SIMPLE MECHANISMS
(PRELIMINARY DRAFT)

TILMAN BÖRGERS

Department of Economics, University of Michigan

JIANGTAO LI

Department of Economics, National University of Singapore

Abstract. In this paper we define and investigate a property of mechanisms that we call "simplicity," and that is meant to capture the idea that, in "simple" mechanisms, strategic choices are easy. We define a mechanism to be "simple" if optimal strategy choices can be based on first order beliefs alone, and there is no need for agents to form higher order beliefs because such beliefs are irrelevant to agents' optimal choices. In mechanisms "first order beliefs" are beliefs about other agents' rationality and their utility. "Higher order beliefs" are beliefs about beliefs, beliefs about beliefs about beliefs, etc. In many mechanisms agents who want to make an optimal choice cannot avoid having to form higher order beliefs. But in some mechanisms there is no need for this. These are the mechanisms that we investigate and characterize in this paper. All dominant strategy mechanism are simple. But many more mechanisms are simple. In particular, simple mechanisms may be more flexible than dominant strategy mechanisms in examples such as the bilateral trade problem and the voting problem.

Date: February 5, 2017.

We are grateful to Heng Liu and Alessandro Pavan for very helpful comments.

1. Introduction

In mechanism design it often seems desirable for the designer to offer a mechanism in which agents face a straightforward choice problem, and need not engage in complex strategic thinking to determine their best choice. It seems more likely that agents make the choices that the designer expects them to make if thinking through the strategic aspects of the mechanism is simple than when it is complicated. Also, agents may be more willing to participate in simple mechanisms. Finally, it may be desirable that the outcomes of a mechanism don't depend too much on the cognitive abilities of agents. All these arguments provide potential reasons for constructing mechanisms in which strategic choices are easy to make.

One class of mechanisms for which one might argue that it is easy to choose a strategy are dominant strategy mechanisms. In such mechanisms, agents need not think at all about the motivations of other players, or other players' rationality. This is because agents have at least one strategy that is optimal regardless of what other players do, and they can just choose such a strategy.[1] Thus, one complication that may make strategy choices difficult is not present. Of course, strategy choices may be complicated also for other reasons. For example, players may not be able to recognize that they have a dominant strategy. Li [9] has recently proposed the notion of "obviously dominant" strategy mechanisms, which are mechanisms for which the task of identifying dominant strategies is, in some sense, easy. Our concept of "simplicity" ignores this potential difficulty of strategic decision making, and for the purposes of this paper we shall regard the choice of a dominant strategy as easy.

For many mechanism design problems the class of dominant strategy mechanisms is quite small, and only includes mechanisms that are rather unattractive for a mechanism designer who wants to maximize, say, revenue, or welfare.[2] The contribution of this paper is to formalize, and characterize, a notion of "simplicity" of mechanisms that has the property that all dominant strategy mechanisms are "simple," but also the class of "simple"

[1]Here, we use the phrase "dominant strategy" in the sense in which it is used in mechanism design theory, that is, a strategy that is optimal regardless of what the other agents choose. This is slightly different from a strategy that is "weakly dominant" or a strategy that is "strictly dominant" as these terms are defined in game theory.

[2]See the examples in Chapter 4 of Börgers [4].

mechanisms is strictly larger than the class of dominant strategy mechanisms. We show that "simple mechanisms" include in applications mechanisms that may be more attractive to a mechanism designer than dominant strategy mechanisms.

To illustrate our idea it is best to consider an example. Suppose a mechanism designer wants to determine rules for trade between two agents, a seller and a buyer. It is known from Hagerty and Rogerson [8] that, under some conditions, the only dominant strategy mechanisms are mechanisms in which the designer chooses the (possibly random) price, without taking into account any of the agents' private information, and then agents choose whether to agree or not to agree to trade at this price. Trade comes about only when both agents agree. Obviously, this is a rather unappealing mechanism for a welfare maximizing mechanism designer who wants efficient trade to take place whenever possible.

An alternative mechanism is that the mechanism designer sets a price cap, but allows the seller to reduce the price. The buyer can then agree or not agree to trade at this potentially reduced price. The seller clearly does not have a dominant strategy. Whether or not to reduce the price, and how far to reduce the price, depends on the seller's beliefs about the buyer's choices. But, regardless of her beliefs, the seller will never reduce the price below her reservation value, and the buyer agrees to trade only if the price is below his reservation value. Thus, in comparison to the fixed price mechanism, the price cap mechanism facilitates more efficient trade.[3]

The "price cap mechanism" is a simple mechanism for the buyer. The buyer accepts the trade if and only if a price below her reservation price is offered. But the seller's problem is arguably not too complicated either. If she believes that the buyer accepts the trade if and only if a price below her reservation price is offered, then all that she has to do is to consider her belief about the buyer's reservation price, and then maximize her expected gain from trade. This problem is equivalent to the standard monopoly problem with a price ceiling as taught in undergraduate microeconomics. For given beliefs, it is a straightforward optimization problem. Our formal definition of "simplicity" will imply that the price cap mechanism is indeed simple.

On the other hand, suppose buyer and sellers were engaged in a double auction as in Chatterjee and Samuelson [6]. The double auction will, in

[3]This mechanism was discussed in Börgers and Smith [5].

our definition, not be "simple." Suppose the seller's reservation price is 5 Dollars. She might ask for more. Ideally, she would like to ask for a price that is as close as possible to the price that is offered by the buyer. But the buyer has the same incentive: he wants to offer a price as close as possible to the seller's ask price. Therefore, the seller has to think about which price the buyer expects the seller to ask for. But then the buyer has to think about the seller's thoughts about which price the buyer expects the seller to ask for, and from here onwards, an infinite hierarch of strategic thinking becomes necessary. Note that this infinite hierarchy of thought was not needed in the price cap mechanism.

In this paper we formally define and investigate a property of mechanisms that we call "simplicity," and that is meant to capture the idea that, in "simple" mechanisms, strategic choices are easy. We shall define a mechanism to be "simple" if choices can be derived from first order beliefs alone, and there is no need for agents to form higher order beliefs because such beliefs are irrelevant to agents' optimal choices. Intuitively, it seems plausible that the need to form beliefs about beliefs, beliefs about beliefs about beliefs, beliefs about beliefs about beliefs about beliefs, etc., is one key reason why making optimal choices in games may often be difficult and require a lot of strategic thinking.

We have to be more specific about which beliefs we are referring to. The "first order beliefs" to which we are referring are every player i's beliefs about the other agents' ($j \neq i$) utility functions, and about other agents' rationality. "Higher order beliefs" are, for example, agent i's beliefs about agent j's beliefs ($j \neq i$) about agent i's utility function, and about agent i's rationality. Thus, we shall call a mechanism simple if player i's beliefs about the other agents' ($j \neq i$) utility functions, and about other agents' rationality, alone, imply which choices are optimal for player i.

To give this definition meaning, we have to say what is the range of first order beliefs that we are willing to consider for agent i. We shall allow a variety of choices for this range, each of which will imply a different concept of "simplicity," but what will be common is that we are assuming that agent i is certain that all other agents are expected utility maximizers.[4] Thus, we

[4] In fact, we shall replace this later by the assumption that agent i is "almost certain" that all other agents are expected utility maximizers. But for the moment we ignore this complication.

can say that we call a mechanism simple if agent i's beliefs about other agents' utility functions, combined with the probability 1 belief that the other agents are rational, is sufficient to determine an optimal choice for agent i.

The previous paragraph contained a very informal use of the language of epistemic game theory. After defining our concept of simplicity precisely, we shall return to its epistemic foundations. Our intuitive motivation of the notion of simplicity presented in this Introduction also involves wording that suggests that we view the choice of a strategy as a mental process that involves "forming beliefs" and "determining optimal choices," and where these steps are potentially costly to the decision maker. This is language not commonly used by game theorists. We shall also return to this aspect of our motivation later in the paper.

