BROWN
Orlando Bravo Center
for Economic Research

Signaling, Screening, and Core Stability

Bravo Working Paper # 2022-001

Yusuke Kamishiro ⓡ Rajiv Vohra ⓡ Roberto Serrano*

January 2022
Revised: August 2022

Abstract. This paper provides a noncooperative approach to core stability in an economy with incomplete information. We study the perfect Bayesian equilibria of an extensive form mechanism that extends the one used by Serrano and Vohra (1997) to implement the core of a complete-information economy. This leads to a version of the core that we refer to as the *sequential core*, which allows for information flows through proposals that can be viewed as signaling devices and/or screening contracts. Equilibrium refinements are then used to provide justifications for the coarse core and the fine core.

JEL Classification: C71, C72, D51.

Key Words: core, incomplete information, perfect Bayesian equilibrium, signaling, screening, Nash program.

*Author names are in random order. Kamishiro: Department of Economics, Kanto Gakuin University; ykamishi@kanto-gakuin.ac.jp, Vohra: Department of Economics, Brown University; rajiv_vohra@brown.edu, Serrano: Department of Economics, Brown University; roberto_serrano@brown.edu. Kamishiro's research is partially supported by KAKENHI (16H03121, 21K01407).

1 Introduction

In their seminal paper on Pareto efficiency with incomplete information, Holmström and Myerson (1983) –HM, henceforth– explain why the standard notion of efficiency does not immediately extend to environments with incomplete information. A feasible state-contingent allocation, or a decision rule, is Pareto efficient if one cannot find another feasible state-contingent allocation that makes everyone better off. HM mention at least three issues that must be addressed in extending this concept to environments with incomplete information: (1) Should feasibility be taken to include incentive compatibility constraints or not? (2) Is the term "better off" applied to ex ante expected utilities, interim expected utilities or ex post utilities? (3) Who is to check whether an improving allocation is available: the agents or an outside observer? Corresponding to how the first two questions are answered, HM propose six different definitions of Pareto efficiency. The third issue relates to whether or not the possibility of an improvement is required to be common knowledge, and we will have more to say about this presently.

The core can be seen as an extension of Pareto efficiency applied to all coalitions, not just the coalition of the whole. In the classical setting, it refers to the set of feasible allocations to which no coalition can find a feasible improvement. Since a core allocation must necessarily be Pareto efficient, all of the difficulties mentioned in HM will also need to be confronted in developing the notion of core stability with incomplete information, yet our results will yield surprising conclusions in this respect. To be sure, the different ways in which these questions have been addressed is one explanation for the plethora of different core concepts that appear in the literature following Wilson (1978). For reviews of this literature, we refer the reader to Forges, Minelli, and Vohra (2002) and Forges and Serrano (2013).

Our concern in this paper is the third issue mentioned in the opening paragraph. We examine core stability when the agents themselves are engaged in seeking potential coalitional improvements. To keep matters simple and following Wilson's (1978) seminal work, we resolve the first two questions by (a) assuming in most of the paper that incentive constraints are not relevant[1]–in the sense that private information eventually becomes publicly known and prohibitive penalties can be imposed on agents who lie about their private information–, and (b) confining attention to the interim stage, where agents have their private information, but do not necessarily know the information of others.[2]

To say that an outsider can check that one allocation rule Pareto dom-

[1]Subsection 6.4 discusses how to extend our results when incentive constraints matter.
[2]Indeed, this is the relevant stage for posing the third question.

inates another is equivalent to saying that it is common knowledge among the agents (or anyone who knows the model) that this is so. Wilson's (1978) definition of the coarse core formalizes this idea for coalitional improvements. A coalition has a coarse improvement over a status-quo if there is a feasible allocation (given the coalition's resources) such that it is common knowledge within the coalition that all its members are better off, in terms of interim expected utility. For the case in which agents within a coalition can share their private information in any arbitrary way, Wilson (1978) defines the fine core.

Unlike Wilson (1978), we make the amount of information transmission in the coalition formation process endogenous (rather than assuming there is *no* information leakage –as in the coarse core– or *any* kind of information leakage –as in the fine core). As argued in HM, if the agents themselves are engaged in collective decision-making at the interim stage, they may well depart from what would be considered an efficient allocation, from the point of view of an uninformed planner.[3] The same applies to coalitional decision-making.

Indeed, in a coalitional setting, Serrano and Vohra (2007) – SV (2007) in the sequel – argue that noncooperative equilibrium theory can be used to pin down the amount of private information agents transmit to each other in the process of making cooperative agreements. The equilibrium strategies in such a game implicitly determine the information transmission that takes place. They build on the observation that in a complete-information economy, a defining feature of a core allocation is that it will not be unanimously rejected within any coalition in favor of an alternative allocation. The same idea can be applied even in an incomplete-information environment. In their analysis, all agents in a coalition vote (simultaneously) to choose either the status-quo or another feasible alternative. A status-quo allocation is said to be *resilient* if there is no such voting game with *some* Bayesian equilibrium in which the status-quo is rejected. (Clearly, this is closely related to HM's durability; see the last section of SV (2007) for details).[4]

One issue with the approach of SV (2007) is that the use of a simultaneous-move voting game cannot avoid a coordination failure, in which a 'bad' status-quo survives because all agents expect others to reject an improvement. We tackle this shortcoming here by insisting on implementation through a sequential game. The complexities of signaling and screening are also absent in SV (2007) since a proposal there is given exogenously, not made by a

[3]This leads them to the study of decision rules that are *durable*.

