D. Hence, either only D is offered, in which case the designer gets an expected payoff of 2/3, or D and U are offered, in which case type u also selects D, and the expected payoff to the designer is again 2/3. □

OA.2.2 An example of strong dominance when $N \geq 2$ and $\mathcal{X}$ contains lotteries

Assume that $\mathcal{Y}$ is the set of deterministic outcomes, $|\mathcal{Y}| \geq 6$, and $\mathcal{X}$ contains all possible lotteries over $\mathcal{Y}$ (agents' cardinal utilities on $\mathcal{Y}$ are extended to $\mathcal{X}$ using expected utility).

We extend the example from the proof of Proposition 3: We set $\mathcal{Y} = \{U, U', M, M', D, D'\}$. There are two players; Player 1 has the same ordinal preferences over $\mathcal{Y}$ as before; Player 2 has two equally likely types, called H and T, but is indifferent between all alternatives in both cases (this particular choice suffices but is not necessary for our construction to work). To simplify the arguments, we will rely on payoffs from the extended real line. Consider the same objective function for the designer as in the proof of Proposition 3 except that the designer's payoff is $-\infty$ when *(i)* the second player's type is H and one of $\{U', M', D'\}$ is implemented, or *(ii)* when the second player's type is T and one of $\{U, M, D\}$ is implemented. Finally, for player 1 we choose any cardinal representation of her ordinal preferences with the best option yielding $+\infty$ and the worst option yielding $-\infty$.

Claim 5. *The optimal simple mechanism gives the designer an expected payoff of at most* $-1/6$.

As before, it suffices to show the claim for the optimal SP mechanism. By the revelation principle, the optimal SP mechanism is a direct mechanism in which it is incentive compatible for player 1 to report her type truthfully (player 2 always reports truthfully since she is indifferent between all possible outcomes). Let p_j^y be the probability of implementing outcome y for type j in the optimal mechanism when $\theta_2 = H$, and q_j^y be the probability of implementing outcome y for type j in the optimal mechanism when $\theta_2 = T$. By optimality of the mechanism, we must have $p_j^y = 0$ for $y \in \{U', M', D'\}$ and $q_j^y = 0$ for $k \in \{U, M, D\}$. By incentive-compatibility, the following inequalities must hold:

$$\begin{aligned} p_u^M &\geq p_m^M, \\ p_m^D &\geq p_d^D, \\ p_d^U &\geq p_u^U. \end{aligned}$$

This implies that

$$p_u^U + p_d^D + p_m^M \leq p_d^U + p_m^D + p_u^M \leq 3 - (p_u^U + p_d^D + p_m^M) \implies p_u^U + p_d^D + p_m^M \leq \frac{3}{2}.$$

Therefore, conditional on $\theta_2 = H$, the expected payoff for the designer cannot exceed 0. Again by incentive-compatibility, we have

$$\begin{aligned} q_u^{U'} &\leq q_m^{U'}, q_d^{U'} \\ q_m^{M'} &\leq q_u^{M'}, q_d^{M'} \\ q_d^{D'} &\leq q_u^{D'}, q_m^{D'}. \end{aligned}$$

This implies that

$$\begin{aligned} 2(q_u^{U'} + q_m^{M'} + q_d^{D'}) \leq q_m^{U'} + q_d^{U'} + q_u^{M'} + q_d^{M'} + q_u^{D'} + q_m^{D'} \leq 3 - (q_u^{U'} + q_m^{M'} + q_d^{D'}) \\ \implies q_u^{U'} + q_m^{M'} + q_d^{D'} \leq 1. \end{aligned}$$

Therefore, conditional on $\theta_2 = T$, the expected payoff for the designer cannot exceed $-1/3$. Overall, in expectation over θ_2, the designer's payoff cannot exceed $-1/6$. (While not important for the proof, this upper bound is achieved by a mechanism that implements any of $\{U', M', D'\}$ (regardless of the report of player 1), when $\theta_2 = T$; and a $1/2 - 1/2$ lottery between *(i)* M and U when $\theta_1 = u$, *(ii)* D and M when $\theta_1 = m$, and *(iii)* U and D when $\theta_1 = d$, when $\theta_2 = H$.)

Claim 6. *The complex mechanism* Γ *from the proof of Proposition 3 guarantees* *the designer an expected payoff of* 0*, and thus strongly dominates the optimal simple mechanism.*

The proof is immediate: Since we have not changed the ordinal preferences for player 1, the same set of strategies are K-undominated. Since player 2 is indifferent between all alternatives, the mechanism is simple for her and it is optimal to report her type truthfully.[41] Finally, the designer obtains the same payoff guarantee because her objective function is the same as in the proof of Proposition 3 on the set of outcomes of Γ.