We shall define simplicity formally in the next section, and we illustrate the definition in Section 3 with examples. In Section 4 we provide the discussion of the conceptual foundations of our definition promised in the previous paragraph. In Section 5 we introduce "menu mechanisms" in which one agent offers the other agent a set of options, and the other agent then chooses one of these options. Menu mechanisms are a generalization of the price cap mechanism mentioned earlier in this Introduction. All menu mechanisms are simple. We show in Section 5 that with two agents and one-sided uncertainty, all simple mechanisms are equivalent in some sense to menu mechanisms. In Section 6 we turn to a special type of simplicity called "universal simplicity." This property combines simplicity with robustness, as it has been defined in the recent literature on robust mechanism design. We provide necessary and sufficient conditions for a mechanism to be "universally simple." Section 6 also offers examples of universally simple mechanisms. It turns out, that all menu mechanisms are universally simple, but there are more universally simple mechanisms than just menu mechanisms. Section 7 discusses an example in detail, and Section 8 concludes with a list of some open questions.

2. Definitions

There are n agents: $i \in I = \{1, 2, \ldots, n\}$, and a finite set A of outcomes. A mechanism consists of finite strategy sets S_i for each agent i, and a function $g : S_1 \times S_2 \times \ldots \times S_n \to A$ that describes for each choice of strategies which outcome will result. We define $S \equiv \prod_{i \in I} S_i$, and, for every $i \in I$, we define S_{-i} to be the set $\prod_{j \neq i} S_j$.

Note that we do not assume that there is a one-to-one correspondence between strategies and preferences, or perhaps, beliefs and preferences. In other words, we allow indirect mechanisms and do not restrict attention to direct mechanisms. In the final section of the paper we describe a version of the revelation principle that applies to our model, and that implies that it is without loss of generality to restrict attention to a certain class of direct mechanisms. We explain there as well that for the analysis offered in this paper, it does not seem to be particularly useful to rely on this revelation principle. Two other reasons for not restriction attention to direct mechanisms are as follows. First, mechanisms such as the price cap mechanism in the Introduction are most naturally described as indirect mechanisms. Second, below we allow infinite sets of possible preferences and beliefs for each agent. We would therefore have to consider infinite mechanisms if we wanted to use the revelation principle. This would prevent us from relying on the simplest version of certain simple results in game theory that we invoke below..

A utility function is a function $u : A \to \mathbb{R}$. We interpret utility functions as von Neumann Morgenstern utility functions. We define $\mathcal{U}$ to be the set of all utility functions on A that are 0-1 normalized, that is, that satisfy: $\min_{a \in A} u_i(a) = 0$ and $\max_{a \in A} u_i(a) = 1$. By choosing this normalization we rule out the uninteresting case that agents are indifferent between all alternatives.

For every agent i there is a non-empty, Borel-measurable set $\mathbf{U}_i \subset \mathcal{U}$ of utility functions that are possible utility functions of agent i. We allow for the possibility that $\mathbf{U}_i \neq \mathcal{U}$ to be able to capture restrictions of the sort that all agents have quasi-linear utility functions, a restriction that sometimes plays an important role in the theory of mechanism design. We define $\mathbf{U} \equiv \prod_{i \in I} \mathbf{U}_i$, and, for every $i \in I$, we define $\mathbf{U}_{-i}$ to be the set $\prod_{j \neq i} \mathbf{U}_j$.

A "utility belief" μ_i of agent i is a Borel probability measure on $\mathbf{U}_{-i}$. We define $\Delta(\mathbf{U}_{-i})$ to be the set of all Borel probability measures on $\mathbf{U}_{-i}$. The set of all possible utility beliefs of agent i is some non-empty subset $\mathbf{M}_i$ of $\Delta(\mathbf{U}_{-i})$. We allow that $\mathbf{M}_i \neq \Delta(\mathbf{U}_{-i})$ to be able to capture restrictions of the sort that all agents believe that other agents' utility functions are stochastically independent, a restriction that sometimes plays an important role in the theory of mechanism design. We define $\mathbf{M} = \prod_{i \in I} \mathbf{M}_i$, and, for every $i \in I$, we define $\mathbf{M}_{-i}$ to be the set $\prod_{j \neq i} \mathbf{M}_j$.

For any given $u_i \in \mathbf{U}_i$ we denote by $UD_i(u_i)$ the set of all strategies that are not weakly dominated for agent i with utility function u_i, where weak dominance may be by a pure or by a mixed strategy. If $u \in \mathbf{U}$, we define $UD(u) \equiv \prod_{i \in I} UD_i(u_i)$, and, for every $i \in I$ and every $u_{-i} \in U_{-i}$, we define $UD_{-i}(u_{-i}) \equiv \prod_{j \neq i} UD_j(u_j)$.

Weak dominance will play an important role in the definitions below. We shall discuss motivations and justifications for this in Section 4 below.

A "strategic belief" $\hat{\mu}_i$ of agent i is a probability measure on S_{-i}. The set of all such probability measures is denoted by $\Delta(S_{-i})$. Given a utility function $u_i \in \mathbf{U}_i$ of agent i, and a strategic belief $\hat{\mu}_i$ of agent i, we denote by $BR_i(u_i, \hat{\mu}_i)$ the set of all strategies in $UD_i(u_i)$ that maximize expected utility in this set.

Definition 1. *A strategic belief $\hat{\mu}_i$ on S_{-i} is "compatible" with a utility belief μ_i if there is a probability measure ν_i on*

$$\prod_{j \neq i} \{(u_j, s_j) \in \mathbf{U}_j \times S_j | s_j \in UD_j(u_j)\}$$

that has marginal μ_i on $\mathbf{U}_{-i}$ and marginal $\hat{\mu}_i$ on S_{-i}.

Intuitively, a strategic belief is compatible with a utility belief if and only if it can be constructed as follows: For each vector of utility functions in the support of the utility belief of agent i distribute the probability assigned to this utility function in some arbitrary way among the strategy combinations of the of the other players that only include strategies that are not weakly dominated for the given utility functions. Then add up for each strategy combinations of the other players the probabilities that have been assigned to that strategy. Thus, the strategic belief reflects the player's utility belief combined with the assumption that the player believes with probability 1 that the other players won't choose weakly dominated strategies.

We denote the set of all strategic beliefs that are compatible with a given utility belief μ_i by $\mathcal{M}_i(\mu_i)$.

Definition 2. *A mechanism is "strategically simple for agent i with utility function $u_i \in \mathbf{U}_i$ and with utility belief $\mu_i \in \mathbf{M}_i$" if*

$$\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i) \neq \emptyset.$$

A mechanism is "strategically simple for agent i given $\mathbf{U}_i$ and $\mathbf{M}_i$" if it is strategically simple for agent i for all utility functions $u_i \in \mathbf{U}_i$ and utility beliefs $\mu_i \in \mathbf{M}_i$. A mechanism is "simple given $\mathbf{U}$ and $\mathbf{M}$" if it is strategically simple for all agents i given $\mathbf{U}_i$ and $\mathbf{M}_i$.

Our definition of simplicity thus requires that there is *at least* one strategy of agent i that is optimal for all strategic beliefs that are compatible with a given utility belief. Intuitively, simplicity means that every agent i can forgo the cost of forming higher order beliefs at no loss in expected utility.

Mechanisms implement social choice correspondences. Traditionally, social choice correspondences map agents' preferences or utilities into selected outcomes. In our setting, it seems natural to extend the domain of social choice correspondences to also include agents' first order beliefs about the other agents' utility functions.

Definition 3. *A "social choice correspondence" is a correspondence:*

$$F : \mathbf{U} \times \mathbf{M} \twoheadrightarrow A.$$

The social choice correspondence implemented by a mechanism is:

$$F(u, \mu) = g\left(\prod_{i \in I}\left(\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i)\right)\right) \text{ for all } (u, \mu) \in \mathbf{U} \times \mathbf{M}.$$

Thus, a mechanism is simple if and only if it implements a non-empty valued social choice correspondence.

