Journal of Economic Behavior and Organization 241 (2026) 107367

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Research Paper

Behavioral implementation by individual-based rights structures: A full characterization ⋆

Ville Korpela [a], Michele Lombardi [b,c], Julius Zachariassen [a,*]

[a] Turku School of Economics, University of Turku, Finland
[b] University of Liverpool Management School, Liverpool, UK
[c] Department of Economics and Statistics, University of Napoli Federico II, Italy

ARTICLE INFO

JEL classification:
C72
D11
D71
D82

Keywords:
Behavioral economics
Implementation
Individual-based rights structure
(Choice) unanimity

ABSTRACT

Behavioral implementation studies the social choice rules that a planner can implement when individuals' choice behavior need not be rational. Existing results cannot determine which social choice rules are behaviorally implementable when individuals' choices exhibit significant choice anomalies such as decoy effect, choice overload, and compromise effect. We improve the situation by fully identifying the class of social choice rules behaviorally implementable by individual-based rights structures. In this framework, the planner designs an individual-based rights structure, which specifies rights of individuals to change the status-quo outcome of the society. It turns out that the key to a full characterization is to understand which condition should replace unanimity when individuals' choices are non-rational.

1. Introduction

Individuals constantly interact with each other through various institutions, be they political, such as elections, or economic, such as markets. Implementation theory studies how to design such institutions in a way that they achieve a socially desirable outcome.

Typically, implementation theory assumes that agents' choice behavior is rational, i.e., it's consistent with the maximization of context-independent preferences. However, there is a wealth of evidence in marketing, psychology, and behavioral economics that individuals' choices need not be consistent with the maximization of a preference relation.[1] In order to capture non-rational choice behavior, we take individuals' choice behavior as the primitive, instead of preferences. That is, in our framework individuals have state-dependent choice correspondences, rather than state-dependent preferences.[2]

In much of the implementation literature, the object of design is a *mechanism*. A mechanism generates a power distribution in a society; it gives a single agent the right to change the current outcome by changing her/his strategy. A generalization of this idea

⋆ This paper originally circulated under the title "Behavioral implementation without unanimity". We thank two anonymous referees and the associate editor for comments that have greatly improved this paper. We also want to thank the participants of the 17th meeting of the Society for Social Choice and Welfare in Paris, the 2024 Conference on Mechanism and Institution Design in Budapest, and seminar participants at Aalto University for their helpful comments.

* Corresponding author.
E-mail addresses: vipeko@utu.fi (V. Korpela), michele.lombardi@liverpool.ac.uk (M. Lombardi), julius.v.zachariassen@utu.fi (J. Zachariassen).

[1] See, for example, Toffler (1970) for choice overload bias, Huber et al. (1982) for attraction effect and (Samuelson and Zeckhauser, 1988) for status quo bias.

[2] Although we will often specify preferences of individuals, this is only to illustrate how the choice behavior is biased.

https://doi.org/10.1016/j.jebo.2025.107367
Received 11 September 2024; Received in revised form 24 October 2025; Accepted 29 November 2025
Available online 5 December 2025

has been developed by Sertel (2001) and Koray and Yildiz (2018), who propose the notion of *individual-based rights structures* as a formalization of what a single agent can or cannot do in a society. This notion is a flexible tool for designing institutions such as constitutions, legal codes, and rules of conduct. With this approach, an implementation problem consists of designing an individual-based rights structure such that its equilibrium outcomes correspond to the outcomes prescribed by the planner. In solving this problem, the planner defines individuals' rights — that is, who may change the current status quo outcome and which alternative outcomes they may replace it with.

In a seminal paper, Koray and Yildiz (2018) show that a full characterization of *social choice rules* (henceforth designated as SCRs) implementable by individual-based rights structures under rational choice behavior is given by two well-known conditions — Maskin monotonicity together with unanimity.[3] Roughly speaking, Maskin monotonicity says that if the society wants to select an outcome in one state, and this outcome rises in the ranking of everyone at another state, then it must still be selected. Unanimity, on the other hand, says that if an outcome is considered the best among all feasible outcomes by everyone at some state, then it must be selected. Combining the result of Koray and Yildiz (2018) with that of de Clippel (2014), who shows that a condition called *consistency* is the natural counterpart of Maskin monotonicity when choice behavior can be non-rational, it is evident that for a full characterization of SCRs that are behaviorally implementable by individual-based rights structures, it is important to understand what is the property that should replace unanimity when choice behavior can be non-rational.

The literature on behavioral economics gives little guidance on this matter beyond one obvious candidate; *choice unanimity*: If an outcome is unanimously chosen from the set of all feasible outcomes at some state, then it must be accepted as socially desirable (i.e. the SCR must select it). When choice behavior is rational, choice unanimity, which in this case coincides with the unanimity used by Koray and Yildiz (2018), is a sensible property of the SCR. However, when choice behavior is non-rational, choice unanimity is neither a necessary condition for implementation nor a sensible property of the SCR. The following example illustrates this point.

Compromise effect: Suppose that three friends are planning to dine out at one of three possible restaurants open tonight in town. The restaurants offer similar food but of different qualities and prices. The first offers high quality albeit expensive food, the second offers medium quality food for a moderate price, and the third offers affordable but low quality food. Let us call them restaurant h, m and l, respectively. Each friend would select option h from $\{l, h\}$ and from $\{m, h\}$. Under rational choice behavior, each friend would also choose h from $\{l, m, h\}$. Since the friends unanimously consider h to be the best outcome, it is impossible to construct a mechanism or a rights structure where the friends would not go to h in at least some equilibrium. This does not matter since, arguably, h is a good choice anyway. However, if friends exhibit a behavioral bias known as the compromise effect, each would choose the middle option m from the set $\{l, m, h\}$ instead of h, although all would choose h from $\{l, h\}$ and $\{m, h\}$.[4] If this bias is known, then the right choice would still be h, although asking friends to select the best option from $\{l, m, h\}$ would not reveal it. This is problematic because the socially desirable restaurant is still h, while any sufficient condition that use choice unanimity would impose to the friends the choice of m.

Seminal papers on behavioral implementation are (Korpela, 2012) and (de Clippel, 2014). Both use the same generalization of Nash equilibrium to a choice function environment.[5] However, both papers fail to handle a large class of choice anomalies because their sufficiency results rely on choice unanimity. Indeed, de Clippel's (2014) main sufficiency result relies directly on choice unanimity, and Korpela's (2012) result relies on the assumption that individuals' choice behavior satisfies the so-called Sen's property α (Sen, 1970).[6] This paper studies behavioral implementation without relying on choice unanimity or Sen's property α.

de Clippel (2014) identifies conditions under which we can design a mechanism for which, at every state, the outcome of each behavioral (Nash) equilibrium coincides with the outcome prescribed by the SCR for that state. In particular, he identifies a key necessary condition—the existence of opportunity sets satisfying consistency with respect to SCR. In this paper, we find that a strengthening of this consistency condition provides a necessary and sufficient condition for behavioral implementation by individual-based rights structures. This strengthening reveals how we can avoid using choice unanimity in the characterization of implementable SCR.

Related literature. Hurwicz (1986) was the first to study implementation problems using choice functions instead of preference relations. He showed that results of Nash implementation do not essentially depend on the assumption of rational choice behavior. In this setup, Korpela (2012) pushes Hurwicz's analysis forward by providing a more suitable notion of behavioral (Nash) equilibrium and showing that the necessary and sufficient conditions of Moore and Repullo (1990) for Nash implementation critically hinge on Sen's property α. de Clippel (2014) studies implementation in behavioral equilibrium when individual's choice behavior can violate Sen's property α. He provides a necessary condition, which is also sufficient when combined with an auxiliary condition.

[3] Unanimity is a necessary condition in Koray and Yildiz (2018) because they impose no restrictions on the preferences that members of the society can have. The necessary condition becomes different if there is a logical way to rule out some of the preferences. This is given in Korpela et al. (2025).

[4] For more on compromise effect, see Example 3.

[5] Hurwicz (1986) was the first to generalize Nash equilibrium into a choice function environment. His solution concept requires that agents make pairwise comparisons of alternatives in the set that can be obtained by unilaterally deviating. A truly choice-based solution concept would require agents to choose simultaneously among all alternatives in this set. What Hurwicz assumes is called binariness in the literature on choice function; the behavior is defined by pairwise comparison (Sen, 1977).

[6] Sen's α is a contraction consistency property that rules out cyclical preferences. The property says that if an option a is chosen from a set X, and a is also available in a subset Y of X, then a must be chosen from Y as well. Thus, in the case of a mechanism, if each agent would select a from the range of the outcome function of the mechanism, they would also select a from any row that includes it, thus making it an equilibrium outcome. Therefore, to implement an SCR when Sen's α holds, choice unanimity must hold in some set (the range of the implementing mechanism).