OA.3 Supplemental Material for Subsection 4.3

In this appendix, we give a formal definition of the uniform shortest-path tree condition. We work in a quasi-linear setting as defined in Subsection 3.2, with X being the set of all (possibly random) physical allocations, and $\mathcal{X} = X \times \mathbb{R}^N$. We say that a decision rule $g : \Theta \to X$ is strategy-proof if there exists a transfer scheme $t = (t_1, ..., t_N)$ such that (g, t) is an SP mechanism (treated as a direct mechanism). The following standard lemma is due to Rochet (1987).

Lemma 2. *A necessary and sufficient condition for a decision rule g to be strategy-proof is the following cyclical monotonicity condition: $\forall i \in \mathcal{N}$, $\forall \theta_{-i} \in \Theta_{-i}$, and for every sequence of types $(\theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,m})$ with $\theta_{i,m} = \theta_{i,1}$, we have*

$$\sum_{n=1}^{m-1} \Big[u_i(g(\theta_{i,n}, \theta_{-i}), \theta_{i,n+1}) - u_i(g(\theta_{i,n}, \theta_{-i}), \theta_{i,n}) \Big] \le 0. \tag{OA.2}$$

We first collect some graph-theoretic terminology in Definition 9.

Definition 9. *Fix a strategy-proof decision rule g, agent $i \in \mathcal{N}$, and θ_{-i}.*

(1) The set of nodes is $\Theta_i \cup \{\theta_0\}$;

(2) For any $\theta_i \in \Theta_i$ and $\theta'_i \in \Big\{\Theta_i \setminus \{\theta_i\}\Big\} \cup \{\theta_0\}$, $\theta_i \to \theta'_i$ is a directed edge with length

$$u_i(g(\theta_i, \theta_{-i}), \theta_i) - u_i(g(\theta'_i, \theta_{-i}), \theta_i);$$

(3) A path from the dummy type θ_0 to $\theta_{i,m} \in \Theta_i$ is a sequence of nodes $P = (\theta_0, \theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,m})$ where (i) $\theta_{i,j} \in \Theta_i$, $\forall j = 1, 2, \ldots, m$; and (ii) $j \neq j' \implies \theta_{i,j} \neq \theta_{i,j'}$.

[41] We could specify alternative preferences for player 2 that would make it a strict K-dominant strategy to report her type truthfully without affecting the previous claim. This is true, for example, if type H strictly prefers all outcomes in $\{U, M, D\}$ to all outcomes in $\{U', M', D'\}$, and type T has the opposite preferences.

Definition 10. *Fix a strategy-proof decision rule g, agent $i \in \mathcal{N}$, and θ_{-i}. A shortest-path tree is the union of shortest paths from the root to all nodes such that if θ_i' belongs to the shortest path from the root θ_0 to some $\theta_i \in \Theta_i$, then the truncation of the path from θ_0 to θ_i' defines the shortest path from θ_0 to θ_i'.*

Definition 11. *We say that the uniform shortest-path tree condition holds if, for each agent $i \in \mathcal{N}$, there is the same shortest-path tree for all strategy-proof decision rules g and other agents' reports θ_{-i}.*

When the uniform shortest-path tree condition is satisfied, we can drop the dependence of the shortest-path tree on g and θ_{-i}. Let E_s denote the collection of the directed edges in the uniform shortest-path tree.

Suppose that the uniform shortest-path tree condition holds. In the optimal SP mechanism design problem, it suffices to consider constraints that correspond to the edges on the uniform shortest-path tree, subject to the decision rule g satisfying the cyclical monotonicity constraint (OA.2). As is standard in the mechanism design literature, we first consider the following relaxed problem, in which we ignore the cyclical monotonicity constraint. A regularity condition on π is then imposed to ensure that some optimal decision rule g^* that solves the relaxed maximization problem automatically satisfies the cyclical monotonicity constraint.

$$\max_{g(\cdot)\in X,\, t_i(\cdot)} \quad \sum_{\theta\in\Theta} \pi(\theta) \sum_{i\in\mathcal{N}} t_i(\theta) \tag{SP-relaxed}$$

$$\text{subject to } \forall i \in \mathcal{N},\ \forall(\theta_i, \theta_i') \in E_s,\ \forall \theta_{-i} \in \Theta_{-i},$$

$$u_i(g(\theta_i, \theta_{-i}), \theta_i) - t_i(\theta_i, \theta_{-i}) \geq u_i(g(\theta_i', \theta_{-i}), \theta_i) - t_i(\theta_i', \theta_{-i}).$$

Definition 12. *We say that π is regular if the cyclical monotonicity constraint (OA.2) is automatically satisfied for some g^* that solves the optimization problem* (SP-relaxed).[42]

[42]To the best of our knowledge, there is no formal definition of regularity in the general environments. Our definition of regularity captures how it has been used in the literature; see for example, Myerson (1981) and Chung and Ely (2007). See Chen and Li (2018) for the primitive condition for the regularity condition in a variety of settings.