3. An Example

To illustrate the precise meaning of our definition, it is perhaps best to start with a simple example that, by itself, is not of particular interest in the theory of mechanism design. In this example $n = 2$, $S_1 = \{T, M, B\}$, $S_2 = \{L, C, R\}$, and $\mathbf{U_i} = \{u_i, \hat{u}_i\}$ for $i = 1, 2$. Each of the four boxes in Example 1 corresponds to one of the four possible vectors of utility functions. Instead of indicating for each combination of strategies which outcome will result, and then specifying utility functions by assigning a utility value to

	L	C	R
T	$1,1$	$0,0$	$1,0$
M	$0,1$	$1,0$	$0,0$
B	$0,0$	$0,1$	$0,0$

(u_1, u_2)

	L	C	R
T	$1,0$	$0,0$	$1,1$
M	$0,0$	$1,0$	$0,1$
B	$0,0$	$0,0$	$0,1$

$(u_1, \hat{u}_2)$

	L	C	R
T	$0,1$	$0,0$	$0,0$
M	$0,1$	$0,0$	$0,0$
B	$1,0$	$1,1$	$1,0$

$(\hat{u}_1, u_2)$

	L	C	R
T	$0,0$	$0,0$	$0,1$
M	$0,0$	$0,0$	$0,1$
B	$1,0$	$1,0$	$1,1$

$(\hat{u}_1, \hat{u}_2)$

Example 1

each outcome, we have combined these two steps, and indicated for each combination of strategies, for each possible vector of utility functions, the utility of each agent from the outcome that results if agents choose that combination of strategies. This allows us to suppress a specification of the set of outcomes A. We indicate in Example 1 the utility of agent 1 first, and then the utility of agent 2. Each agent's set of possible utility beliefs, $\mathbf{M}_i$, consists of only one probability measure, namely for $i = 1$ the probability measure μ_1 that assigns probability 0.4 to u_2, and for $i = 2$ the probability measure μ_2 that assigns probability 0.6 to u_1.

We shall argue that this example represents a simple mechanism. First observe that $UD_1(u_1) = \{T, M\}$, $UD_1(\hat{u}_1) = \{B\}$, $UD_2(u_2) = \{L, C\}$ and $UD_2(\hat{u}_2) = \{R\}$. The strategic beliefs of player 1 that are compatible with the given utility belief of agent 1 are all beliefs of the form: $\hat{\mu}_1(L) = x, \hat{\mu}_1(C) = 0.4 - x, \hat{\mu}_1(R) = 0.6$, where $x \in [0, 0.4]$. This is because, if agent 2 has utility function u_2, which according to the utility belief occurs with probability 0.4, agent 2 may either play L or C, so that agent 1 may divide the probability 0.4 in arbitrary ways among L or C when forming his strategic belief, but when agent 2 has utility function $\hat{u}_2$, which occurs with probability 0.6, player 2 will play R. Similarly, agent 2's strategic beliefs that are compatible with the given utility belief are all beliefs $\hat{\mu}_2$ of the form $\hat{\mu}_2(T) = y$, $\hat{\mu}_2(M) = 0.6 - y$, and $\hat{\mu}_2(B) = 0.4$, where $y \in [0, 0.6]$.

Now suppose player 1 has utility function u_1. Then the expected utility from strategy T under any strategic belief compatible with the given utility belief is $x \cdot 1 + (0.4 - x) \cdot 0 + 0.6 \cdot 1$, and the expected utility from strategy M is $x \cdot 0 + 0.4 - x \cdot 1 + 0.6 \cdot 0$. Thus, the set of best responses, $BR_1(u_1, \hat{\mu}_1)$, equals $\{T\}$ for all strategic beliefs compatible with the given utility belief, i.e. regardless of the value of x. Trivially, the intersection of the best response sets for all possible $\hat{\mu}_1$ is non-empty. Intuitively, even though expected utilities are affected by how much of the probability 0.4 player 1 allocates to L, and how much he allocates to C, the optimal choice is unaffected by this, because player 1 has to allocate the large probability 0.6 to the case that player 2 chooses R, and in that case T is optimal. We conclude that the mechanism is strategically simple for agent 1 with utility function u_1 and utility belief μ_1. We also have trivially: $BR_1(\hat{u}_1, \hat{\mu}_1) = \{B\}$ for all $\hat{\mu}_1$ that are compatible with the given utility belief, so that the mechanism is also trivially strategically simple for agent 1 with utility function $\hat{u}_1$ and utility belief μ_1.

Next consider player 2's perspective. If player 2 has utility function u_2, then his expected utility from L is $y \cdot 1 + (0.6 - y) \cdot 1 + 0.4 \cdot 0 = 0.6$, whereas it is $y \cdot 0 + (0.6 - y) \cdot 0 + 0.4 \cdot 1 = 0.4$ from strategy C. Thus, the set of best responses, $BR_1(u_2, \hat{\mu}_2)$ is $\{L\}$, for all strategic beliefs compatible with the given utility belief, i.e. regardless of what y is, and the simplicity criterion is again satisfied. Indeed, y does not enter player 2's expected payoff because the expected payoff difference between L and C is the same regardless of which of his two undominated strategies player 1 plays when he has utility function u_1. We conclude that the mechanism is thus strategically simple for player 2 with utility function u_2 and belief μ_2, and it is obvious that it is also strategically simple for player 2 with utility function $\hat{u}_2$ and belief μ_2. Thus, the mechanism is indeed simple.

Consider now the case in which for each agent the sets $\mathbf{U}_i$ have only a single element, and therefore also the sets $\mathbf{M}_i$ are singletons. These games resemble complete information games as commonly defined in game theory, except that, unlike the theory of complete information games, our theory does not assume any common knowledge. Therefore, we shall refer to these games as "certainty games." Lemma 1 applies to all such games. It says that such games are simple only if after eliminating for each agent the weakly dominated strategies then, in the remaining game, each agent has an always best response. This condition is obviously also sufficient. Thus, simple "certainty games" can be solved in two steps.

Example 1 is not of particular interest as an example in mechanism design because the sets of possible utility functions of the two agents are very small, having only two elements. Our definitions even allows the case that the sets $\mathbf{U}_i$ are singletons. Then, of course, also the sets $\mathbf{M}_i$ are singletons. In this case, mechanisms resemble complete information games as commonly defined in game theory, except that, unlike the theory of complete information games, our theory does not assume any common knowledge. In this case mechanisms are simple if and only if, after deleting weakly dominated strategies for each player, in the remaining game each player has a dominant strategy where we use the expression "dominant strategy" as in mechanism design, i.e. it refers to a strategy that is always optimal, regardless of what the other players do. Loosely speaking, when all sets $\mathbf{U}_i$ are singletons, simplicity is equivalent to dominance solvability in two steps.

4. Discussion

Our intuitive discussion of the notion of simplicity studied in the Introduction invoked the concepts of first and higher order beliefs about other players' utility functions and about their rationality. This suggests that our approach to simplicity should be describable in the language of epistemic game theory. To a certain extent it is.

Consider the universal type space of epistemic game theory, assuming that for each agent i the space of uncertainty is $\mathbf{U}_{-i} \times A_{-i}$. Suppose we consider in the universal type space all states in which the following are true: (i) agent i is an expected utility maximizer; (ii) agent i has utility function u_i; (iii) agent i's beliefs about other agents' utility functions is given by μ_i, and (iv) agent i believes with certainty that all other agents are expected utility maximizers. Note that (iii) and (iv) pin down agent i's first order beliefs, and that there are no conditions for agent i's higher order beliefs. Thus, if in all states in which (i)-(iv) hold, player i were to make the same choice, then player i's first order beliefs, together with knowledge of his utility function and knowledge that he is rational, would allow us to predict player i's choice.

Had we defined simplicity in this way, then it would be straightforward to show that a game is strategically simple for player i with utility function u_i and utility belief μ_i if and only if the set

$$\arg\max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}) \hat{\mu}_i(a_{-i})$$

has the same, single element for all strategic beliefs $\hat{\mu}_i$ for which there is a probability measure ν_i on

$$\prod_{j \neq i} \{(u_j, s_j) \in \mathbf{U}_j \times S_j | s_j \in RAT_j(u_j)\}$$

that has marginal μ_i on $\mathbf{U}_{-i}$ and marginal $\hat{\mu}_i$ on S_{-i}. Here, we define

$$RAT_j(u_j) = \bigcup_{\hat{\mu}_i \in \Delta(\mathcal{U}^{n-1})} \arg\max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}) \hat{\mu}_i(a_{-i}).$$

The condition in our definition differs in two ways from the condition that what we have just described. The first is that in our definition "rationality" means not only expected utility maximization, but also as not playing a weakly dominated strategy. This first point lead us to consider the sets $BR_i(u_i, \hat{\mu}_i)$ for $\hat{\mu}_i \in \mathcal{M}_i(\mu_i)$ rather than the sets described above. The second is that, rather than requiring these sets to always have the same, single element for all beliefs that we consider, we have required that the sets have at least one element in common.