[4]When incentive constraints are imposed, resilient allocations characterize (a) the *credible core* of Dutta and Vohra (2005), and (b) a core concept very close to Myerson's (2007) if one allows a mediator to randomize over coalitions it approaches for a vote. Without incentive constraints, resilient allocations coincide with the fine core.

particular agent.

We shall focus on the perfect Bayesian equilibria (PBE) of an extensive-form mechanism that is essentially the same as one used by Serrano and Vohra (1997), henceforth SV (1997), to implement the core of an economy with complete information. This mechanism has the desirable property that it is motivated closely by the very description of the core (under complete information). The game's main feature is that, after establishing a status-quo through unanimous agreement in the initial Stage, its final Stage allows any agent to make a proposal of a feasible allocation to a coalition. The proposal is implemented if and only if it is accepted unanimously. Under complete information, SV (1997) show that the set of subgame perfect equilibrium outcomes of the mechanism coincides with the core.

Therefore, taking this mechanism as a launching platform to explore larger domains, we supplant it into a model with incomplete information and rely on its PBE outcomes to suggest a new version of the core. [5] While we are not able to provide a complete characterization of the status-quo PBE outcomes in general settings, we provide useful upper and lower bounds. If coalitions consist of only two agents, as in the case of pairwise stability, these bounds coincide and we have a complete characterization.[6] In general, we refer to the set identified in the best lower bound as the *sequential core*.[7] To sustain allocations as equilibrium outcomes and to knock out others as being not supported, the sequential core allows for complex informational flows between a proposer and responders, with information flowing in either direction, but always under the restrictions imposed by the PBE notion. A result that appears surprising at first is that the sequential core may contain interim inefficient allocations. This stems from the large multiplicity of PBE outcomes (see Example 6 below).

We show that, while there is no logical relationship between the sequential core and the two core notions in Wilson (1978), two PBE refinements justify the coarse core and the fine core; see Subsections 6.1 and 6.2, respectively.

Our paper is related to Okada (2012), which relies on the sequential equi-

[5]PBE arise in which the outcome is different from the status quo, something that was not found under complete information. Those PBE have an unappealing lack of robustness feature (see Example 4), and therefore, we confine our attention to status-quo PBE outcomes.

[6]It is, however, important to bear in mind that the equivalence between pairwise stability and core stability that applies to matching models under complete information may not hold in the presence of incomplete information. Larger coalitions may allow for exchange that is contingent on information not available to smaller coalitions. See for instance Example 2 in Forges (2004) and Example 8 in Liu (2020).

[7]We alert the reader to the fact that the term "sequential core" has also been applied to study a different notion of the core, in a setting with time and (symmetric) uncertainty; see, for example, Gale (1978) and Habis and Herings (2011).

libria of a coalitional bargaining game to motivate the *signaling core*. One important difference is that Okada (2012) concentrates on the signaling aspect of information transmission, while we open the door more broadly to information leaking not just from the proposer to responders but in the other direction as well, as in the phenomenon of screening. We discuss the relationship with Okada (2012) at length in Subsection 6.3. Similarly, the *type-agent core* of de Clippel (2007) is based on an adaptation of the Rothschild-Stiglitz competitive screening model; see Subsection 6.1.

We follow Wilson (1978), and much of the related literature, in assuming that coalitional agreements cannot be renegotiated ex-post. In contrast, Forges (1994) studies a notion of posterior efficiency that makes use of the information revealed by the outcome of a mechanism. In a similar vein, Liu et al. (2014) and Liu (2020) study stable matching when agents observe outcomes and draw inferences from the absence of pairwise blocking.[8] Liu (2022) extends this theory to more general coalitional games. Example 3 in Liu (2020) illustrates well how this approach differs from our framework. Another important difference is that Liu (2020) does not seek to implement stable outcomes through a noncooperative game. Yet, his definition of stability is couched in terms of a matching outcome and beliefs – "off-path beliefs" for counterfactual pairwise blocking and "on-path" beliefs" in the absence of such deviations – which makes for a close connection with our reliance on PBE. In particular, his Bayesian consistency condition is akin to the credible updating rule that we use in Subsection 6.2.

Our results offering support to the sequential core are confirmed in an important robustness check, as sequential core allocations correspond to stationary PBE outcomes of an infinite-horizon coalitional bargaining game, an extension of our basic mechanism that is very close to the model in Moldovanu and Winter (1995).

The rest of the paper is organized as follows. Section 2 presents three illustrative examples. Section 3 introduces the model of an interim economy with asymmetric information. Section 4 introduces our mechanism, and Section 5 contains our results for our basic mechanism. Section 6 discusses the connections between the sequential core and other relevant core notions. Section 7 shows the robustness of our results by ruling out integer games in the basic mechanism and considering coalitional bargaining with an infinite horizon. Section 8 concludes.

[8] See also Fernandez, Rudov, and Yariv (2022), who study a centralized one-to-one matching problem with incomplete information, and show that several results on the Deferred Acceptance algorithm (such as the rural hospital theorem) would break down.

2 Some Illustrative Examples

This Section contains a few simple examples that illustrate how, in our context, private information may get transmitted through a PBE. It may be that only certain types of a proposer offer a particular proposal in equilibrium. This is the well-known phenomenon of signaling. It is also possible that only certain types of a responder would accept a proposal, which conveys information to the proposer. In other words, screening may be the avenue through which information gets transmitted. For these reasons, we should not expect the PBE outcomes of a sequential mechanism to coincide with the coarse core. These examples are only meant to be suggestive, as they are highly stylized and abbreviated versions of the actual sequential game we shall study. The reader may move directly to the next Section without loss of continuity.