When agents have incomplete information, Saran (2011) introduces an appropriate notion of equilibrium, and fully identifies the set of choice behavior (beyond rationality) over which the revelation principle holds. Barlo and Dalkıran (2023) fill the gap left in Saran (2011) by studying (full) implementation in behavioral interim equilibrium. Recently, Barlo and Dalkıran have made a lot of advances in the analysis of the choice function framework, studying ex-post behavioral implementation (Barlo and Dalkıran, 2024), robust behavioral implementation (Barlo and Dalkıran, 2023b), as well as some more subtle questions (Altun et al., 2023; Barlo and Dalkıran, 2022). In addition, Hagiwara (2025) studies implementation problems in behavioral subgame-perfect equilibrium, Hayashi et al. (2023) study implementation problems in behavioral strong (Nash) equilibrium, while Chambers and Yenmez (2017) and Caspari and Khanna (2025) study matching problems with nonstandard choice behaviors.[7] As for rights structures, Yildiz (2025) has explored the possibilities of implementation using two-stage rights structures.[8]

The rest of the paper is organized as follows. Section 2 provides formal definitions and explains how standard mechanisms and individual-based rights structures are connected. Section 3 discusses several choice anomalies and gives examples of SCRs that are behaviorally implementable by individual-based rights structures. Section 4 presents our necessary and sufficient condition called extended consistency. Section 5 studies an example on limited willpower and shows how the condition can be applied. Section 6 presents some results on convergence before Section 7 concludes.

2. The setup

The environment consists of a set of agents $N = \{1, \dots, n\}$, where n is at least 2, a (non-empty) set of outcomes X and a set of possible *states of the world* $\theta \in \Theta$. Let $\mathcal{X} = \{A \subseteq X \mid A \neq \emptyset\}$ be the collection of all non-empty subsets of X. Agents' choice behavior is dependent on the state of the world $\theta \in \Theta$. The choice behavior of agent $i \in N$ at state θ is described by a choice correspondence $C_i(\cdot\,;\theta) : \mathcal{X} \twoheadrightarrow \mathcal{X}$ that assigns a non-empty set of chosen outcomes $C_i(A, \theta) \subseteq A$ to each choice set $A \in \mathcal{X}$. Agent i's choice correspondence is *rational* at θ if there exists a complete and transitive preference relation $\succeq_i^\theta$ such that $C_i(A, \theta) = \{x \in A \mid x \succeq_i^\theta y \text{ for all } y \in A\}$ for each $A \in \mathcal{X}$. Henceforth, we denote by $C_i(\cdot\,;\succeq_i^\theta)$ the choice behavior generated by maximization with respect to the preference relation $\succeq_i^\theta$. We focus on complete information environments where the true state of the world is common knowledge among agents but unknown to the planner.

The goal of the planner is represented by a *social choice rule* (SCR) $F : \Theta \twoheadrightarrow X$ that maps each state of the world $\theta \in \Theta$ to a set of socially acceptable (or optimal) outcomes $F(\theta) \subseteq X$. Let $F(\Theta)$ denote the range of F. The domain Θ of an SCR F is called a *rational domain* if the choice correspondence of each agent is rational at all possible states of the world. Otherwise the domain is called a *behavioral domain*. The following is a central property of SCRs.

Definition 1. An SCR F satisfies **choice unanimity** if at all states $\theta \in \Theta$, and for all outcomes $x \in X$, if $x \in C_i(X, \theta)$ holds for all $i \in N$, then $x \in F(\theta)$.

In words, if all agents would select outcome x from the set of all possible outcomes at state θ, then x must be socially acceptable at that state (i.e., $x \in F(\theta)$). Although choice unanimity is a desirable property under rational choice behavior, it may not be desirable under non-rational choice behavior.

In standard implementation, the object of design is a *mechanism*, which is a pair (M, g), where $M = \Pi_{i\in N} M_i$ is the joint message space and $g : M \to X$ is the outcome function. A strategy for agent i is the choice of a message in M_i. Agent i's *opportunity set*, given a strategy profile m_{-i} for the other agents, is given by $O_i(m_{-i}) = \{g(m_i, m_{-i}) | m_i \in M_i\} \subseteq X$. This is the set of outcomes i can induce by a unilateral deviation, given m_{-i}.

Definition 2. A strategy profile $m^*(\theta)$ is a **behavioral (Nash) equilibrium** of the (strategic-form) game induced by the mechanism (M, g) at state θ if $g(m^*(\theta)) \in C_i(O_i(m^*_{-i}), \theta)$ holds for all $i \in N$.

Let us denote the set of all behavioral equilibria of a mechanism (M, g) at θ by $NE(M, g, \theta)$. The mechanism (M, g) *implements* the SCR F in behavioral equilibrium if $F(\theta) = g(NE(M, g, \theta))$ holds at all states $\theta \in \Theta$. If such a mechanism exists, the SCR is said to be *behaviorally implementable by a mechanism.*

In our framework, the object of design is an *individual-based rights structure* instead of a mechanism.

Definition 3. An individual-based rights structure is a tuple $\Gamma = (S, \gamma, h)$, where S is the set of all possible states of the society, $h : S \to X$ is the outcome function, and $\gamma : S \times S \to 2^{\mathcal{N}}$ (a possibly empty-valued correspondence) is an *individual-based code of rights*, where $\mathcal{N} = \{\{i\} : i \in N\}$ and $2^{\mathcal{N}}$ is the set of all possible subsets of $\mathcal{N}$.

Subsequently, for each pair of distinct states of the society $s, t \in S$, γ specifies the set of single agents that are entitled to replace s with t. The rights structure tells us what everyone can do given the current configuration of the society.

Given an individual-based rights structure Γ, for any state of the society $s \in S$, we define i's opportunity set at s, denoted by $O_i(s)$, as $O_i(s) \equiv \{h(t) \in X \mid \{i\} \in \gamma(s, t)\} \cup \{h(s)\}$. This is the set of outcomes that i can induce by unilaterally replacing s with another state of the society where i is entitled to move.

Definition 4. A state of the society s^* is a **behavioral equilibrium** of Γ at state θ if $h(s^*) \in C_i(O_i(s^*), \theta)$ holds for all $i \in N$.[9]

[7] This is an incomplete list. We also refer the reader to de Clippel (2022) and Saran (2016).

[8] For more on implementation via rights structures, see also Lombardi et al. (2023), Savva (2021a,b), and Korpela et al. (2025).

[9] Notice that, since rights structures are a generalization of mechanisms, Definition 2 refers to a behavioral equilibrium induced by a special class of rights structures. See footnote 16 for a further discussion of the connection between mechanisms and rights structures.

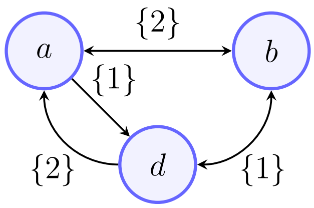

Fig. 1. Status quo bias.

Let us denote the set of all behavioral equilibria of Γ at θ by $BE(\Gamma, \theta)$. We say that an individual-based rights structure Γ *implements* SCR F if $h(BE(\Gamma, \theta)) = F(\theta)$ holds at all states $\theta \in \Theta$. If such an individual-based rights structure exists, the SCR is said to be *behaviorally implementable by an individual-based rights structure*.

How are individual-based rights structures and mechanisms related? On this front, Koray and Yildiz (2018) provide us with a result that illustrates the connection between individual-based rights structures and mechanisms. They show that implementation by individual-based rights structures is equivalent to implementation by *deviation-constrained mechanisms*. These are standard mechanisms (M, g) with an additional constraint function $\mathcal{D}_i : M \to M_i$ that maps each strategy profile m to a subset $\mathcal{D}_i(m) \subseteq M_i$ of messages that agent i is allowed to play. Thus, the deviation-constrained mechanism is simply a mechanism, where at every strategy profile m, the mechanism assigns for each agent a subset of messages where she is allowed to deviate. Although the result of Koray and Yildiz (2018) is obtained under the assumption that agents' choice behavior is rational, it extends to our set-up.[10]

3. Examples

In this section, we present three examples of SCRs that are behaviorally implementable by an individual-based rights structure, but cannot be studied using the sufficiency results of de Clippel (2014) or Korpela (2012).[11]

Example 1. *Status quo bias.* Let $N = \{1, 2\}$ and $X = \{a, b, d\}$. Outcomes a and b represent new policy options, while d is the status quo. The agents can be susceptible to the status quo bias, i.e., they choose the status quo alternative from the grand set X, even if they might choose differently from a smaller set that includes the status quo.[12] We assume that the agents' choice behavior aligns with their preferences when making pairwise comparisons. For simplicity, we assume that agents have identical choice behavior. This is only to illustrate why choice unanimity becomes problematic under non-rational choice behavior. The state of the world $\theta \in \Theta$ defines agents' preferences and determines who is biased. The SCR is such that $F^1(\theta) = \{x\}$ if, and only if, x is the most preferred outcome for the agents.[13] The individual-based rights structure is such that $S = X$, the outcome function is the identity map, and $\gamma(a, d) = \gamma(d, b) = \gamma(b, d) = \{1\}$, $\gamma(a, b) = \gamma(b, a) = \gamma(d, a) = \{2\}$. Fig. 1 illustrates it.

To see that this rights structure implements F^1, consider first $a \in X$. The opportunity sets of agent 1 and 2 are $\{a, d\}$ and $\{a, b\}$, respectively. Since the agents make pairwise comparisons, they are not influenced by the status quo bias. If a is their most preferred outcome, they choose it from their respective opportunity sets. If not, they make different choices from their opportunity sets, so a is not a behavioral equilibrium. Similar logic holds for all other outcomes in X. Notice that if both agents were biased and, e.g., a was their most preferred outcome, they would choose d from X. Then, choice unanimity would require that d is the socially optimal outcome even though a was the most preferred outcome for both agents.