The first point is primarily for practical reasons, but it can be seen . Some of our examples are most naturally interpreted as the normal forms of extensive form games, and by ruling out weakly dominated strategies we rule out some strategies that violate the most basic versions of sequential rationality. Note that the primary objection against the use of weakly dominated strategies, that the order of elimination of weakly dominated strategies matters, does not arise in our setting, because we do not iterate the elimination of weakly dominated strategies. We could provide an epistemic interpretation of our use of weak dominance. For this, we would replace point (iv) above by two conditions: every agent's strategic beliefs have full support on the product of the other players' strategy sets, and every agent attaches probability p to the hypothesis that all other players are expected utility maximizers with strategic beliefs that have full support. We would then consider the limit for p tending to 1. We do not elaborate this argument.[5] Note that the second of the two new assumptions does involve a condition on second order beliefs: every agent believes that every agent has full support beliefs. One might thus argue that implicitly our construction does include a (perhaps mild) restriction for second order beliefs.

[5]A related argument, but concerning common p-belief of rationality with full support beliefs in games, has been formalized by Frick and Romm [7].

The second difference between our definition of simplicity in this paper, and the definition outlined above, allows us to apply our definition even in cases in which an agents' beliefs and preferences are, in knife-edge cases, such that the agent is indifferent between several optimal actions, so that the set of best responses cannot be a singleton, and is also perhaps non-robust under belief perturbations.

Intuitively, one might also motivate our requirement in Definition 2 that the intersection of the best response sets be non-empty by the argument that the formation of beliefs, and the determination of best responses, as a costly intellectual process. Thus, once an agent discovers that there is one strategy that is optimal regardless of the agent's higher order beliefs, the agent may stop forming beliefs and determining best responses, and simply choose that strategy. The idea that belief formation is a costly process has been advocated, for example, by Binmore [2, pp. 129-132], who argues that achieving the consistency that is necessary for having a well-defined prior requires many costly iterations of attempted belief formation.[6]

5. Menu Games

Our first and straightforward result describes a necessary condition for a mechanism to be simple.

Lemma 1. *Consider a mechanism that is strategically simple given* $\mathbf{U}$ *and* $\mathbf{M}$. *Suppose for some agent* $i \in I$ *the set* $\mathbf{M}_i$ *contains for some* $u_{-i} \in \mathbf{U}_{-i}$ *the probability measure* μ_i *that places probability 1 on* u_{-i}. *Then for every* $u_i \in \mathbf{U}_i$ *there is a strategy* $s_i \in S_i$ *such that*

$$u_i(g(s_i, s_{-i})) \geq u_i(g(s_i', s_{-i}))$$

for all $s_{-i} \in \prod_{j \neq i} UD_j(u_j)$ *and all* $s_i' \in S_i$.

Proof. For every $u_{-i} \in \mathbf{U}_{-i}$ let μ_i be the probability measure that places probability 1 on u_{-i}. Let $s_{-i} \in \prod_{j \neq i} UD_j(u_j)$. Then $\hat{\mu}_i$, the probability measure that places probability 1 on s_{-i} is compatible with μ_i, and is therefore contained in $\mathcal{M}_i(\mu_i)$. This implies that every strategy in

$$\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i)$$

[6]Binmore asserts that in many circumstances, including games, it is actually impossible to carry out this process.

	L	C_1	C_2	R
U	a	a	b	b
D	b	c	b	c

EXAMPLE 2

has the property described in Lemma 1. For simple mechanisms there is at least one such strategy. $\square$

We shall describe the property of the strategy s_i in Lemma 1 as being an "always best response to strategies in $\prod_{j\neq i} UD_j(u_j)$." Note that the strategy s_i in Lemma 1 may depend on $\prod_{j\neq i} UD_j(u_j)$.

An interesting class of games to which Lemma 1 applies are two agent games with one-sided incomplete information, that is, games in which $n = 2$, $\#\mathbf{U}_1 = 1$, and $\#\mathbf{U}_2 > 1$. Lemma 1 implies that if such a game is to be simple, after agent 1's weakly dominated strategies are eliminated, for every utility function $u_2 \in U_2$ agent 2 has to have an always best reply. We now demonstrate that in certain cases all simple mechanisms are in a certain sense equivalent to the normal form of a game in which agent 1 offers agent 2 a menu of alternatives to choose from, and then agent 2 chooses one of these alternatives.

Definition 4. *Let $n = 2$. A mechanism is a "menu mechanism" if for some subset $\mathcal{A}$ of the set of all non-empty subsets of A, the mechanism is the normal form of the following perfect information game: First, agent 1 chooses some element A_1 (a "menu") of $\mathcal{A}$. Then agent 2 picks an element of A_1.*

Example 2 is an example of a menu-mechanism. It is the normal form of the following mechanism: first agent 1 chooses either the set $\{a, b\}$, or $\{b, c\}$, and then agent 2 picks an alternative from the set chosen by agent 1. Agent 2's column choices in Example 2 correspond to contingent plans, specifying both what to choose from $\{a, b\}$ and what to choose from $\{b, c\}$.

Proposition 1. *Suppose $n = 2$, $\#\mathbf{U}_1 = 1$ and for all $u_2 \in \mathbf{U_2}$: $u_2(a) \neq u_2(a')$ for all $a, a' \in A$ with $a \neq a'$. Then for every simple mechanism there is a menu mechanism that implements the same social choice correspondence.*

Proof. Consider any simple mechanism. Define a corresponding menu mechanism as follows: The "menus" agent 1 can choose from are all sets $A_1 \subseteq A$ for which there is some $s_1 \in S_1$ such that: $A_1 = \{a \in A | g(s_1, s_2) =$

a for some $s_2 \in S_2\}$. Agent 2's strategies are all possible functions that assign to each menu A_1 a choice $a \in A_1$ from that menu. We claim that this menu mechanism implements the same social choice correspondence as the given simple mechanism. To see this note that, in the menu game, for each utility function $u_2 \in \mathbf{U_2}$, agent 2 will choose a strategy that assigns to each menu A_1 the alternative that maximizes u_2 in A_1. By assumption there is a unique such alternative. Therefore, agent 1's expected utility for every choice s_2 in the original mechanism is the same as agent 1's expected utility for the corresponding menu in the menu mechanism. This implies the Proposition. $\square$

Let us consider an application of Proposition 1. Suppose a monopolist faces a single buyer with utility $v-p$ if the buyer buys an indivisible product from the monopolist and pays price p, and utility 0 otherwise. We also assume that when $v = p$, the buyer strictly prefers to trade. The monopolist does not know v. All values of v between 0 and 1 are possible, and all probability distributions on $[0, 1]$ are possible beliefs of the seller. The buyer knows that the monopolist does not attach any value to the object, and that the seller maximizes expected revenue. Suppose we are interested in simple mechanisms that seller and buyer can use to determine whether, and, if so, at which price they trade. Proposition 1 implies that it is without loss of generality to menu mechanisms.

Let us assume that we only consider mechanisms such that for every strategy of the seller there exists at least one strategy of the buyer such that the outcome is no trade, and the buyer does not make any payment to the seller. Thus, in a menu mechanism, every menu that the seller might offer must include this outcome. In any menu there only two alternatives that might potentially maximize the buyer's utility: the outcome of no trade, and the outcome at which the object changes hands at the lowest possible price in that menu. This means that we can think of menu mechanisms as price setting games: The seller proposes a price from some given set of admissible prices. The buyer chooses a function $f : \mathbb{R}_+ \rightarrow \{Y, N\}$ indicating for each possible price whether he wants to say "yes" or "no" to that price. Trade comes about at the price proposed by the seller if the buyer says "yes" to the price proposed by the seller, and otherwise no trade comes about. After weakly dominated strategies are eliminated, in this game the buyer has an always optimal strategy for every possible value of v: accept the trade if $p \geq v$, and reject it otherwise. The seller's optimal choice does not rely on second order beliefs about the buyer. It only depends on her beliefs about v.