Our first example is similar to one used by Lee and Volij (2002) to motivate their solution concept in contrast to the coarse core. It also serves to show the power of signaling in our context.

EXAMPLE **1.** *(Signaling).* Consider a two-agent, two-commodity economy. Only agent 1 has private information: he has two types: $T_1 = \{t_1, t_1'\}$, and each type occurs with probability 1/2. Let $x_i = (x_{i1}, x_{i2})$ denote agent i's commodity bundle. Both agents have identical, state-independent utility functions; $u_i(x_i; s) = \min\{x_{i1}, x_{i2}\}$ for all s and for $i = 1, 2$.[9]

The endowments of the agents, ω, and another allocation, y are shown in the next table.

Agent	Endowment ω		Allocation y	
	t_1	t_1'	t_1	t_1'
1	$(1, 1)$	$(0, 2)$	$(1, 0)$	$(1, 1)$
2	$(2, 0)$	$(2, 0)$	$(2, 1)$	$(1, 1)$

Consider a game in which the informed agent, agent 1, can either propose y or allow ω, the status-quo, to prevail. If he proposes y, agent 2, who is uninformed, must accept or reject the proposal. Rejection leads to ω being implemented and acceptance results in y.

The sequential game is depicted in Figure 1.

Since agent 2 prefers y to ω in *both* states, regardless of her beliefs at her information set, when offered y she must accept. This means that there is

[9]In this example and some of the others in the sequel, utility functions are not strictly monotonic, one of our assumptions. This is done purely for expositional simplicity.

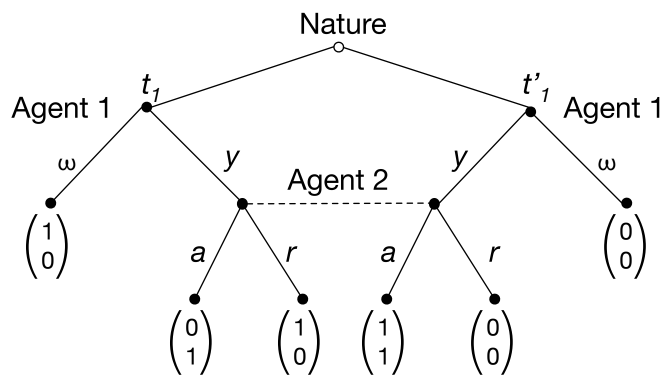

Figure 1: Signaling

a unique PBE, one in which type t_1 will select the status-quo, while type t_1' will propose y, which will be accepted.[10]

Since the initial endowment gives the informed agent the highest possible utility in state t_1, it belongs to the coarse core. But it cannot be supported as a PBE outcome, and does not belong to the sequential core.[11]

◇

EXAMPLE **2.** *(Multiple Equilibria).* Suppose the endowments and preferences remain the same as in Example 1 but the alternative allocation under consideration is y', which differs from y only in state t_1:

$$y_1'(t_1) = (3, 0), \quad y_1'(t_1') = (1, 1)$$
$$y_2'(t_1) = (0, 1), \quad y_2'(t_1') = (1, 1)$$

The new extensive form game is depicted in Figure 2.

The equilibrium outcome similar to the one in Example 1 remains but there is now another PBE in which the outcome is ω (for both states) and the agents are unable to capture the gains from trade in state t_1'. In this equilibrium, agent 2 always rejects and agent 1 chooses ω. This is supported by the off-path belief that if y' is offered, agent 2 assigns probability 1 that the type is t_1. The ruling out of 'bad equilibria' is, of course, an important element of full implementation and we will have to be attentive to this issue when we develop our results. Note that in this example, it is possible for the informed agent to make the case that he would have no incentive to propose

[10]This Example, as well as the next two, are simple in the sense that PBE impose no restrictions on the off-path beliefs; PBE are the same as weak PBE.

[11]Formal definitions of PBE, the coarse core, and the sequential core appear in Sections 3, 4, and 5 below.

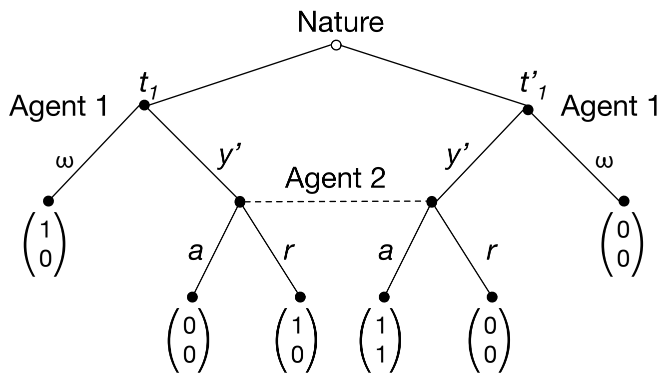

Figure 2: Multiple PBE

y' in state t_1, and forward induction would rule out the bad equilibrium. The only PBE satisfying the intuitive criterion of Cho and Kreps (1987) is one in which type t_1 maintains ω, while type t_1' offers y', which is accepted. $\diamond$

Our third example illustrates the importance of screening in our context; see also Example 9 below.