Example 2. *Choice overload.* Let $N = \{1, 2\}$ and $X = \{a, b, c\}$. The preferences of agents are represented in Table 1. Let us assume that agents' choices can be influenced by the choice overload bias.[14]

Then, let us suppose that agents choose the outcome $\{c\}$ when they choose from X. We let the SCR be such that $F^2(\theta) = \{b\}$ and $F^2(\theta') = \{a\}$. We define Γ by setting $S = X$, defining the outcome function h as the identity map, and defining the individual-based rights structure as follows: $\gamma(b, c) = \gamma(c, b) = \{1\}$ and $\gamma(a, c) = \gamma(c, a) = \{2\}$. Fig. 2 illustrates it.

Since agents' opportunity sets are always different from X, they choose according to their preferences. To see that Γ implements F^2 in behavioral equilibrium, let us first consider outcome b. The opportunity sets of agents 1 and 2 are $\{b, c\}$ and $\{b\}$, respectively. If the state is θ, agents choose b from their opportunity sets. If, instead, the state is θ', agent 1 chooses c, and so b is not a behavioral equilibrium. Now consider a. Opportunity set for agent 1 is $\{a\}$ and for agent 2 is $\{a, c\}$. Agents both choose a at θ'. However, at θ, agent 2 chooses c. Observe that c is never a behavioral equilibrium. Thus, Γ implements F^2 in behavioral equilibrium. Finally, since agents would choose c from X, choice unanimity would require c to be socially optimal. Thus, F^2 does not satisfy choice unanimity, and therefore we cannot use any results that rely on this condition.

[10] We omit the proof that would be a word-by-word translation of the proof of Koray and Yildiz (2018). The reader can think of the deviation-constrained mechanism as a standard mechanism that operates under a system of law that prohibits agents from taking certain actions.

[11] For notational clarity, we denote the SCR in example $i = 1, 2, 3$ by F^i.

[12] The literature on status quo bias suggests that the tendency to choose the status quo option increases as the size of the choice set increases. See, for example, Samuelson and Zeckhauser (1988), Tversky and Shafir (1992), and Iyengar and Lepper (2000)

[13] Notice that even if agents have identical preferences, the planner cannot simply ask their opinion — this would amount to choosing from X, which leads to the wrong outcome when they are biased.

[14] See, for instance, Toffler (1970).

Table 1
Preferences of agents 1 and 2.

θ		θ'	
1	2	1	2
a	b	a	b
b	c	c	a
c	a	b	c

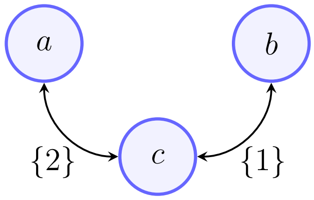

Fig. 2. Choice overload.

Table 2
Preferences of agents 1, 2, and 3.

θ			θ'		
1	2	3	1	2	3
a	c	c	a	c	b
c	b	a	b	a	a
b	a	b	c	b	c

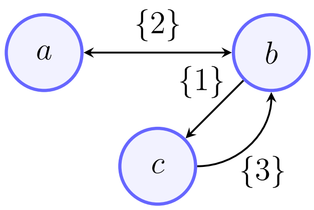

Fig. 3. Compromise effect.

Example 3. *Compromise effect*. Let $N = \{1, 2, 3\}$ and $X = \{a, b, c\}$, where outcomes describe new environmental policies. The preferences of agents are given in Table 2. The planner knows that the agents can be sensitive to the compromise effect. The compromise effect is a behavioral bias that can manifest, when the alternatives in a choice set can be evaluated by at least two measurable characteristics (e.g., price and quality), and some alternative in the choice set represents a compromise between these two characteristics.[15] In our example, suppose the agents have to decide how to clean up a local river. Suppose policy a represents an expensive but effective alternative, while policy c is an inexpensive but less effective alternative. Lastly, policy b represents a middle ground option between a and c. In this case, if an agent is biased, she chooses b from X. The planner only knows that some agents may suffer from this bias. Let F^3 be such that $F^3(\theta) = \{c\}$ and $F^3(\theta') = \{a\}$.

The implementing individual-based rights structure is such that $S = X$, the outcome function is the identity map, and $\gamma(a, b) = \gamma(b, a) = \{2\}$, $\gamma(b, c) = \{1\}$, $\gamma(c, b) = \{3\}$. Fig. 3 illustrates it.

Let us verify that Γ implements F^3 in behavioral equilibrium. At $\{a\}$, the opportunity set for agents 1 and 3 is $\{a\}$, while for agent 2 it is $\{a, b\}$. At θ', for all agents, $\{a\}$ is the best option in their opportunity sets, so it is a behavioral equilibrium. However, at θ, agent 2 can find a profitable deviation. Similarly, at $\{c\}$, the opportunity set for agents 1 and 2 is $\{c\}$, while for agent 3 it is $\{b, c\}$. At θ', $\{c\}$ is not a behavioral equilibrium because 3 has incentives to move to $\{b\}$. However, it can easily be checked that $\{c\}$ is a behavioral equilibrium at θ. Finally, since b is never a behavioral equilibrium, this rights structure implements F^3. Finally, since agents would choose b from X, choice unanimity would require b to be socially optimal. Thus, F^3 does not satisfy choice unanimity, and again we cannot use any results that rely on this condition.

[15] This is an informal definition, for more on compromise effect, see, e.g., Simonson (1989) and Galeotti et al. (2022).

1 \ 2	L	R
U	a	b
D	a	c

Fig. 4. Implementing mechanism for Example 2.

A (3)

1 \ 2	L	R
U	c	a
D	b	c

B (3)

1 \ 2	L	R
U	b	a
D	a	c

Fig. 5. Implementing mechanism for Example 3.

Before presenting our main result, let us note that Examples 2 and 3 discussed above can also be implemented in behavioral equilibrium by designing a mechanism. However, the SCR of Example 1 cannot be behaviorally implemented via any mechanism.[16]

The mechanism in Fig. 4 implements the SCR of Example 2, whereas the mechanism in Fig. 5 works for Example 3. In the mechanism in Fig. 4, agent 1 is the row-agent, so s/he chooses between row U and row D, whereas agent 2 is the column-agent and chooses between columns L and R. The outcome in each box is the outcome of the strategy profile to which the box corresponds. By using the preferences displayed in Table 1, it can be checked that (U, R) is the unique behavioral equilibrium strategy profile at θ, whereas (D, L) is the unique behavioral equilibrium strategy profile at θ'. Thus, F^2 of Example 2 is behaviorally implementable by a mechanism.

In the mechanism in Fig. 5, agent 1 is the row-agent, agent 2 is the column-agent, and agent 3 is the matrix-agent because s\he chooses between matrices A and B. The outcome in each box is the outcome of the strategy profile to which the box corresponds. By using the preferences displayed in Table 2, one can check that (U, L, A) is the unique behavioral equilibrium strategy profile at θ, whereas (U, R, B) is the unique behavioral equilibrium strategy profile at θ'. Thus, F^3 of Example 3 is also behaviorally implementable by a mechanism.

4. Main result: A full characterization

In this section, we provide a necessary and sufficient condition for behavioral implementation by individual-based rights structures. This condition is based on the following key necessary condition for behavioral implementation given in de Clippel (2014).

Definition 5. A collection of sets $\mathcal{O}^* = \{\mathcal{O}_i(a,\theta) \subseteq X \mid \theta \in \Theta, a \in F(\theta), i \in N\}$ satisfies **consistency** with respect to $F : \Theta \twoheadrightarrow X$, if the following requirements are satisfied for all $\theta, \theta' \in \Theta$ and all $a \in F(\theta)$:

(i) $a \in C_i(\mathcal{O}_i(a,\theta), \theta)$ for all $i \in N$.
(ii) If $a \in C_i(\mathcal{O}_i(a,\theta), \theta')$ for all $i \in N$, then $a \in F(\theta')$.

Maskin (1999); circulated since 1977) showed that in a rational domain any Nash implementable SCR satisfies a remarkably strong invariance condition, now widely referred to as Maskin monotonicity. It requires that if x is F-optimal at state θ, preferences change from θ to θ', while x does not fall in any agent's preference ordering relative to other alternatives, then x must remain F-optimal at state θ'. Consistency is an extension of Maskin monotonicity beyond the rational domain.[17] Suppose that F is behaviorally implementable by an individual-based rights structure. If a is a behavioral equilibrium outcome at θ, the equilibrium state of the society $s \in S$ supporting it defines an opportunity set for each agent i, denoted by $O_i(s)$, which represents the set of outcomes that agent i can induce by unilaterally deviating from s. From the definition of the behavioral equilibrium, each agent i must choose a from $O_i(s)$ at θ. Moreover, if there is an alternative state of the world θ' such that every agent i chooses a from $O_i(s)$ at θ', then s forms a behavioral equilibrium also at θ'. Hence, a is still an F-optimal outcome at θ' if F is behaviorally implementable by an

[16] The proof is relegated to Appendix. Notice that any SCR which is behaviorally implementable by a mechanism, is also behaviorally implementable by an individual-based rights structure. The implementing rights structure can be constructed from the implementing mechanism; select strategy-profiles as states of the society, let the outcome in any state of the society be the outcome that would prevail if all agent select the corresponding strategy in the mechanism, and then allow each agent to move from one state of the society to another (or from one strategy-profile to another) if they can induce the new strategy profile by changing their own strategy. This shows that mechanisms are in fact special kind of rights structures.