6. Universal Simplicity

A class of mechanisms of particular interest is the class of all mechanisms that are simple when *all* conceivable utility functions and beliefs are considered. We call such mechanisms "universally simple." For universally simple games the mechanism designer's analysis of the mechanism is robust to changes in agents' belief hierarchies, and is thus robust in the sense of Bergemann and Morris [1].

Definition 5. *A mechanism is "universally simple" if it is strategically simple for all agents i given $\mathbf{U}_i = \mathcal{U}$ and $\mathbf{M} = \Delta(\mathcal{U}^{n-1})$.*

In this section we give a characterization of universally simple mechanisms. We first need some additional terminology and notation.

Definition 6. *Let R_i be a linear, that is, total, transitive, and anti-reflexive, order on A. A strategy $s_i \in S_i$ of agent i is called "weakly dominated given R_i" if there is another strategy $\hat{s}_i \in S_i$ such that for all $s_{-i} \in S_{-i}$*

$$g(\hat{s}_i, s_{-i}) R_i g(s_i, s_{-i}),$$

and, for some $s_{-i} \in S_{-i}$

$$g(\hat{s}_i, s_{-i}) R_i g(s_i, s_{-i}) \text{ and } g(\hat{s}_i, s_{-i}) \neq g(s_i, s_{-i}).$$

We denote by $UD_i(R_i) \subseteq S_i$ the set of all strategies of agent i that are not weakly dominated given R_i.

When a list of linear orders $(R_1, R_2, \ldots, R_n)$ is given, we denote, for every $i \in I$, by $UD_{-i}(R_{-i})$ the set $\prod_{j \neq i} UD_j(R_j)$.

Theorem 1. *A mechanism is universally simple if and only if the following condition is satisfied: If for every $i \in I$ R_i is a linear order on A, then there is some agent i^* such that for every strategy $s_{i^*} \in UD_{i^*}(R_{i^*})$ there is an alternative $a \in A$ such that*

$$g(s_{i^*}, s_{-i^*}) = a \text{ for all } s_{-i^*} \in UD_{-i^*}(R_{-i^*}).$$

In words the condition that is necessary and sufficient for universal simplicity says the following. Whenever we fix a vector of preferences $(R_1, R_2, \ldots, R_n)$, if we consider the mechanism restricted to the strategy sets $UD_i(R_i)$ for all $i \in I$, then, in the restricted mechanism, some agent i^* is a dictator, that is, for each of the alternatives that are possible when agents choose their strategies from $UD_i(R_i)$ agent i^* has an action that enforces that alternative if all other agents choose from $UD_i(R_i)$, and each of agent i^*'s actions enforces some alternative.

Proof. Sufficiency is obvious. We only prove necessity. We proceed by establishing a sequence of claims.

CLAIM 1: Let $u_i \in \mathbf{U}_i$, $u_{-i} \in \mathbf{U}_{-i}$, and let μ_i be a utility belief such that $\mu_i(\{u_{-i}\}) > 0$. Suppose $s_i, s_i' \in \bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i)$. Then for all $s_{-i}, s_{-i}' \in UD_{-i}(u_{-i})$:

$$u_i(g(s_i, s_{-i})) - u_i(g(s_i', s_{-i})) = u_i(g(s_i, s_{-i}')) - u_i(g(s_i', s_{-i}')).$$

PROOF OF CLAIM 1: Suppose the assertion were not true. Then there are $s_{-i}, s_{-i}' \in UD_{-i}(u_{-i})$ such that:

$$u_i(g(s_i, s_{-i})) - u_i(g(s_i', s_{-i})) > u_i(g(s_i, s_{-i}')) - u_i(g(s_i', s_{-i}')).$$

Pick any $\hat{\mu}_i \in \mathcal{M}_i(\mu_i)$ that places strictly positive probability on s_{-i} and s_{-i}'. Because s_i and s_i' are both in $BR_i(u_i, \hat{\mu}_i)$ both strategies must yield the same expected utility under $\hat{\mu}_i$. Now suppose we vary $\hat{\mu}_i$ such that it places ε probability more than $\hat{\mu}_i$ on s_{-i} and ε probability less than $\hat{\mu}_i$ on s_{-i}', leaving all other probabilities unchanged. If we choose $\varepsilon > 0$ and sufficiently small, we can vary $\hat{\mu}_i$ in this way so that it remains an element of $\mathcal{M}_i(\mu_i)$, and so that for the modified belief s_i is a strictly better response than s_i'. This contradicts $s_i' \in \bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i)$. □

CLAIM 2: Let $u_i \in \mathbf{U}_i$, $u_{-i} \in \mathbf{U}_{-i}$, and let μ_i, μ_i' be any two utility beliefs such that $\mu_i(\{u_{-i}\}) > 0$ and $\mu_i'(\{u_{-i}\}) > 0$. Suppose $s_i \in \bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i)$ and $s_i' \in \bigcap_{\hat{\mu}_i' \in \mathcal{M}_i(\mu_i')} BR_i(u_i, \hat{\mu}_i')$. Then for all $s_{-i}, s_{-i}' \in UD_{-i}(u_{-i})$:

$$u_i(g(s_i, s_{-i})) - u_i(g(s_i', s_{-i})) = u_i(g(s_i, s_{-i}')) - u_i(g(s_i', s_{-i}')).$$

PROOF OF CLAIM 2: Claim 2 follows from repeated application of Claim 1 if we can find a sequence of utility beliefs of agent i, μ_i^k $(k = 2, \ldots, K)$, and strategies of agent i, s_i^k $(k = 1, 2, \ldots, K)$, where $K \geq 2$, such that

$s_i^1 = s_i$, $s_i^K = s_i'$, for every $k \in \{2, \dots, K\}$ the utility belief μ_i^k places positive probability on u_i, and for every $k \in \{2, \dots, K\}$ both s_i^{k-1} and s_i^k are elements of $\bigcap_{\hat{\mu}_i^k \in \mathcal{M}_i(\mu_i^k)} BR_i(u_i, \hat{\mu}_i^k)$. We shall construct such a sequence.

We set $s_i^1 = s_i$. For every $\alpha \in [0,1]$ define $\mu_i(\alpha) \equiv \alpha\mu_i + (1-\alpha)\mu_i'$. Define $\alpha^2 \equiv \sup\{\alpha \in [0,1] | s_i \in \bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i(\alpha))} BR_i(u_i, \hat{\mu}_i)\}$. Define s_i^2 to be any strategy in S_i that is an element of $\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i(\alpha^2+\varepsilon))} BR_i(u_i, \hat{\mu}_i)\}$ for a sequence of $\varepsilon > 0$ tending to zero. Then, by upper hemicontinuity of the correspondence of best responses, s_i^1 and s_i^2 are both contained in $\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i(\alpha^2))} BR_i(u_i, \hat{\mu}_i)\}$. We define μ_i^2 to be $\mu_i(\alpha_2)$. Note that, because μ_i and μ_i' attach strictly positive probability to u_{-i}, and because μ_i^2 is a convex combination of μ_i and μ_i', also μ_i^2 places strictly positive probability on u_{-i}.

Now suppose for $k \geq 2$ we had already constructed strategy s_i^k and a corresponding utility belief μ_i^k and suppose $\mu_i^k = \mu_i(\alpha^k)$. Define $\alpha^{k+1} \equiv \sup\{\alpha \in (\alpha^k, 1] | s_i^k \in \bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i(\alpha))} BR_i(u_i, \hat{\mu}_i)\}$. Define s_i^{k+1} to be any strategy in S_i that is an element of $\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i(\alpha^{k+1}+\varepsilon))} BR_i(u_i, \hat{\mu}_i)\}$ for a sequence of $\varepsilon > 0$ tending to zero. Then, by upper hemicontinuity of the correspondence of best responses, s_i^k and s_i^{k+1} are both contained in $\bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i(\alpha^{k+1}))} BR_i(u_i, \hat{\mu}_i)\}$. We define μ_i^{k+1} to be $\mu_i(\alpha_{k+1})$. Note that μ_i^{k+1} places strictly positive probability on u_{-i}.