Example **3.** *(Screening).*
The only modification to Example 2 is that agent 2 is the proposer. She must decide whether or not to propose y'. If she proposes y', then agent 1 either accepts y' or enforces the status-quo. Since the proposer is uninformed, there is no longer any possibility of signaling in the usual sense. However, the alternative allocation, y', has the property that it screens out agent 1 of type t_1 from accepting the proposal; only type t_1' will accept the proposal. So, even though agent 2 is uninformed, she can safely delegate decision-making to agent 1. The only PBE is one in which agent 2 proposes y', which is accepted by t_1' and rejected by t_1. $\diamond$

3 Preliminaries

The basic model of an exchange economy with asymmetric information can be formulated as follows. Let T_i denote the (finite) set of agent i's types. The interpretation is that $t_i \in T_i$ denotes the *private information* possessed by agent i. With $N = \{1, \ldots, n\}$ as the finite set of agents, let $T = \prod_{i \in N} T_i$ denote the set of all information states. We use the notation t_{-i} to denote $(t_j)_{j \neq i}$. Similarly, $T_{-i} = \prod_{j \neq i} T_j$, and for any nonempty $S \subseteq N$, $T_S = \prod_{j \in S} T_j$ and $T_{-S} = \prod_{j \notin S} T_j$. We assume that agents have a common prior probability distribution q defined on T, and that no type is impossible, i.e.,

$q(t_i) > 0$ for all $t_i \in T_i$ for all i. At the interim stage, nature chooses $t \in T$, and each agent i knows her type, t_i. Hence, conditional probabilities will be important: for each $i \in N$ and $t_i \in T_i$, the conditional probability of $t_{-i} \in T_{-i}$, given t_i is denoted $q(t_{-i} \mid t_i)$.

The consumption set of agent i is $X_i \subseteq \mathcal{R}^l_+ \setminus \{0\}$.[12] Agent i's utility function in state t is denoted $u_i(\cdot, t) : X_i \times T \mapsto \mathcal{R}_+$. We assume that for all $i \in N$, $u_i(\cdot, t)$ is continuous and strictly monotonic in the sense that $u_i(x_i, t) > u_i(x'_i, t)$ if $x_i > x'_i$, i.e., $x_{ij} \geq x'_{ij}$ for all commodities j with at least one strict inequality. The endowment of agent i of type t_i is $\omega_i(t_i) \in X_i$. While i's endowment can vary with her type, it does not depend on others' information. We assume that for all i and all t_i, $\omega_i(t_i) > 0$.

We can now define an interim exchange economy as $\mathcal{E} = \langle (u_i, X_i, \omega_i, T_i)_{i \in N}, q \rangle$.

For coalition $S \subseteq N$, a feasible (state-contingent) S-allocation, $x : T \mapsto \prod_{i \in S} X_i$, consists of a commodity bundle for each agent in S in each state such that $\sum_{i \in S} x_i(t) \leq \sum_{i \in S} \omega_i(t_i)$ for all $t \in T$, and satisfying that $x(t_S, t'_{-S}) = x(t_S, t''_{-S})$ for all $t_S \in T_S$ and for all $t'_{-S}, t''_{-S} \in T_{-S}$. (The latter assumption ensures that the set of feasible allocations for a coalition is independent of the information held by the complement). We denote by A_S the set of feasible state-contingent allocations of S. The state-contingent allocations for N are simply referred to as allocations, and denoted A.

Given $x \in A$, the interim utility of agent i of type t_i is:

$$U_i(x \mid t_i) = \sum_{t_{-i} \in T_{-i}} q(t_{-i} \mid t_i) u_i(x_i(t_{-i}, t_i), (t_{-i}, t_i)).$$

Wilson (1978) defines two notions of the core at the interim stage depending on whether or not members of a coalition can share their private information. The definitions we present next follow Wilson (1978) in treating the case in which private information is eventually verifiable and incentive constraints are not relevant. These definitions are simply translations of Wilson's definitions from the language of discernible fields of events to informational types; see the discussion in Forges, Minelli, and Vohra (2002).

When information sharing is not possible, a coalitional objection is required to be commonly known to all members of a coalition. For coalition S, consider an event of the form $E = \prod_{i \in S} E_i \times \prod_{j \notin S} T_j$, where $E_i \subseteq T_i$ for all $i \in S$ and $q(E) > 0$ for all i. We refer to such an event as an *admissible* event for coalition S. An admissible event is said to be *common knowledge* for S if

$$q(t'_{-i} \mid t_i) = 0 \text{ for all } i \in S,\ t_i \in E_i \text{ and } (t'_{-i}, t_i) \notin E.$$

[12] Excluding the zero bundle is mainly for convenience; see Remark 1 below.

Coalition S has a *coarse objection* to $x \in A$ if there exists a common knowledge event E for S and an S-allocation $y \in A_S$, such that

$$U_i(y_i \mid t_i) > U_i(x_i \mid t_i) \text{ for all } t_i \in E_i, \text{ for all } i \in S.$$

The *coarse core* is the set of all feasible allocations to which no coalition has a coarse objection.

In order to consider the possibility that an objection may be directed at an event that is not necessarily commonly known, we need some additional notation. Suppose E is an admissible event for coalition S. Let $q(t \mid E, t_i)$ denote the updated conditional probability of an agent whose type is $t_i \in E_i$ and who believes that the true state lies in E; set to 0 the probability of any state not in E and apply Bayes' rule. For an allocation x define $U_i(x_i \mid E, t_i)$ as the corresponding updated conditional expected utility.[13]

When agents in a coalition can share their private information in an arbitrary manner, Wilson (1978) defines the corresponding notions of an objection and the core as follows.