[17] See Lemma 3 in Barlo and Dalkıran (2022) for a proof.

individual-based rights structure. This explains why the existence of a collection of sets satisfying consistency with respect to an SCR is a necessary condition for behavioral implementation; for any $\theta \in \Theta$, $a \in F(\theta)$, and $i \in N$, we can set $\mathcal{O}_i(a,\theta) = O_i(s^*)$, where s^* is any behavioral equilibrium at θ such that $h(s^*) = a$.

Our condition is an extension of consistency and it can be stated as follows.

Definition 6. A collection of sets $\mathcal{O} = \mathcal{O}^* \cup \{\mathcal{O}_i(b) \subseteq B \mid b \in B \backslash F(\Theta), i \in N\}$ satisfies **extended consistency** with respect to $F : \Theta \twoheadrightarrow X$, if the following requirements are satisfied:

(i) The collection $\mathcal{O}^*$ satisfies consistency with respect to $F : \Theta \twoheadrightarrow X$.
(ii) For all $\theta \in \Theta$ and all $b \in B \backslash F(\Theta)$, $b \notin C_i(\mathcal{O}_i(b), \theta)$ for some $i \in N$.

Extended consistency with an SCR F requires not only the existence of a collection of sets satisfying consistency with respect to F, but also the existence of sets of type $\mathcal{O}_i(b)$ where b is never an equilibrium outcome. Part (i) of extended consistency is satisfied by an F that is behaviorally implementable by individual-based rights structure for exactly the same reason as before, since the sets $O_i(s)$ where agents can deviate from a state of the society s are subsets of B by definition. To check part (ii), let us consider any state $\theta \in \Theta$, and any outcome $b \in B \setminus F(\Theta)$. Since b is in the range of the rights structure, there must exist $s \in S$ such that $h(s) = b$. Set $\mathcal{O}_i(b) = O_i(s)$. Part (ii) of our condition requires that b cannot be the outcome chosen from $\mathcal{O}_i(b)$ at θ. The reason why this must hold is that if it were chosen from $O_i(s)$ at θ by every agent i, then s would be a behavioral equilibrium at θ, and so $b = h(s) \in F(\theta)$, yielding a contradiction.

The argument from the previous paragraph illustrates that the existence of a collection of sets satisfying extended consistency with respect to an SCR is necessary for behavioral implementation. Next, we show that in addition to being necessary, extended consistency with respect to an SCR is also a sufficient condition for behavioral implementation by an individual-based rights structure, which implies our main result.

Theorem 1. *(Full characterization). F is behaviorally implementable by an individual-based rights structure if, and only if, there exists a collection of sets satisfying extended consistency with respect to F.*

Proof. "*Only if*": Omitted as straightforward extension of the arguments already presented informally.

"*If*": Suppose there exists a collection of sets satisfying extended consistency with respect to F. We construct an individual-based rights structure Γ that implements F. To this end, let $\psi : X \to \Theta$ be any function such that $\psi^{-1}(\theta) \in F(\theta)$ for all $\theta \in \Theta$. This function maps every outcome $a \in F(\Theta)$ to exactly one state of the world $\psi(a) \in \Theta$ such that a is socially optimal at it. Let the set of states of the society be

$$S \equiv \{(a,\theta) \in X \times \Theta | a \in F(\theta)\} \cup B,$$

and define the outcome function $h : S \to X$ by setting $h((a,\theta)) = a$ for all $(a,\theta) \in S \setminus B$ and $h(b) = b$ for all $b \in B$. Finally, define the code of rights $\gamma : S \times S \to 2^N$ by the following rules. For all $i \in N$:

(1) $\{i\} \in \gamma((a,\theta), s)$ iff $h(s) \in \mathcal{O}_i(a,\theta)$;
(2) for all $a \in F(\Theta)$, $\{i\} \in \gamma(a,s)$ iff $h(s) \in \mathcal{O}_i(a,\psi(a))$;
(3) for all $b \in B \setminus F(\Theta)$, $\{i\} \in \gamma(b,s)$ iff $h(s) \in \mathcal{O}_i(b)$;
(4) $\gamma(s,s') = \emptyset$ for all other states s and s'.

Let us show that this individual-based rights structure implements F in behavioral equilibrium.

Let us first show that $F(\theta) \subseteq h(BE(\Gamma,\theta))$ holds for all $\theta \in \Theta$. Fix any $\theta \in \Theta$ and any $a \in F(\theta)$. Let us show that (a,θ) is a behavioral equilibrium at θ. By construction, a unilateral deviation from (a,θ) can happen only via rule 1. Then, for each agent $i \in N$, agent i's opportunity set $O_i((a,\theta))$ is equal to $\mathcal{O}_i(a,\theta)$. Since part (i) of de Clippel's (2014) consistency implies that $a \in C_i(\mathcal{O}_i(a,\theta),\theta)$ for all $i \in N$, it follows that $a \in C_i(O_i((a,\theta)),\theta)$ for all $i \in N$. Thus, (a,θ) is a behavioral equilibrium at θ.

For the converse, fix any $\theta \in \Theta$, and any $s^* \in BE(\Gamma,\theta)$. We proceed according to the following cases:

Case 1: $s^* = (b,\theta')$ for some $\theta' \in \Theta$ and some $b \in F(\theta')$.
Case 2: $s^* \in B \cap F(\Theta)$.
Case 3: $s^* \in B \setminus F(\Theta)$.

Case 1: $s^* = (b,\theta')$ for some $\theta' \in \Theta$ and some $b \in F(\theta')$

Then, $h(s^*) = b$. According to γ, a unilateral deviation from s^* can happen only via rule 1. Fix any $i \in N$ and any $c \in \mathcal{O}_i(b,\theta')$. Then, agent i has the power to move from s^* to any s such that $h(s) = c$. Since this holds for all $c \in \mathcal{O}_i(b,\theta')$, we have that $\mathcal{O}_i(b,\theta') = O_i(s^*)$. Since the choice of agent i was arbitrary, it follows that $\mathcal{O}_i(b,\theta') = O_i(s^*)$ for all $i \in N$. Since s^* is a behavioral equilibrium at θ, it follows that $b \in C_i(O_i(s^*),\theta)$ for all $i \in N$. Since $\mathcal{O}_i(b,\theta') = O_i(s^*)$ for all $i \in N$, it follows that $b \in C_i(\mathcal{O}_i(b,\theta'),\theta)$ for all $i \in N$. Part (ii) of de Clippel's (2014) consistency implies that $b \in F(\theta)$.

Case 2: $s^* \in B \cap F(\Theta)$

Then, according to γ, a unilateral deviation from s^* can happen only via rule 2. Fix any $i \in N$. Fix any $b \in \mathcal{O}_i(s^*,\psi(s^*))$. Then, agent i has the power to move from s^* to any s such that $h(s) = b$. Since this holds for all $b \in \mathcal{O}_i(s^*,\psi(s^*))$, we have that $\mathcal{O}_i(s^*,\psi(s^*)) = O_i(s^*)$. Since the choice of agent i was arbitrary, it follows that $\mathcal{O}_i(s^*,\psi(s^*)) = O_i(s^*)$ for all $i \in N$. Since s^* is a behavioral equilibrium at θ, it follows that $s^* \in C_i(O_i(s^*),\theta)$ for all $i \in N$. Since $\mathcal{O}_i(s^*,\psi(s^*)) = O_i(s^*)$ for all $i \in N$, it follows that $s^* \in C_i(\mathcal{O}_i(s^*,\psi(s^*)),\theta)$ for all $i \in N$. Again, part (ii) of de Clippel's (2014) consistency implies that $s^* \in F(\theta)$.

Case 3: $s^* \in B \backslash F(\Theta)$

Part (ii) of extended consistency requires that $s^* \notin C_i(\mathcal{O}_i(s^*), \theta)$ holds for some $i \in N$. It then follows directly, that $s^* \in B \setminus F(\Theta)$ cannot be an equilibrium.

We conclude that $h(BE(\Gamma, \theta)) \subseteq F(\theta)$ holds for all $\theta \in \Theta$. Thus, Γ implements F in behavioral equilibrium. □

de Clippel (2014) shows that consistency can be transformed into a sufficient condition by strengthening consistency and assuming choice unanimity (de Clippel, 2014, Proposition 2B). Our characterization avoids this by replacing choice unanimity with part (ii) of extended consistency. However, de Clippel (2014) has also another sufficient condition that does not use choice unanimity.