Note that by construction, in the sequence of strategies no strategy is ever repeated. Moreover, the sequence α^k is strictly increasing and will reach 1 in a finite number of steps. If at some point while $\alpha^k < 1$ we can choose $s_i^k = s_i'$, then we stop the construction, and set $k = K$. Alternatively, the construction ends when $\alpha^k = 1$. At that point we define $K = k+1$, $\alpha^{k+1} = 1$, and $s^{k+1} = s_i'$. Our sequence will then have the required properties. $\square$

Claim 3: For every agent i, for every linear order R_i on A, there exists a utility function u_i^* that represents R_i, such that for every $s_i \in UD(R_i)$ and for every u_{-i} there exists a strategic belief $\hat{\mu}_i$ that is compatible with a utility belief μ_i with $\mu_i(u_{-i}) > 0$ and such that:

$$BR_i(u_i^*, \hat{\mu}_i) = \{s_i\}.$$

Moreover, the utility functions u_i^* can be chosen such that $u_i^*(a) - u_i^*(b) \neq u_i^*(c) - u_i^*(d)$ for all $(a,b), (c,d) \in A^2$ with $(a,b) \neq (c,d)$.

Proof of Claim 3: By the Lemma and the remark in the first paragraph of the proof of that Lemma, in Börgers [3], for every strategy $s_i \in UD_i(R_i)$ there exist a utility function u_{s_i} that represents R_i (that is, $u_i(a) \geq u_i(b) \Leftrightarrow aR_ib$), and a full support strategic belief μ_i, such that s_i is the unique maximizer of expected utility given that belief. We can construct a utility belief with which μ is compatible, and, because μ_i has full support, and therefore gives strictly positive probability to a combination of strategies that are not weakly dominated given u_{-i}, we can find a utility belief with which μ_i is compatible, and that assigns strictly positive probability to u_{-i}.

It remains to be shown that the utility functions u_{s_i} can be chosen to be the same for all strategies $s_i \in UD_i(R_i)$. This follows directly from the argument in the proof of Proposition 1 in Weinstein [11] if we can show that there are a regular utility function u_i^* and, for every $s_i \in UD_i(R_i)$, a concave function $f_{s_i} : \mathbb{R} \rightarrow \mathbb{R}$, such that $u_i^* = f_{s_i}(u_{s_i})$ for all $s_i \in UD_i(R_i)$. This assertion is sufficient because, by Weinstein's argument, if we subject utility to a concave transformation, then the set of not weakly dominated strategies cannot shrink. We shall therefore now proceed to prove this assertion.

We first construct u_i^*. Enumerate the elements of A as $a_1, a_2, \ldots, a_L$ such that $a_L R_i a_{L-1} R_i a_{L-2} R_i \ldots R_i a_1$. We pick u_i^* to satisfy the following, where the first two lines are a normalization:

$$\begin{aligned} u_i^*(a_1) &= 0 \\ u_i^*(a_2) &= 1 \\ &\cdots \end{aligned}$$

$$u_i^*(a_{\ell-1}) < u_i^*(a_\ell) < u_i^*(a_{\ell-1}) + \ldots$$

$$\ldots (u_i^*(a_{\ell-1}) - u_i^*(a_{\ell-2})) \min_{s_i \in U(R_i)} \frac{u_{s_i}(a_\ell) - u_{s_i}(a_{\ell-1})}{u_{s_i}(a_{\ell-1}) - u_{s_i}(a_{\ell-2})}.$$

Note that the right most term in the inequality is strictly larger than the left term, so that u_i^* can be constructed, and will be monotonically increasing, and thus compatible with R_i. Because u_i^* is defined by inequalities, it is also clear that we can choose u_i^* so that it is regular.

We now turn to the construction of the functions f_{s_i}. For every s_i we set $f_{s_i}(u_{s_i}(a_\ell)) = u_i^*(a_\ell)$ for all $\ell = 1, 2, \ldots, L$. This defines f_{s_i} for a finite number of elements of $\mathbb{R}$ only. However, it is clear that we can extend f_{s_i} to a concave piecewise linear function on $\mathbb{R}$ if it satisfies the following concavity condition for the points in which it is defined:

$$\frac{f_{s_i}(u_{s_i}(a_\ell)) - f_{s_i}(u_{s_i}(a_{\ell-1}))}{u_{s_i}(a_\ell) - u_{s_i}(a_{\ell-1})} \leq \frac{f_{s_i}(u_{s_i}(a_{\ell-1})) - f_{s_i}(u_{s_i}(a_{\ell-2}))}{u_{s_i}(a_{\ell-1}) - u_{s_i}(a_{\ell-2})}$$

for all $\ell \geq 2$. By the definition of f_{s_i}, this inequality is equivalent to:

$$\frac{u_i^*(a_\ell) - u_i^*(a_{\ell-1})}{u_{s_i}(a_\ell) - u_{s_i}(a_{\ell-1})} \leq \frac{u_i^*(a_{\ell-1}) - u_i^*(a_{\ell-2})}{u_{s_i}(a_{\ell-1}) - u_{s_i}(a_{\ell-2})} \Leftrightarrow$$

$$u_i^*(a_\ell) \leq u_i^*(a_{\ell-1}) + \dots$$

$$\dots (u_i^*(a_{\ell-1}) - u_i^*(a_{\ell-2})) \frac{u_{s_i}(a_\ell) - u_{s_i}(a_{\ell-1})}{u_{s_i}(a_{\ell-1}) - u_{s_i}(a_{\ell-2})}$$

which holds by construction. □

CLAIM 4: For every agent i, for every linear order R_i on A, and for every $u_{-i} \in \mathbf{U}_{-i}$ either

(i) there is for every strategy $s_i \in UD(R_i)$ an alternative a such that $g(s_i, s_{-i}) = a$ for all $s_{-i} \in UD_{-i}(u_{-i})$,

or

(ii) there is for every strategy combination $s_{-i} \in UD_{-i}(u_{-i})$ an alternative a such that $g(s_i, s_{-i}) = a$ for all $s_i \in UD_i(R_i)$,

or both.

PROOF OF CLAIM 4: If we represent R_i by the utility function u_i^* from Claim 3. Pick any two $s_i, s_i' \in UD(R_i)$. By Claim 3 there are a strategic belief $\hat{\mu}_i$ that is compatible with a utility belief μ_i with $\mu_i(u_{-i}) > 0$ such that: $BR_i(u_i^*, \hat{\mu}_i) = \{s_i\}$, and a strategic belief $\hat{\mu}_i'$ that is compatible with a utility belief μ_i' with $\mu_i'(u_{-i}) > 0$ such that: $BR_i(u_i^*, \hat{\mu}_i) = \{s_i\}$. Obviously, this implies $s_i \in \bigcap_{\hat{\mu}_i \in \mathcal{M}_i(\mu_i)} BR_i(u_i, \hat{\mu}_i)$ and $s_i' \in \bigcap_{\hat{\mu}_i' \in \mathcal{M}_i(\mu_i')} BR_i(u_i, \hat{\mu}_i')$. Therefore, by Claim 2 for all $s_{-i}, s_{-i}' \in UD_{-i}(u_{-i})$:

$$u_i^*(g(s_i, s_{-i})) - u_i^*(g(s_i', s_{-i})) = u_i^*(g(s_i, s_{-i}')) - u_i^*(g(s_i', s_{-i}')). \qquad (*)$$

Suppose first that the two sides in (*) equal zero. Because R_i has no indifferences, it follows that $g(s_i, s_{-i}) = g(s_i', s_{-i})$ for all $s_{-i} \in UD_{-i}(u_{-i})$. Suppose first that for some $a \in A$ we have: $g(s_i, s_{-i}) = a$ for all $s_{-i} \in UD(u_{-i})$. Then $g(s_i', s_{-i}) = a$ for all $s_{-i} \in UD_{-i}(u_{-i})$, and, moreover, for any other strategy $\tilde{s}_i \in UD_i(R_i)$ the equation (*) can hold only if there is some alternative $\tilde{a} \in A$ such that $g(\tilde{s}_i, s_{-i}) = \tilde{a}$ for all $s_{-i} \in UD(u_{-i})$. This is implied by: $u_i^*(a) - u_i^*(b) \neq u_i^*(c) - u_i^*(d)$ for all $(a, b), (c, d) \in A^2$ with $(a, b) \neq (c, d)$. Thus, we have obtained Case (i).