Coalition S is said to have a *fine objection* to $x \in A$ if there exists an admissible event E for S, and an S-allocation y such that

$$U_i(y_i \mid E, t_i) > U_i(x_i \mid E, t_i) \text{ for all } t_i \in E_i, \text{ for all } i \in S.$$

The *fine core* consists of all allocations $x \in A$ to which no coalition has a fine objection.

The fine core is clearly a subset of the coarse core. Other notions of the core that allow for some (but not arbitrary) information transmission may lie between the fine core and the coarse core. See, for instance, the coarse+ core defined by Lee and Volij (2002) or the signaling core of Okada (2012).

As Wilson (1978) shows, the coarse core of an economy is nonempty under standard assumptions, including convexity and continuity of preferences. While these assumptions are not sufficient to guarantee the nonemptiness of the fine core, Dutta and Vohra (2005) show that quasilinearity of preferences is able to restore it.

4 Implementation of Interim Cores

We assume that the designer knows the data of the economy, $\mathcal{E}$, but not the information possessed by the agents. To implement allocations in some version of the core of this economy, we construct an extensive game form that is inspired by the one constructed in SV (1997) to fully implement the core in a complete-information economy.

[13]Note that, if E is a common-knowledge event, then $U_i(x_i \mid E, t_i) = U_i(x_i \mid t_i)$.

The game form consists of two stages with observable actions:

- In Stage 0, every agent i chooses simultaneously from the choice set $M_i^0 = A \times \mathcal{N} \times \Pi(N-i)$, where $\mathcal{N}$ is the set of integers and $\Pi(N-i)$ is the set of permutations of agents other than i. A typical Stage-0 choice of agent i is denoted $m_i^0 = (x^i, k_i, \pi_i)$. Note that x^i refers to agent i's announcement of a state-contingent allocation, i.e., $x^i = (x_j^i)_{j \in N}$. Let $m^0 = (m_i^0)$ represent the profile of Stage-0 messages and let $1(m^0)$ be the lowest indexed agent who announces the highest integer and $n(m^0)$ be highest indexed agent who announces the lowest integer. As will be explained next, through the announcement of the integer and the permutation of other agents, any agent can become the first in line and make a proposal to the others, which she wants to be responded to according to the protocol she announces.

  For each i and t_i, fix $\epsilon_i(t_i) \in \mathcal{R}^l$ such that $\epsilon_i(t_i) > 0$ and $\omega_i(t_i) - \epsilon_i(t_i) \in X_i$. This is possible because $\omega_i(t_i) > 0$ for all i and t_i.

  All agents observe the announcements (x^i, k_i, π_i). If for any i and j, $x^i \neq x^j$, the outcome is that agent $n(m^0)$ receives $\omega_{n(m^0)}(t_i) - \epsilon_{n(m^0)}(t_i)$, and all other agents receive their initial endowments. If $x^i = x^j = x^*$ for all i and j in N, proceed to Stage 1. In this case, we will refer to x^* as the status-quo.

- In Stage 1, agent $1(m^0)$ chooses a coalition S containing $1(m^0)$, and $y \in A_S$. Let $S = \{1(m^0), 2(m^0), \ldots, k(m^0)\}$, where $(2(m^0), \ldots, k(m^0))$ is consistent with $\pi_{1(m^0)}$, when restricted to S. Then, the other members of S respond sequentially to this proposal (starting with agent $2(m^0)$ and going up to $k(m^0)$) by either accepting it or rejecting it.

  Let t be the true type profile (only known ex post). Then, if all members of S accept y, coalition S is assigned $y(t)$ and all agents not in S are assigned their initial endowments. If any agent in S rejects the proposal, the final outcome is $x^*(t)$. This completes the description of the mechanism.

Notice that the mechanism is feasible, i.e., it prescribes a feasible allocation at every terminal node. With respect to the mechanism in SV (1997), the only departure is the slightly different way in which the endogenous protocol is determined, in terms of an integer and a permutation of the other agents (as opposed to a permutation of all agents; this change is necessary due to incomplete information). Since our interest is to construct a single mechanism whose set of PBE outcomes corresponds to a notion of the core, we resort to the integer game and choice of protocol in stage 0. Integer games can be criticized for their unnatural strategic properties, and so it is worth pointing out that we can obtain a similar result, based on a class of mechanisms identical to the one just described, but that dispense with these devices. Each mechanism in that class is indexed by a protocol of agents, which is exogenously given; see Subsection 7.1. Similarly, as a confirmation of our findings, Subsection 7.2 dispenses with the simultaneous announcements of Stage 0 and studies an infinite-horizon coalitional bargaining model that simply repeats indefinitely the rules in Stage 1 of the foregoing mechanism.

Given that the agents' announcements are public, the information sets in which beliefs are to be specified in on- and off-equilibrium paths concern only Nature's move and they are described as follows: for each agent i of type t_i, she must assign probabilities to the type profiles $t_{-i} \in T_{-i}$. Notice that this is so for the Stage 1 proposer and for agents with the role of responders whose information set is reached: this happens only if all previous responders have accepted the proposal made by the first agent in the protocol.