Proposition 1 (de Clippel, 2014, Proposition 2B’)**.**

If there exists a collection of opportunity sets satisfying consistency with respect to an SCR F, and there exist sets $(X_i)_{i \in N}$ *such that* $x \in F(\theta)$ *for each* x, θ *with* $\left|\{i \in N \mid x \in C_i(X_i \cup \{x\}, \theta)\}\right| \geq n-2$*, then* F *is behaviorally implementable by a mechanism.*

It is easy to see that this condition is stronger than extended consistency. If this condition holds, then we can set $B = X$ and select $\mathcal{O}_i(b) = X_i \cup \{b\}$ for all $b \in X \setminus F(\Theta)$ and all $i \in N$ to satisfy extended consistency. Moreover, this condition is far too strong for small n, while our full characterization makes no assumptions about n. In the next section we look at an example on limited willpower that does not always satisfy 2B’; thus important cases are left out.

Before closing this section, let us note that extended consistency coincides with de Clippel’s (2014) consistency when the range of the SCR F is equal to the set of outcomes. This means that there are no outcomes that are merely used to generate incentives ($X \setminus F(\Theta) = \emptyset$).

Corollary 1. *Let* $F(\Theta) = X$*.* F *is behaviorally implementable by an individual-based rights structure if, and only if, there exists a collection of opportunity sets satisfying consistency with respect to* F*.*

Proof. $F(\Theta) = X$ implies that part (ii) of extended consistency is automatically satisfied. □

5. Limited willpower

This section reconsiders the example of limited willpower presented by de Clippel (2014) to illustrate how our results can be applied.

Consider a group of n agents who have a common long-term goal. These agents face tempting outcomes that they may select over the common long-term goal. Each agent can exercise limited willpower, which is characterized by the number of tempting outcomes an agent can ignore over the long-term goal.[18] Agents’ long-term preferences are represented by an ordering $\succ_L$ on X and short-term cravings by $\succ_S$ on X. Let an integer k_i determine agent i’s willpower. Agent i’s choice out of any set $A \subseteq X$ is the most preferred outcome with respect to $\succ_L$ among those outcomes that are dominated by at most k_i outcomes according to $\succ_S$. The following illustrative example was given by de Clippel (2014): If the willpower of i is $k_i = 1$ and long-term goal and short-term cravings are such that, salad $\succ_L$ pizza $\succ_L$ burger and burger $\succ_S$ pizza $\succ_S$ salad, agent i will choose pizza from the set {salad, burger, pizza}.

Suppose that a state of the world $\theta \in \Theta$ specifies the common long-term goal as well as agents’ possibly distinct short-term cravings. We define the SCR F as a function that chooses the common long-term goal from the set of all outcomes at any state $\theta \in \Theta$. When agents are faced with the set of all possible outcomes that is large enough, they cannot exercise willpower, which leads them to choose the tempting short-term outcome. We show that the SCR described above satisfies extended consistency if $\sum_{i \in N} k_i \geq |X|$.[19]

Theorem 2. *Assume that* $\sum_{i \in N} k_i \geq |X|$*. Then,* F *is behaviorally implementable by an individual-based rights structure.*

Proof. To prove this claim we must identify the sets in the definition of extended consistency. For any $\theta \in \Theta$, $a \in F(\theta)$, and $i \in N$, select the sets $\mathcal{O}_i(a, \theta)$ in such a way that $|\mathcal{O}_i(a, \theta)| = k_i + 1$, $a \in \mathcal{O}_i(a, \theta)$, and $X = \bigcup_{i \in N} \mathcal{O}_i(a, \theta)$. This construction is feasible because of the assumption that $\sum_{i \in N} k_i \geq |X|$. The sets $\mathcal{O}_i(b)$, where $b \in X \setminus F(\Theta)$, are chosen similarly; $|\mathcal{O}_i(b)| = k_i + 1$, $b \in \mathcal{O}_i(b)$, and $X = \bigcup_{i \in N} \mathcal{O}_i(b)$. Finally, let us set $B = X$. These sets form the collection $\mathcal{O}$.

We are left to check that $\mathcal{O}$ satisfies extended consistency with respect to F. Since every set $\mathcal{O}_i(a, \theta)$ contains $k_i + 1$ outcomes, any outcome in these sets is dominated by at most k_i outcomes with respect to $\succ_S$. It follows that agent i chooses the best outcome with respect to $\succ_L$ from the set $\mathcal{O}_i(a, \theta)$. Next, consider any two states of the world $\theta, \theta' \in \Theta$. Fix any $a \in F(\theta)$ such that $a \in C_i(\mathcal{O}_i(a, \theta), \theta')$ for all $i \in N$. We must show that $F(\theta') = \{a\}$. Since every agent i now selects a as the most preferred outcome with respect to $\succ_L$ from the set $\mathcal{O}_i(a, \theta)$ at θ', it follows that $F(\theta') = \{a\}$ must indeed be the case. Thus, part (i) of extended consistency is satisfied. To verify also part (ii), select any $b \in X \setminus F(\Theta)$. For the sake of contradiction, suppose that all agents select b from $\mathcal{O}_i(b)$ at some state $\theta \in \Theta$. Since all agents now select the best outcome according to $\succ_L$ (for the same reason as before), b must then be the common long-term goal; a contradiction. Thus also part (ii) holds. □

6. Convergence

Convergence from non-equilibrium to equilibrium states of the society is significant in the design of rights structures as no initial state of the society has been specified. In this section, we require the implementing individual-based rights structure to be endowed

[18] For more on willpower, see Çgedik (2022).

[19] de Clippel (2014) also makes this assumption. Additionally, notice that the SCR does not satisfy condition 2B’, when $|N| = |X| = 3$ and $k_i = 1$.

with a well-defined convergence property. Indeed, we identify conditions which guarantee that an SCR behaviorally implementable by an individual-based rights structure is behaviorally implementable by one where each non-equilibrium state of the society is connected to an equilibrium state of the society via a finite path of movements. Formally, we require that the implementing rights structure satisfies the following *path existence* condition.

Definition 7. (*Path Existence*). An individual-based rights structure $\Gamma = (S, h, \gamma)$ satisfies **path existence** if, at each state of the world $\theta \in \Theta$, and any non-equilibrium state $s_0 \notin BE(\Gamma, \theta)$, there exists $a \in F(\theta)$ and a **path** (a finite sequence of states) $s_0, s_1, \dots, s_m$ with $s_m \in BE(\Gamma, \theta)$ and $h(s_m) = a$ such that for each $k \in \{1, 2, \dots, m\}$, $\{i\} \in \gamma(s_{k-1}, s_k)$, $h(s_k) \in C_i(O_i(s_{k-1}), \theta)$, and $h(s_{k-1}) \notin C_i(O_i(s_{k-1}), \theta)$ for some $i \in N$.

Koray and Yildiz (2018) propose a similar convergence condition on a rational domain. The difference is that they require every equilibrium state of the society to be obtainable from any non-equilibrium state of the society, while we require that only some equilibrium states of the society must be obtainable.[20]

We want to emphasize that for some particular instances the canonical rights structure that was used in the proof of Theorem 1 can satisfy path existence. The limited willpower example of the previous section is a case in point. In fact, for this example, there always exists a one-step path to equilibrium from any non-equilibrium state of the society. To see this, select the collection of sets $\mathcal{O}$ that is used in the proof of Theorem 1 as the collection of sets that was used in the proof of Theorem 2. Then, for any non-equilibrium state of the society s in the rights structure, there exists some state of the world, say $\theta' \in \Theta$, such that by unilaterally deviating each agent i can obtain exactly the outcome in the set $\mathcal{O}_i(F(\theta'), \theta')$. The joint long-term goal of the agents must be in at least one of these sets by definition of the collection $\mathcal{O}$. This agent wants to select the long-term goal from his or her opportunity set because it includes exactly $k_i + 1$ outcomes, and therefore is willing to move directly back to equilibrium. Often, however, we need a specifically tailored rights structure to guarantee convergence.

Let us first consider an economic environment where the best outcome of one particular agent is the worst outcome of all other agents. In a resource allocation problem, for instance, the best outcome of an agent is to get all available resources, which is the worst outcome of everybody else. The following definition is due to Hayashi et al. (2023).

Definition 8. (*Strongly competitive environment*). The implementation environment is **strongly competitive** if there exists a profile of outcomes $(x[i])_{i \in N}$, with $x[i] \in X$ for each $i = 1, \dots, n$ and $x[i] \neq x[j]$ for all $i \neq j$, such that the following two properties hold for all $i, j \in N$ with $i \neq j$, all $\theta \in \Theta$, and all $A \subseteq X$:

(i) If $x[i] \in A$, then $\{x[i]\} = C_i(A, \theta)$.
(ii) If $x[i] \in A$ and $x[i] \in C_j(A, \theta)$, then $A = C_j(A, \theta)$.

Consider a resource allocation problem where X is the set of all resource allocations and w is the aggregate endowment. Thus, for any $x \in X$, x_i represent resources given to agent i. Let $x[i]$ be such that $x[i]_i = w$ and $x[i]_j = 0$ for all $j \neq i$. Part (i) requires that if agent i is called to make a choice from a set A containing $x[i]$, then $x[i]$ is uniquely chosen from it. This holds because in all other outcomes some resources are given to other agents. Part (ii) requires that if agent j deems $x[i]$ choosable from A, then s/he cannot reject any other outcome from A. This holds because agent j will choose $x[i]$ only if his or her resources are 0 in all outcomes of A.

On top of this assumption on the implementation environment we need the following weak assumption to prove our convergence result.