Next suppose that the two sides in (*) equal zero, but that $g(s_i, s_{-i}) \neq g(s_i, s'_{-i})$ for some $s_{-i}, s'_{-i} \in UD_{-i}(u_{-i})$. Then, if we replace s'_i in (*) by some other $\tilde{s}_i \in UD(R_i)$, (*) can only hold if for $\tilde{s}_i$ both sides are zero, and thus $g(s_i, s_{-i}) = g(\tilde{s}_i, s_{-i})$. Thus, we have obtained Case (ii).

Now suppose that the two sides in (*) do not equal zero. Using again the regularity of u^*_i, we can conclude that: $g(s_i, s_{-i}) = g(s_i, s'_{-i})$ and $g(s'_i, s_{-i}) = g(s'_i, s'_{-i})$ for all $s_{-i}, s'_{-i} \in UD(u_{-i})$. Another use of the regularity of u^*_i implies then that, if we replace in $*$ strategy s'_i by some other strategy $\tilde{s}_i \in UD(R_i)$, then (*) can only hold if $g(\tilde{s}_i, s_{-i}) = g(\tilde{s}'_i, s'_{-i})$ for all $s_{-i}, s'_{-i} \in UD(u_{-i})$. Thus, we are in Case (i). □

Claim 5: Suppose for every agent j we have a linear order R_j on A. Then, for every agent i, either

(i) there is for every strategy $s_i \in UD(R_i)$ an alternative a such that $g(s_i, s_{-i}) = a$ for all $s_{-i} \in UD_{-i}(R_{-i})$,

or

(ii) there is for every strategy combination $s_{-i} \in UD_{-i}(R_{-i})$ an alternative a such that $g(s_i, s_{-i}) = a$ for all $s_i \in UD_i(R_i)$,

or both.

Proof of Claim 5: Claim 5 follows from Claim 4 if we represent for each j with $j \neq i$ the linear order R_j by the utility function u^*_j referred to in Claim 3. □

Completing the Proof of Theorem 1: The claim is obviously true if there is an alternative a such that $g(s) = a$ for all $s \in UD(R)$. Therefore from now on we restrict attention in this proof to the case that there are two alternatives $a \neq b$ such that $g(s) = a$ for some $s \in UD(R)$ and $g(s') = b$ for some other $s' \in UD(R)$.

We shall say that agent $i \in I$ "has no influence" if for every $s_{-i} \in UD_{-i}(R_{-i})$ there is an $a \in A$ such that $g(s_i, s_{-i}) = a$ for all $s_i \in UD_i(R_i)$, and we shall say that agent i is a dictator if agent i has the property ascribed to agent i^* in Theorem 1. By Claim 5 every agent i either has no influence, or is a dictator.

Next note that it cannot be that there is more than one dictator. A dictator can enforce any of the alternatives contained in $\{g(s)|s \in UD(R)\}$. We have assumed that there are at least two such alternatives, say a and

	L	RL	RRL	RRR
L	a	a	a	a
RL	a	b	b	b
RRL	a	c	b	b
RRR	a	c	c	a

Example 3

b. Having two dictators leads to a contradiction if one of them chooses an action that enforces a, and the other one chooses an action that enforces b.

Finally note that it cannot be that all agents have no influence. Recall that we are considering the case in which there are two alternatives $a \neq b$ such that $g(s) = a$ for some $s \in UD(R)$ and $g(s') = b$ for some other $s' \in UD(R)$. Consider the sequence of n strategy combinations s^k obtained by switching sequentially first agent 1, then agent 2, etc. from strategy s_i to strategy s'_i. Thus, $s^1 = (s'_1, s_2, \ldots, s_n)$, $s^2 = (s'_1, s'_2, s_3 \ldots, s_n)$, etc. Define $s^0 = s$. Because $g(s^0) \neq g(s^n)$, there must be some k such that $g(s^k) \neq g(s^{k-1})$. But this means that by construction agent k has influence. Hence agent k must be a dictator. □

7. An Example

Theorem 1 suggests the following classification of universally simple mechanisms. Suppose we denote by $I^*(R_1, R_2 \ldots, R_n)$ the set of agents to whom the condition in Theorem 1 applies, that is, who are dictators, if the preference profile is $(R_1, R_2, \ldots, R_n)$. Then a first type of universally simple mechanisms is the type for which there is some i^* that is contained in $I^*(R_1, R_2 \ldots, R_n)$ for all $(R_1, R_2 \ldots, R_n)$. The second type of universally simple mechanisms in which no such single i^* exists.

The menu mechanism in Example 2, which we discussed earlier, is a universally simple mechanism of the first type. Example 3 is an example of a universally simple mechanism of the second type. There are two agents in Example 3, $i = 1, 2$, and three alternatives, $A = \{a, b, c\}$.

To analyze Example 3 we first determine for each agent, for each ordinal preferences of that agent, the set of weakly undominated strategies of that agent. We begin with agent 1. If agent 1 ranks a highest, L is weakly dominant. If agent 1 ranks b highest, RL is weakly dominant. If agent 1's preference is *cab*,[7] then *RRR* is weakly dominant. If agent 1's preference is *cba*, then *RRL* and *RRR* are the not weakly dominated strategies.

If agent 2 ranks a highest, L is weakly dominant. If agent 2's preference is *bac*, then *RRR* is weakly dominant. If agent 2's preference is *bca*, then *RRL* is weakly dominant. If agent 2's preference is *cab*, then L and *RL* are the not weakly dominated. If agent 2's preference is *cba*, then *RL* is weakly dominant.

We now show that the mechanism in Example 3 satisfies the condition in Theorem 1. Whenever one agent has a dominant strategy, then obviously the other agent is a dictator. It remains to consider the case in which neither agent has a dominant strategy, which is when agent 1 ranks alternatives *cba*, and agent 2 ranks alternatives *cab*. Then the undominated strategies are *RRL* and *RRR*, and L and *LR*. In this case agent 2 is a dictator, and thus by Theorem 1 this is a universally simple mechanism.

But note that agent 2 is not in all cases the dictator. If agent 1 has two undominated strategies, but agent 2 has a weakly dominant strategy, then obviously only agent 1 is the dictator. Thus, Example 3 is a universally simple mechanism of the second type.

We can also show directly that the mechanism in Example 3 is universally simple. For this, we have to show that agents can determine their expected utility maximizing strategies on the basis of first order beliefs about other agents' utility functions only. This is obvious if an agent has a weakly dominant strategy. We have thus only two cases in which there are multiple not weakly dominated strategies. The first is that agent 1's preference is *cba*. Then *RRL* and *RRR* are the not weakly dominated strategies. Without loss of generality, suppose agent 1's von Neumann Morgenstern utility function is given by: $u_1(c) = 1, u_1(b) = x$ and $u_1(a) = 0$. Denote by p the probability that agent 1 assigns to the event that agent 2's preference is *bac*, and thus that agent 2 chooses *RRR*, and denote by q the probability that agent 1 assigns to the event that agent 2's preference is *bca*, and thus that agent 2 chooses *RRL*. Agent 1 will choose *RRL* if and only if: $(p + q)x \geq q$. Note that for given x agent 1's optimal choice only depends on his first order belief, as required by simplicity.

[7]Here, and in the following, this notation means that agent 1 ranks c above a above b.

The second case in which there are multiple undominated strategies is that agent 2 has preferences *cab*. Without loss of generality assume agent 2's utility function is: $u_2(c) = 1, u_2(a) = y, u_2(b) = 0$, and assume that he attaches probability $\tilde{p} \in [0, 1]$ to the event that player 1 ranks b highest, and thus chooses RL, and probability r to the event that player 1 ranks c highest, and hence plays RRL or RRR. Then player 2 will choose L if and only if: $(q + r)y \geq r$. Again, we see that agent 2's optimal choice only depends on his first order belief, as required by simplicity. This completes the direct argument that Example 3 is a universally simple mechanism.