5 Perfect Bayesian Equilibria and Core Stability

5.1 The Equilibrium Concept

To define a perfect Bayesian equilibrium (PBE) in the game induced by the mechanism in Section 4, let $\sigma = (\sigma_j)_{j \in N}$ be a strategy profile, where each $\sigma_j : I_j \mapsto A_{js}$ – here, I_j denotes the set of agent j's information sets, and A_{js} denotes the action set for agent j in period s. Each agent j holds beliefs $\beta_j(\iota_{js})$ at each of her information sets ι_{js} : $\beta_j(\iota_{js})$ is a probability distribution over the decision nodes of the information set ι_{js}. Since, other than the type profiles chosen by Nature, actions of all agents are observable, β_j is a probability distribution over T_{-j}. The collection $(\beta_j(\iota_{js}))_{\iota_{js} \in I_J}$ for all information sets ι_{js} in the set of agent j's information sets is agent j's system of beliefs.

For simplicity, we follow Fudenberg and Tirole (1991) and assume independent types (we discuss this issue below): for every $t \in T$, $q(t) = \prod_{i \in N} q_i(t_i)$. The history at period s (within Stage 1) is denoted $h^s = (a_0, a_1, \ldots a_{s-1})$.

A profile of strategies $\sigma = \{\sigma_j; j \in N\}$ and a system of beliefs $\beta = \{\beta_j; j \in N\}$ is a *perfect Bayesian equilibrium* (PBE) if they satisfy the following properties:

(P.1) Sequential rationality: For all j and for all information sets ι_{js}, σ_j prescribes a best response to σ_{-j} given the beliefs $\beta_j(\iota_{js})$.

(P.2) Bayes' rule is used to update beliefs whenever possible. In other words, when the history changes from h^s to h^{s+1}, and a_j^s is consistent with the new history, Bayes' rule is used to compute the beliefs of all agents regarding agent j: $\beta(t_j \mid h^{s+1})$. Note that this is stronger than using Bayes' rule only on the equilibrium path.

(P.3) Posterior beliefs are independent and all types of agent i have the same beliefs: for all $t_i \in T_i$ and all histories h^s, $\beta_i(t_{-i} \mid t_i, h^s) = \prod_{j \neq i} \beta_i(t_j \mid h^s)$.

(P.4) Agents i and j have the same beliefs about agent $k \neq i, j$: $\beta_i(t_k \mid h^s) = \beta_j(t_k \mid h^s) = \beta(t_k \mid h^s)$. If agent k is the last (and only) agent to have moved the history to h_s, we denote by $\beta^k(h_s)$ the probability distribution over T_k representing the (common) beliefs of all the other agents about k's type.

(P.5) "No signaling what you don't know" in the sense that beliefs about agent k's type must depend only on agent k's actions. In particular, if the only agent to have moved at stage s is agent j, then for $k \neq j, i$, for all histories h^s, $\beta_i(t_k \mid h^s, a_s) = \beta_i(t_k \mid h^s)$.

This is the definition of a PBE in Fudenberg and Tirole (1991) adapted to our model. Generalizing the notion of PBE without using independent types would require changes to properties (P.3) and (P.5), as follows. Condition (P.3) is modified so that it no longer requires independence. Condition (P.5) should capture the general idea that a unilateral move by agent j conveys information about other agents' types only to the extent that it conveys information about j's type. In other words, for $i \neq j$, i's posterior *conditional* beliefs about agents other than j remain unchanged.[14]

(P.3') All types of agent i have the same beliefs: $\beta_i(t_{-i} \mid t_i, h^s) = \beta_i(t_{-i} \mid t_i', h^s) = \beta_i(t_{-i} \mid h^s)$ for all $t_i, t_i' \in T_i$ for all histories h^s.

(P.5') "No signaling what you don't know": Let j be the only agent who moves in period $s+1$. Then, for $i \neq j$, $\beta_i(t_{-\{i,j\}} \mid t_i, t_j, h^s, a_s) = \beta_i(t_{-\{i,j\}} \mid t_i, t_j, h^s)$.

While we refrain from complicating the writing of our proofs to incorporate correlated types, we shall feel free to make use of examples that are not restricted to independent types.

[14] Condition (P.4) would also need to be modified, but it is not really essential for our result even under independent types, as its only role is to simplify our notation; we shall not bother to write its modification.

Given the particular structure of our mechanism, it is useful to consider certain kinds of PBE that appear frequently in our analysis and identify the information sets that are off the equilibrium path. Consider a PBE (σ, β) in which Stage 1 always begins with a proposal (S', y'). Suppose agent i deviates and makes a proposal $(S, y) \neq (S', y')$. (This deviation may also require i to announce the highest integer at Stage 0 to become the proposer). Now, in the continuation game that follows, the first responder, if any, would be at an information set that is off the equilibrium path. Let β^i be the first responder's beliefs about i. Since Bayes' rule cannot be applied at this information set, β^i can be arbitrary in equilibrium. By conditions (P.3) and (P.4), all responders must have the same beliefs, β^i, about the deviator/proposer's type.

Next, consider a continuation game immediately following a proposal (S, y) by agent i. (This must, of course, mean that $|S| \geq 2$). Let the number of responders be denoted k and suppose, according to the protocol resulting from the strategies in Stage 0, they are ordered $\{j_1, \ldots, j_k\}$, where j_1 is the first responder and j_k the last. There are two kinds of responders' strategies immediately following (S, y) that are of particular interest:

Case (a): For every responder j, there is a nonempty set of types $E_j \subseteq T_j$ that accept the proposal. When agent j_2 moves, following an acceptance by j_1, her equilibrium beliefs about j_1 can be derived from Bayes' rule because E_{j_1} is nonempty. The same argument applies to any other responders that follow. By (P.2), therefore, the beliefs of any responder j about another responder k are pinned down by applying Bayes' rule to the event E_k.