Definition 9. An SCR F is **non-dictatorial** if for all $\theta \in \Theta$, all $z \in F(\theta)$, and all $i \in N$, we have $\{z\} = C_j(\{z, x[i]\}, \theta)$ for some $j \neq i$.

This condition says that there cannot exist an agent who gets all resources at some state of the world. If agent $j \neq i$ selects $x[i]$ when also z is available, then s\he gets nothing at allocation z either. Thus, if all agents other than i are willing to select $x[i]$ instead of z, then it must be that agent i gets all resources at z. Usually SCRs are called non-dictatorial if there exists at least one state of the world like this, while we require that the SCR is non-dictatorial at every state of the world.

The following result says that in a strongly competitive environment with two or more agents, any SCR F behaviorally implementable by an individual-based rights structure that is non-dictatorial can be behaviorally implemented by an individual-based rights structure satisfying the condition of path existence.

Theorem 3. *Assume that the environment is strongly competitive. Any non-dictatorial SCR behaviorally implementable by an individual-based rights structure is behaviorally implementable by an individual-based rights structure satisfying path existence.*

Proof. See Appendix. □

To stress that convergence is not a severe constraint even more, we give another condition that guarantees convergence in social choice environments such as voting. The first assumption says that for every outcome x in the range of F and any outcome z outside it, at least one agent selects x and rejects z from the set $\{x, z\}$.

Definition 10. The environment is **pairwise non-problematic** for F, if for all $\theta \in \Theta$, all $x \in F(\Theta)$, and all $z \in X \setminus F(\Theta)$, we have $\{x\} = C_i(\{x, z\}, \theta)$ for at least one agent $i \in N$.

[20] Koray and Yildiz (2018) show that a property called winner monotonicity guarantees convergence in rational domains. We cannot use a generalization of this because our notion of convergence is weaker.

It is worth noting that all applications discussed above satisfy this definition. Under rational choice behavior this condition requires that F must select Pareto-optimal outcomes. Thus, when agents make unbiased choices in small sets, a presumption that often holds for common behavioral biases, this condition is often satisfied.

Our second condition is a generalization of a well-known condition in social choice theory, namely the *tops-only property* (Chatterji and , 2011). In rational domains tops-only property requires that an SCR can select only outcomes that are top-ranked by at least one agent. We replace the "top-ranked" by requiring that an SCR can select only outcome that at least one agent would select if he would be allowed to select among all possible outcomes.

Definition 11. An SCR F satisfies the **behavioral tops-only property** when for all $\theta \in \Theta$, and all $x \in F(\theta)$, if $C_i(X,\theta) \nsubseteq X \setminus F(\Theta)$ for all $i \in N$, then $\{x\} = C_j(X,\theta)$ for some $j \in N$.

The above definition is satisfied by Examples 2 and 3 discussed above.

Theorem 4. *Assume that the environment is pairwise non-problematic for an SCR F that satisfies the behavioral tops-only property. If F is behaviorally implementable by an individual-based rights structure, then it is behaviorally implementable by an individual-based rights structure satisfying path existence.*

Proof. See Appendix. □

7. Conclusion

We have derived a complete characterization of SCRs that are behaviorally implementable by individual-based rights structures. A rights structure captures the idea of what individuals can or cannot do in a society and it is a flexible tool for designing institutions such as constitutions, legal codes, and rules of conduct. Our necessary and sufficient condition, called extended consistency, is a strengthening of de Clippel's (2014) consistency. The conditions coincide if every outcome is socially optimal at some state of the world.

The best way to see how our result complements the existing literature is to compare it simultaneously with Koray and Yildiz (2018) and de Clippel (2014). In their main result (Koray and Yildiz, 2018) provide a full characterization of SCRs that are implementable by individual-based rights structures when choice behavior is rational; this is given by two well-known conditions — Maskin monotonicity together with unanimity. de Clippel (2014), on the other hand, shows that the natural counterpart of Maskin monotonicity when choice behavior can be non-rational is consistency. Since the definition of extended consistency has two parts, de Clippel's consistency (item (i)), and another property (item (ii)), it immediately reveals the condition that must replace unanimity when choice behavior can be non-rational.

Data availability

No data was used for the research described in the article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Appendix

Proof of the non-existence of a mechanism for Example 1

Any mechanism implementing the SCR in Example 1 must satisfy the following conditions:

(1) $X \subseteq \bigcup_{i\in N} O_i(m_{-i})$ for all $m \in M$.
(2) Let $\theta \in \Theta$ be such that both agents are biased and $F^1(\theta) \in \{a,d\}$. For any $m^*(\theta) \in M$ such that $g(m^*(\theta)) \in \{a,b\}$, it must be that $|O_i(m^*_{-i}(\theta))| \leq 2$ holds for both $i \in N$.

Condition (1) simply states that agents must always be able to deviate to their most preferred outcome. Condition (2) states that whenever both agents are biased and $\{a\}$ or $\{b\}$ should be an equilibrium outcome, the opportunity set of either agent cannot contain more than 2 elements. Otherwise the status quo bias would lead to a deviation to $\{d\}$.

Intuition of the proof is that since agents can only make pairwise comparisons, the row-column structure of a mechanism guarantees the existence of unwanted equilibria.

We will now show that no implementing mechanism can be constructed. Fix any $\theta \in \Theta$ such that both agents are biased and $F^1(\theta) = \{a\}$. Condition (2) requires that in any implementing mechanism, where $\{a\}$ is an equilibrium outcome, the opportunity sets of agents 1 and 2 must be $\{a,d\}$ and $\{a,b\}$. Without loss of generality, suppose 1 is the column agent and 2 is the row agent. The equilibrium described above is defined by the intersection of a row containing outcomes $\{a,d\}$ and a column containing $\{a,b\}$. Then, from any $\{a\}$ and $\{d\}$ on this row, the row agent 2 must be able to deviate to $\{b\}$ for condition (1) to hold. This defines at least two columns that contain outcomes $\{a, b\}$ and $\{d, b\}$. Similarly, from any $\{b\}$ on any such aforementioned $\{a, b\}$ column, column agent 1 must be able to deviate to $\{d\}$. This implies that the mechanism has a column and a row that both contain outcomes $\{b, d\}$. The intersection of these two consists of an outcome, from which either agent can't deviate to $\{a\}$, violating condition (1), and creating an unwanted equilibria. Thus, any implementing mechanism cannot satisfy simultaneously conditions (1) and (2). □

Proof of Theorem 3

Suppose that a non-dictatorial SCR F is behaviorally implementable by an individual-based rights structure in a strongly competitive environment. We need to construct an individual-based rights structure Γ that satisfies path existence and implements F in behavioral equilibrium. Let the set of states of the society be

$$S \equiv \{(a,\theta) \in X \times \Theta | a \in F(\theta)\} \cup X \cup \{(x[i], \theta, a) \mid i \in N, \theta \in \Theta, a \in F(\theta)\}.$$

Define the outcome function $h : S \to X$ as $h((a,\theta)) = a$ for all $(a,\theta) \in S$, $h(b) = b$ for all $b \in X$, and $h((x[i], \theta, a)) = x[i]$ for all $(x[i], \theta, a) \in S$. Finally, define the code of rights $\gamma : S \times S \to 2^N$ by the following rules (here we assume that the agents are ordered from 1 to n and denote $n + 1 = 1$). For all $i, j \in N$:

(1) $\{i\} \in \gamma((a,\theta), b)$ iff $b \in \mathcal{O}_i(a,\theta)$;
(2) $\{i\} \in \gamma(b, (x[j], \theta, a))$ for any $b, (x[j], \theta, a) \in S$ and $i, j \in N$;
(3) $\{i\} \in \gamma((x[j], \theta, a), (\theta, a))$ with $i \neq j$;
(4) $\{i\} \in \gamma((x[j], \theta, a), (x[i], \theta', b))$ for all $(x[j], \theta, a), (x[i], \theta', b) \in S$ and $i \neq j$;
(5) $\gamma(s, s') = \emptyset$ for all other $s, s' \in S$.

Γ is well-defined because there exists a collection of sets satisfying extended consistency with respect to F. Let us verify that this individual-based rights structure implements F.

Let us first show that $F(\theta) \subseteq h(BE(\Gamma, \theta))$. Fix any $\theta \in \Theta$ and any $a \in F(\theta)$. Let us show that (a,θ) is a behavioral equilibrium at θ. By construction, a unilateral deviation from (a,θ) can happen only via rule 1. Then, for each agent $i \in N$, agent i's opportunity set $O_i(a,\theta)$ is equal to $\mathcal{O}_i(a,\theta)$. Since part (i) of de Clippel's (2014) consistency implies that $a \in C_i(\mathcal{O}_i(a,\theta), \theta)$ for all $i \in N$, it follows that $a \in C_i(O_i(a,\theta), \theta)$ for all $i \in N$. Thus, (a,θ) is a behavioral equilibrium at θ.

For the converse, suppose that θ is the true state of the world. Suppose that $s \in BE(\Gamma, \theta)$. First, let us notice that $s \notin X$ and that s cannot be of the type $(x[j], \theta, a)$ because our assumption that the environment is strongly competitive environment and each agent i can get $x[i]$ by unilaterally deviating via rule 2 if $s \in X$ or via rule 4 if $s = (x[j], \theta, a)$. Then, let us suppose that $s \in \{(a,\theta) \in X \times \Theta | a \in F(\theta)\}$. Let $s = (c, \theta')$.