Example 3 is actually the reduced normal form of an extensive game that we show in Figure 1. The extensive game in Figure 1 is a game of perfect information. Strategy "L" of player 1 in Example 3 corresponds to the strategy in which player 1 chooses "left" at his initial node. Strategy "RL" of player 1 in Example 3 corresponds to the strategy in which player 1 chooses "right" at his initial node, but "left" at his second node. All other strategies are labelled in the analogous way.

It is also interesting to show the social choice correspondence implemented by Example 3. We show that correspondence in Figure 2. In Figure 2 rows correspond to preferences of agent 1, columns correspond to preferences of agent 2. Note that there are some Pareto inefficiencies. For example, when both preferences have preferences *cab*, then agent 2 might stop the game at her first move, thus opting for a, and this would be rational if she mistakenly believes that agent 1 will choose b if given the chance. This would result in outcome a, yet both agents prefer c over a.

When there is multiplicity in Figure 2, then it depends on players' von Neumann Morgenstern utility functions and their beliefs which alternative is picked. For example, when player 1's preferences are *cba*, and player 2's preferences are *bac*, then alternative b will be chosen if $(p + q)x \geq q$, and otherwise a will be chosen. This follows from the above calculations. Thus, b wil be chosen if agent 1's von Neumann Morgenstern utility from b is sufficiently high, and if agent 1's probability that player 2's preference is *bac* is high in comparison to his probability that agent 2's preference is *bca*. Otherwise, a will be chosen.

We can reinterpret Example 3 as a bilateral example. Suppose "a" stands for "no trade," "b" stands for "trade at a low price," and "c" stands for "trade at a high price." Suppose agent 1 is the seller, and agent 2 is the

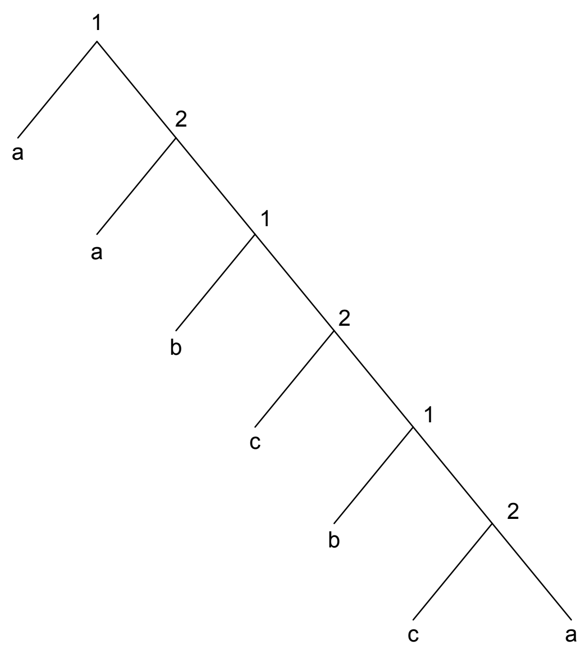

Figure 1: Extensive form for Example 3

buyer. Then it seems natural to consider the domain in which agent 1 always prefers c over b, and agent 2 always prefers b over c. Then, the implemented social choice correspondence is as shown in Figure 3.

The social choice correspondence implemented by Example 3 in the bilateral trade example has the following interpretation: if at least one player is unwilling to trade at any price, then no trade takes place. If the buyer only wants to trade at the high price, and the seller only wants to trade at the low price, then no trade takes place. If the buyer and the seller are willing to trade at both prices, then trade at either price is possible. If the seller is only willing to trade at the high price, but the buyer is willing to trade at both prices, then the trade takes place at the high price. If the buyer is only willing to trade at the low price, but the seller is willing to trade at

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

Figure 2: Social Choice Correspondence Implemented by Example 3

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

Figure 3: Social Choice Correspondence Implemented by Example 3 in the Bilateral Trade Example

both prices, then trade may take place at the low price, or it may not take place at all. This last case is an inefficiency. Our earlier calculations showed that this inefficiency occurs if the seller attaches a high probability to the event that the buyer is willing to buy at the high as well as the low price in comparison to the probability that the buyer is only willing to buy at the low price (as is actually the case).

The bargaining protocol described by Figure 1 is this: First, both agents are given the chance to quite the trade. Then the seller can accept the low price. If he does, trade takes place at that price. Otherwise, the buyer can accept the high price. If he does, then trade takes place at that price. Otherwise, the seller is given another chance to accept the low price. If he does, then trade takes place at that price. Otherwise, the buyer makes a final choice: he can either accept the high price, or quit the trading game.

The strategies that agents choose are as follows: If the seller prefers no trade to either the high or the low price, then the seller quits right away. If the seller is only willing to trade at the high price, then the seller participates, but continues to insist on the high price until the game ends either with the buyer accepting the high price, or rejecting it. If the seller is willing to trade at both prices, then the seller accepts the low price at the third decision node of the seller if the seller attaches a high probability to the event that the buyer is willing to buy only at the low price in comparison to being willing to trade at both prices. Otherwise, he again refuses the low price until the end.

If the buyer prefers no trade to either the high or the low price, then the buyer quits right away. If the buyer is only willing to trade at the low price, then she waits for the seller to accept the low price, and if the seller does not, then she refuses trade in the final round. If the buyer is willing to trade at both prices, then she also waits for the final round, but accepts trade in the final round.

Thus, an inefficient outcome arises if the seller is willing to accept either the high or the low price, but insists on the high price, believing that the buyer is likely willing to trade at the high price, and if, in fact, the buyer is only willing to trade at the low price, but not at the high price. In this case, no trade comes about, even though both sides would have preferred ex post to trade at the low price.

In comparison to the menu game, in which the seller can lower the price, in the extensive form game that corresponds to Example 3, the seller can first try to keep the price high, see the buyer refuse, and then, in response, lower the price. In the menu game, the seller has only one choice to lower the price. Another interesting feature of the extensive form in Figure 1 is that once both agents have chosen not to opt out at their initial nodes, the seller will never be forced to trade at the low price, if he doesn't want to, but he cannot avoid trade at the high price, if the buyer agrees to such a trade. Equally, the buyer can avoid trade at the high price, but he cannot avoid trade at the low price is the seller is willing to make such a trade.

8. Conclusion

This paper has proposed a new class of mechanisms. These mechanisms are "simple" to analyze for agents, and thus also for the mechanism designer. In Section 4 we have introduced a "robust" analysis of such mechanisms. A priority of our research agenda is to characterize in various settings the set of all simple mechanisms. In this context, this paper leaves a couple of questions open. In particular, in Section 3, it would be useful to characterize all simple mechanisms for the buyer-seller example with two-sided uncertainty.

In Section 4 it would be good to obtain a characterization of all universally simple mechanisms of the first type. Over the long run, our objective is to introduce a concept of optimality of simple mechanisms, and to study optimal simple mechanisms.

References

[1] Dirk Bergemann and Stephen Morris, Robust Mechanism Design, *Econometrica* 73 (2005), 1771-1813.

[2] Ken Binmore, *Rational Decisions*, Princeton and Oxford: Princeton University Press, 2009.

[3] Tilman Börgers, Pure Strategy Dominance, *Econometrica* 61 (1993), 423-430.

[4] Tilman Börgers, *An Introduction to the Theory of Mechanism Design*, Oxford: Oxford University Press, 2015.

[5] Tilman Börgers and Doug Smith, Robustly Ranking Mechanisms, *American Economic Review*, Papers and Proceedings 2012, 325-329.

[6] Kalyan Chatterjee and William Samuelson, Bargaining under Incomplete Information, *Operations Research* 31 (1983), 835-851.

[7] Mira Frick and Assaf Romm, Rational Behavior under Correlated Uncertainty, *Journal of Economic Theory* 160 (2015), 56-71.

[8] Kathleen Hagerty and William Rogerson, Robust Trading Mechanisms, *Journal of Economic Theory* 42 (1987), 94-107.

[9] Shenwu Li, Obviously Strategy-Proof Mechanisms, working paper, March 2016.

[10] David Pearce, Rationalizable Strategic Behavior and the Problem of Perfection, *Econometrica* 52 (1984), 1029-1050.

[11] Jonathan Weinstein, The Effect of Changes in Risk Attitude on Strategic Behavior, *Econometrica* 84 (2016), 1881-1902.