Case (b): The equilibrium strategies following the proposal (S, y) require every responder of every type to reject whenever she arrives at a decision node. If responder m arrives at a decision node, with all her predecessors having unexpectedly accepted, she cannot apply Bayes' rule and so her beliefs about her predecessors could be arbitrary, though still subject to conditions (P.3), (P.4), and (P.5).

To be sure, there are other continuation games with 0 probability information sets, but it will not be necessary for us to describe in detail the beliefs that are permissible there.[15]

[15] For instance, there is a continuation game in which, after j_1 unexpectedly accepts the proposal, some of agent $j_2's$ types accept. Then all responders other than j_1 could have arbitrary, but identical, beliefs about j_1. However, all responders that follow j_2 must hold beliefs about j_2 that are derived from Bayes' rule.

5.2 Constructing Equilibria: A Lower Bound on the Set of PBE Outcomes

In general, a notion of core stability that corresponds to PBE must allow for informational leakages to go both ways, from the proposer to the responders and vice versa, but only in ways that are acceptable, according to the rules for equilibrium updating of beliefs in a PBE. In this subsection, we present two results. The first one features particularly simple PBEs, while the second uses a more complex construction in order to support a larger set of outcomes.

Suppose (σ, β) is a PBE in which Stage 1 always begins with a proposal (S', y'). Consider a deviation by agent i in which she makes a proposal $(S, y) \neq (S', y')$. As we noted in the previous Subsection, in the continuation game that follows this proposal, Bayes' rule does not impose any restrictions on the (common) beliefs that the responders have about agent i. Let these beliefs be denoted β^i, i.e., $\beta^i(t_i)$ is the posterior probability assigned by the other agents to agent i being of type t_i. By (P.5), any responder to (S, y) must continue to hold prior beliefs about agents not in S. Suppose every responder also continues to hold prior beliefs about any other responder. Then the interim utility of responder j from an allocation x is:

$$U_j(x \mid t_j, \beta^i) = \sum_{t_{-j} \in T_{-j}} \beta^i(t_i) \prod_{k \neq i,j} q(t_k) u_j(x_j(t_{-j}, t_j), (t_{-j}, t_j)).$$

To be sure that a proposal by i to coalition S will be accepted with positive probability, it must be the case that, regardless of the beliefs that the responders hold about i's type, there exists a set of types for whom it is rational to accept.

Suppose agent i makes a proposal (S, y) when the status-quo is x, and β^i represents the beliefs held by $j \in S - i$ about i's types. We say that a nonempty event $E(\beta^i) = \prod_{j \in S-i} E_j(\beta^i) \times T_{-S}$, where $E_j(\beta^i) \subseteq T_j$ for all $j \in S - i$, is *a plausible event given* β^i if, for all $j \in S - i$,

$$U_j(y \mid t_j, \beta^i, E(\beta^i)) > U_j(x \mid t_j, \beta^i, E(\beta^i)) \text{ if and only if } t_j \in E_j(\beta^i). \quad (1)$$

We will refer to the types in $E_j(\beta^i)$ as j's *plausible types given* β^i.

Given a status-quo x, a proposal (S, y) by agent $i \in S$ is a *weak sequential objection* to x, if for every belief β^i there exists a plausible event $E(\beta^i)$ satisfying (1) and there exists $t_i(\beta^i) \in T_i$ such that

$$U_i(y \mid t_i(\beta^i), E(\beta^i)) > U_i(x \mid t_i(\beta^i), E(\beta^i)). \quad (2)$$

This formalizes the idea that an objection proposal by i is such that regardless of the beliefs that the responders have about i's types, there is positive probability that all agents in S stand to gain. While such an improvement holds for all β^i, the set of plausible types of responders as well as a proposer type may well depend on β^i, as expressed by the notation $E_j(\beta^i)$ and $t_i(\beta^i)$. Note that each responder is assumed to take account of the other plausible types of responders solely through Bayesian updating, by conditioning on $E(\beta^i)$. Because the beliefs about other responders are the prior, and not some other arbitrary beliefs, we refer to such an objection as 'weak'. (We will later consider a stronger form of an objection that will allow responders to hold other kinds of beliefs about fellow responders).

An allocation x is in the *strong sequential core* if there does not exist a weak sequential objection against it.

This is related to a concept introduced in Lee (1998), later extended in Lee and Volij (2002). Lee (1998) defines an *individualistic objection* (S, y) to x if there is $t_i \in T_i$ for some $i \in S$ that gains in interim utility while all types of agents $j \in S$, $j \neq i$ have an improvement for every possible type of agent i. In other words, *all* types of agents other than i are guaranteed an improvement regardless of i's type. This means that every $j \in S, j \neq i$, does better whatever her beliefs about i's type. Hence, while an individualistic objection (S, y) to x by i requires *all* types of agents $j \neq i$ to gain, a weak sequential objection allows for only *some* of the types of $j \in S - i$ – those in $E_j(\beta^i)$ – to gain. An individualistic objection is therefore a weak sequential objection, with $E_j(\beta^i) = T_j$ for all j in $S - i$ and all β^i, and the strong sequential core is a subset of the individualistic core. The difference between the strong sequential core and the individualistic core lies in the fact that the latter does not account for screening – information about the responder being transmitted by the equilibrium strategy to the proposer. We will show below that this inclusion relationship may be strict; there may be an allocation in the individualistic core that is not in the strong sequential core (see Remark 4).