Fix any $i \in N$. Fix any $b \in \mathcal{O}_i(c, \theta')$. Then, agent i has the power to move from c to b and to obtain it via rule 1. Since this holds for all $b \in \mathcal{O}_i(c, \theta')$, we have that $\mathcal{O}_i(c, \theta') = O_i(s)$. Since the choice of agent i was arbitrary, it follows that $\mathcal{O}_i(c, \theta') = O_i(s)$ for all $i \in N$. Since s is a behavioral equilibrium at θ, it follows that $c \in C_i(O_i(s), \theta)$ for all $i \in N$. Since $\mathcal{O}_i(c, \theta') = O_i(s)$ for all $i \in N$, it follows that $c \in C_i(\mathcal{O}_i(c, \theta'), \theta)$ for all $i \in N$. Part (ii) of de Clippel's (2014) consistency implies that $c \in F(\theta)$.

We conclude that also $BE(\Gamma, \theta) \subseteq F(\theta)$ holds. Thus, Γ behavioral implements F.

Finally, let us check that Γ satisfies path existence. Fix any $\theta \in \Theta$.

Fix any $s \notin BE(\Gamma, \theta)$. Then, $b \notin C_i(O_i(s), \theta)$ for some $i \in N$. There are three cases to be considered.

- Suppose that $s = (b, \theta')$. Since $b \notin C_i(O_i(b, \theta'), \theta)$, it holds that $c \in C_i(O_i(b, \theta'), \theta)$. This agent can use rule 1 to move from (b, θ') to c, and s/he has incentives to do so. Then, any agent j can obtain $x[j]$ by moving via rule 2 to $(x[j], a, \theta)$. Then, $x[j] \in O_j(c)$. S/he has incentives to move because the environment is strongly competitive, and so $\{x[j] = C_j(O_j(c), \theta)\}$ and $c \notin C_j(O_j(c), \theta)$. Since F is non-dictatorial, it follows that there exists $k \neq j$ such that $C_k(\{x[j], a\}, \theta) = \{a\}$. Thus, agent k can move away from $(x[j], a, \theta)$ via rule 3 by moving to (a, θ). Note that, by construction, $O_k((x[j], a, \theta)) = \{x[j], a\}$, and so $\{a\} = C_k(O_k((x[j], a, \theta)), \theta)$. Thus, there is a finite path from $(b, \theta') \notin BE(\Gamma, \theta)$ to the equilibrium $(a, \theta) \in BE(\Gamma, \theta)$.
- By the same arguments, it can be shown that there is a finite path to equilibrium $(a, \theta) \in BE(\Gamma, \theta)$ when $s = b \in B$—via rules 2 and 3.
- Suppose that $s = (x[j], b, \theta')$. Since the environment is strongly competitive, there exists $i \neq j$ such that $C_i(A \cup \{x[i]\}, \theta) = \{x[i]\}$ for all $A \in \mathcal{X}$. Then, $C_i(O_i((x[j], b, \theta')), \theta) = \{x[i]\}$. Agent i can move via rule 4 to $(x[i], a, \theta)$ and is incentivized to do so. Since F is non-dictatorial, it follows that there exists $j \neq i$ such that $C_j(\{x[i], a\}, \theta) = \{a\}$. By the arguments provided at the end of the case $s = (b, \theta')$, we see that there exists a finite path from $(x[j], b, \theta') \notin BE(\Gamma, \theta)$ to the equilibrium $(a, \theta) \in BE(\Gamma, \theta)$.

Since the choice of $s \notin BE(\Gamma, \theta)$ was arbitrary, we conclude that Γ satisfies path existence.

Proof of Theorem 4

Let the premises hold. Suppose that the environment is pairwise non-problematic for an SCR F that satisfies the behavioral tops-only property. We need to construct a rights structure Γ that satisfies path existence and implements F. Let the set of states of the society be

$$S \equiv \mathcal{S} \cup F(\Theta) \cup \{(y, \theta, a) \mid y \in X \setminus F(\Theta), (a, \theta) \in \mathcal{S}\}$$

where $\mathcal{S} \equiv \{(a,\theta) \in X \times \Theta | a \in F(\theta)\}$. Define the outcome function $h : S \to X$ as $h((a,\theta)) = a$ for all $(a,\theta) \in \mathcal{S}$, $h(b) = b$ for all $b \in F(\Theta)$, and $h(y, \theta, a) = y$ for all $(y, \theta, a) \in S$. Finally, define the code of rights $\gamma : S \times S \to 2^N$ by the following rules:

(1) $\{i\} \in \gamma((a,\theta), b)$ iff $b \in \mathcal{O}_i(a,\theta)$,
(2) $\{i\} \in \gamma((a,\theta),(y,\theta,a))$ iff $y \in \mathcal{O}_i(a,\theta)$,
(3) $\{i\} \in \gamma(b,s)$ for all $b, s \in S$ and $i \in N$,
(4) $\{i\} \in \gamma((y,\theta,a),(a,\theta))$ for all $(y,\theta,a),(a,\theta) \in S$ and $i \in N$,
(5) $\gamma(s,s') = \emptyset$ for all other $s, s' \in S$.

Γ is well-defined because there exists a collection of sets satisfying extended consistency with respect to F. Let us verify that this individual-based rights structure implements F.

Let us first show that $F(\theta) \subseteq h(BE(\Gamma,\theta))$. Fix any $\theta \in \Theta$ and any $a \in F(\theta)$. Let us show that (a,θ) is a behavioral equilibrium at θ. By construction, a unilateral deviation from (a,θ) can happen only via rules 1-2. Then, for each agent $i \in N$, agent i's opportunity set $O_i(a,\theta)$ is equal to $\mathcal{O}_i(a,\theta)$. Since part (i) of de Clippel's (2014) consistency implies that $a \in C_i(\mathcal{O}_i(a,\theta),\theta)$ for all $i \in N$, it follows that $a \in C_i(O_i(a,\theta),\theta)$ for all $i \in N$. Thus, (a,θ) is a behavioral equilibrium at θ.

For the converse, suppose that θ is the true state of the world. Take any $s \in BE(\Gamma,\theta)$. Let us show that $h(s) \in F(\theta)$.

First, let us note that $s \neq (y,\theta',a)$. To see it, suppose that $s = (y,\theta',a)$. Since the environment is pairwise non-problematic, there exists an agent i such that $\{a\} = C_i(\{a,y\},\theta)$. Since this agents has also the power to move from (y,θ',a) to (a,θ') via rule 4 and $O_i(s) = \{y,a\}$, agent i has found a profitable unilateral deviation, which is a contradiction.

Suppose that $s = a \in F(\Theta)$. Assume, to the contrary, that $a \notin F(\theta)$. Fix any $i \in N$. Fix any $b \in X$. By rule 3, s/he can move from a to b. It follows that $O_i(s) = X$. Since $s \in BE(\Gamma,\theta)$, it follows that $a \in C_i(X,\theta)$, and so $C_i(X,\theta) \nsubseteq X\backslash F(\Theta)$. Since the choice of i was arbitrary, we have that $C_i(X,\theta) \nsubseteq X\backslash F(\Theta)$ for all $i \in N$. Fix any $b \in F(\theta)$. Since $a \notin F(\theta)$, it holds that $a \neq b$. Since F satisfies behavioral top-only property and since $C_i(X,\theta) \nsubseteq X\backslash F(\Theta)$ for all $i \in N$, it follows that $C_i(X,\theta) = \{b\}$ for some $i \in N$. However, since we have already established that $a \in C_i(X,\theta)$ for all $i \in N$, this is a contradiction.

Finally, suppose that $s = (a,\theta')$. Fix any $i \in N$. Fix any $b \in \mathcal{O}_i(a,\theta')$. By construction, agent i can move to b via rule 1. Since the choice of b was arbitrary, it holds that $\mathcal{O}_i(a,\theta') \subseteq O_i(a,\theta')$. Since agent i can move away from (a,θ') either via rule 1 or via rule 2, and since s/he can obtain only outcome in $\mathcal{O}_i(a,\theta')$, it follows that $\mathcal{O}_i(a,\theta') = O_i(a,\theta')$. Since $s \in BE(\Gamma,\theta)$, it follows that $a \in C_i(O_i(a,\theta'),\theta)$. Since the choice of agent i's was arbitrary, we have that $a \in C_i(O_i(a,\theta'),\theta)$ for all $i \in N$. Part (ii) of de Clippel's (2014) consistency implies that $a \in F(\theta)$.

We conclude that $BE(\Gamma,\theta) = F(\theta)$. Since the choice of θ was arbitray, we have that Γ behavioral implements F.

Finally, let us check that Γ satisfies path existence. Fix any $s \notin BE(\Gamma,\theta)$. Then, $h(s) \notin C_i(O_i(s),\theta)$ for some $i \in N$. There are three cases to be considered.

Case 1: $s = (a_0,\bar{\theta})$.
Case 2: $s = (y,\bar{\theta},a_0)$.
Case 3: $s = a_0$.

Case 1: $s = (a_0,\bar{\theta})$.

Suppose that $s = (a_0,\bar{\theta})$. By construction, it holds that $O_i(a_0,\bar{\theta}) = \mathcal{O}_i(a_0,\bar{\theta})$. Since $a_0 \notin C_i(O_i(a_0,\bar{\theta}),\theta)$, it holds that $a_1 \in C_i(O_i(a_0,\bar{\theta}),\theta)$ for some $a_1 \in X$. By rule 1, this agent i has the power and incentives to move from $(a_0,\bar{\theta})$ to a_1. Let us proceed according to whether $a_1 \in F(\Theta)$.

Sub-case 1.1: $a_1 \in F(\Theta)$.

By rule 3, every agent i can obtain $O_i(a_1) = X$. If $a_1 \in BE(\Gamma,\theta)$, we have found a path to the equilibrium state a_1. Otherwise, suppose that $a_1 \notin BE(\Gamma,\theta) = F(\theta)$. Then, $a_1 \notin C_i(O_i(a_1),\theta)$ for some $i \in N$.

- Suppose that $C_i(X,\theta) \nsubseteq X\backslash F(\Theta)$ for all $i \in N$. Take any $a_2 \in F(\theta)$. Since F satisfies behavioral top-only property, it follows that $C_j(X,\theta) = \{a_2\}$ for some $j \in N$. This agent j can move from a_1 to (a_2,θ) via rule 3 and he is incentivized to do so. Thus, path existence is satisfied.
- Suppose that $C_j(X,\theta) \subseteq X\backslash F(\Theta)$ for some $j \in N$. Then, $a_1 \notin C_j(X,\theta)$. Take any $a_2 \in C_j(X,\theta)$. Then, agent j has incentives to induce rule 3 and attain a_2. Let us consider the state (a_2,θ,a_3). By rule 3, agent j has the power and incentives to move from a_1 to (a_2,θ,a_3). Since we are at (a_2,θ,a_3), it follows that every agent i can induce only rule 4, and so $O_i(a_2,\theta,a_3) = \{a_2,a_3\}$. Since $a_3 \in F(\theta)$ and $a_2 \in X\backslash F(\Theta)$, it follows from our supposition that the environment is pairwise non-problematic that there exists an agent i such that $\{a_3\} = C_i(\{a_2,a_3\},\theta)$, and so $a_2 \notin C_i(\{a_2,a_3\},\theta)$. By rule 4, this agent i has the power and incentives to move from (a_2,θ,a_3) to (a_3,θ).

Sub-case 1.1: $a_1 \notin F(\Theta)$.

By rule 2, agent i has the power to move from $(a_0,\bar{\theta})$ to (a_1,θ,a_2) and he also has incentives to do so because $a_0 \notin C_i(O_i(a_0,\bar{\theta}),\theta)$ and $a_1 \in C_i(O_i(a_0,\bar{\theta}),\theta)$. By rule 4, any agent i has the power to move from (a_1,θ,a_2) to (a_2,θ). Note that $O_i(a_1,\theta,a_2) = \{a_1,a_2\}$. Since the environment is pairwise non-problematic, there exists an agent i such that $\{a_2\} = C_i(\{a_1,a_2\},\theta)$, and so $a_1 \notin C_i(\{a_1,a_2\},\theta)$. Then, there is an agent i who has incentives and power to move from (a_1,θ,a_2) to (a_2,θ). Since $(a_2,\theta) \in BE(\Gamma,\theta)$, we found a path to an equilibrium state.

Case 2: $s = (y,\bar{\theta},a_0)$.

Suppose that $s = (y,\bar{\theta},a_0)$. Consider the state $(a_0,\bar{\theta})$. By rule 4, any agent i has the power to move from $(y,\bar{\theta},a_0)$ to $(a_0,\bar{\theta})$. Note that $O_i(y,\bar{\theta},a_0) = \{y,a_0\}$. Since the environment is pairwise non-problematic, there exists an agent i such that $\{a_0\} = C_i(\{a_0,y\},\theta)$. Then, there is an agent i who has incentives and power to move from $(y, \bar{\theta}, a_0)$ to $(a_0, \bar{\theta})$. We proceed according to whether $(a_0, \bar{\theta}) \in BE(\Gamma, \theta)$. If $(a_0, \bar{\theta}) \in BE(\Gamma, \theta)$, then we found a path to an equilibrium state. Suppose that $(a_0, \bar{\theta}) \notin BE(\Gamma, \theta)$. By using the same arguments used in *Case 1*, we have that there exists a path to an equilibrium state.

Case 3: $s = a_0$.

Suppose that $s = a_0$. By rule 3, every agent i can obtain $O_i(a_0) = X$. Since $a_0 \notin BE(\Gamma, \theta)$, it follows that $a_0 \notin C_i(X, \theta)$ and $a_1 \in C_i(X, \theta)$ for some $a_1 \in X$. Then, agent i has the power and incentives to move from a_o to a_1. Let us proceed according to whether $a_1 \in F(\Theta)$. By using the same arguments used in *Case 1*, we see that we converge to an equilibrium state.

Since the choice of $s \notin BE(\Gamma, \theta)$ was arbitrary, we conclude that Γ satisfies path existence.

References

Altun, O.A., Barlo, M., Dalkıran, N.A.D., 2023. Implementation with a sympathizer. Math. Soc. Sci. 121, 36–49.
Barlo, M., Dalkıran, N.A.D., 2023b. Robust behavioral implementation. Mimeo.
Barlo, M., Dalkıran, N.A.D., 2022. Computational implementation. Rev. Econ. Design 26 (4), 605–633.
Barlo, M., Dalkıran, N.A.D., 2023. Behavioral implementation under incomplete information. J. Econ. Theory 203, bibinfopages105738.
Barlo, M., Dalkıran, N.A.D., 2024. Ex-post behavioral implementation. https://ssrn.com/abstract=4932196.
Caspari, G., Khanna, M., 2025. Nonstandard choice in matching markets. Int. Econ. Rev. 66 (2), 757–786.
çgedik, G. K.L., 2022. Implementation with Individuals Limited Willpower (2022) Master's Thesis. Sabancı University.
Chambers, C.P., Yenmez, M.B., 2017. Choice and matching. Am. Econ. Rev: Microecon. 9, 126–147.
Chatterji, S., Sen, A., 2011. Tops-only domains. Econ. Theory 46 (2), 255–282.
de Clippel, G., 2014. Behavioral implementation. Am. Econ. Rev. 104, 2975–3002.
de Clippel, G., 2022. Departures from preference maximization, violations of the sure-thing principle, and relevant implications. Mimeo.
Galeotti, F., Montero, M., Poulsen, A., 2022. The attraction and compromise effects in bargaining: experimental evidence. Manage. Sci. 68, 2987–3007.
Hagiwara, M., 2025. Behavioral subgame perfect implementation. J. Econ. Behav. Org. 233, 106992.
Hayashi, T., Jain, R., Korpela, V., Lombardi, M., 2023. Behavioral strong implementation. Econ. Theory 76 (4), 1257–1287.
Hurwicz, L., 1986. On the Implementation of Social Choice Rules in Irrational Societies. Vol. I. Cambridge University Press, Cambridge.
Iyengar, S.S., Lepper, M.R., 2000. When choice is demotivating: can one desire too much of a good thing? J. Pers. Soc. Psychol. 79, 995.
Koray, S., Yildiz, K., 2018. Implementation via rights structures. J. Econ. Theory 176, 479–502.
Korpela, V., 2012. Implementation without rationality assumptions. Theory Decis. 72, 189–203.
Korpela, V., Lombardi, M., Savva, F., 2025. The difference between the weak core and the strong core from the design point of view. Econ. Theory 79 (4), 1255–1281.
Lombardi, M., Savva, F., Zivanas, N., 2023. Implementation in strong core by codes of rights. Soc. Choice Welfare 60 (3), 503–515.
Maskin, E., 1999. Nash equilibrium and welfare optimality. Rev. Econ. Stud. 66, 23–38.
Moore, J., Repullo, R., 1990. Nash implementation: a full characterization. Econometrica 58, 1083–1099.
Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. J. Risk Uncertain. 1, 7–59.
Saran, R., 2011. Menu-dependent preferences and revelation principle. J. Econ. Theory 146, 1712–1720.
Saran, R., 2016. Bounded depths of rationality and implementation with complete information. J. Econ. Theory 165, 517–564.
Savva, F., 2021a. Conditional rights and implementation Available at SSRN. https://ssrn.com/abstract=3466832.
Savva, F., 2021. Motives and implementation with rights structures. Econ Lett 204, 109913.
Sen, A., 1970. Collective Choice and Social Welfare. Holden Day, San Francisco.
Sen, A., 1977. Social choice theory: a re-examination. Econometrica 45, 53–89.
Sertel, R.M., 2001. Designing rights: invisible hand theorems, covering and membership. Mimeo. Bogazici University.
Simonson, I., 1989. Choice based on reasons: the case of attraction and compromise effects.. J. Consumer Res. 16, 158–174.
Toffler, A., 1970. Future Shock. New York. Random House.
Tversky, A., Shafir, E., 1992. Choice under conflict: the dynamics of deferred decision. Psychol. Sci. 3, 358–361.
Yildiz, K., 2025. Nash bargaining solution is implementable via two-stage rights structures. J. Econ. Behav. Org. 238, 107214.