As a first step in constructing PBEs of the mechanism, we show that every strong sequential core allocation is an outcome of a PBE:

PROPOSITION **1.** *Every allocation in the strong sequential core is an outcome of a PBE of the mechanism.*

Proof. Let x be an allocation in the strong sequential core. We claim that it is supported by the following PBE strategies and beliefs. At Stage 0, x is unanimously agreed to and every agent $i \in N$ announces the integer 0 and some protocol. At Stage 1, the winner of the integer game proposes (N, x), which is rejected by all, so the outcome is x.

If agent i were to deviate at Stage 0 and announce an allocation different from x, she would receive, at best, her initial endowment. If this is profitable, coalition $\{i\}$ has a weak sequential objection, which contradicts the assumption that x is in the strong sequential core.[16] And once (N, x) is proposed, at Stage 1 there is no profitable deviation for any responder because the outcome, x, cannot be changed. The only remaining possibility for a profitable deviation is that an agent, say i, deviates to be the proposer and makes a proposal $(S, y) \neq (N, x)$. We will now describe the off-path beliefs and actions that will make such a deviation unprofitable.

Since x is in the strong sequential core, there must exist a belief β^i such that there does not exist a plausible event $E(\beta^i)$ and $t_i(\beta^i)$ such that conditions (1) and (2) both hold. For any proposal (S, y), pick any such β^i to be the beliefs about i following the proposal. This is permissible in defining a PBE because this proposal is off-path. To describe the actions and beliefs of the responders to such a proposal there are two cases to consider.

Case (a): There exists a plausible event $E(\beta^i)$ given β^i, with plausible types $E_j(\beta^i)$, such that (1) is satisfied for each $j \in S - i$. Fix any such set of plausible types $E_j(\beta^i)$ and let each $j \in S - i$ accept if and only if $t_j \in E_j(\beta^i)$.

Case (b): There does not exist a set of plausible types satisfying (1). Then each responder of every type rejects the proposal whenever required to take an action.

We first check the responders' incentives. Consider case (a). Since $E_j(\beta^i) \neq \emptyset$ for every $j \in S - i$, the strategies we have specified imply that, except for the first responder, each responder's information set – in the continuation game following the proposal (S, y) – is reached with positive probability. By (P.2), this means that while all responders hold beliefs β^i about i, at all information sets for responders other than the first one, beliefs are determined by applying Bayes' rule. Given the specified strategies and beliefs of other responders, the interim expected utility of a responder of type t_j from accepting the proposal is derived from y over $E(\beta^i)$ and x over its complement. Her interim expected utility from rejecting the proposal is $Uj(x \mid t_j, \beta^i)$. In other words, (1) implies that a plausible type cannot gain by unilaterally rejecting the proposal and implementing x rather than y, and a nonplausible type cannot gain by unilaterally accepting the proposal. That is, no responder has a profitable deviation from the proposed strategies.

In case (b), a deviation can change the outcome only if there is a positive probability event that all the responders accept. This means that if there is a profitable deviation, then for every $j \in S - i$, the set of types that accepts, $\hat{E}_j$, is nonempty. Let the beliefs of all the responders about other responders continue to be their prior beliefs.[17] But then, the fact that types

[16]In other words, the strong sequential core satisfies interim individual rationality.

[17]Note that because the proposed strategies require every responder to always reject at every information set that follows, every responder who follows the first one is responding over a 0 probability event and could, in principle, hold arbitrary beliefs about her prede-

in $\hat{E}_j$ accept means that for all $j \in S - i$, the inequalities in (1) hold for the event $\hat{E} = \prod_{j \in S-i} \hat{E}_j \times T_{-S}$. In other words, $\hat{E}$ is a plausible event, a contradiction.

Next, we check i's (i.e., the proposer's) incentives in the deviation. Clearly, a proposal that belongs to case (b) cannot improve upon x. In case (a), if agent i proposes $(S, y) \neq (N, x)$, there exists a plausible event $E(\beta^i)$. Since x belongs to the strong sequential core and condition (1) holds, condition (2) cannot hold, i.e., agent i cannot profit from this deviation. This completes the proof that the strategies and off-path beliefs we have described are a PBE, with outcome x. □

Next, we proceed to construct more involved PBEs. Indeed, while a weak sequential objection allows for arbitrary (updated) beliefs regarding the proposer, it does not consider the possibility that a responder hold arbitrary beliefs about other responders; beliefs about other responders were derived from Bayes' rule, conditioning on $E(\beta^i)$. In general, in some continuation games, a PBE would allow for such updated off-path beliefs. In such cases, the strong sequential core may not capture all the PBE outcomes of the mechanism. We will therefore need to strengthen a weak sequential objection to allow for additional updating of beliefs. Thus, it seems appropriate to use the protocol that determines the order of moves to be incorporated into the definition of an objection.

Let π be a protocol or order of agents in coalition S. We will often refer to the first agent in S according to π as the 'proposer'. For each $j \in S$, let $P_j(\pi)$ denote the set of predecessors of j according to π. For every $i, j \in S, i \neq j$, let β_j^i denote the belief of agent j regarding agent i's type. (Beliefs about agents not in S remain the prior beliefs). Given (S, π), a collection of beliefs $\beta_S = (\beta_j^i)_{i,j \in S}$ is said to be *consistent* if:

(i) For all $j \in S$ $\beta_j^k = q^k$ for all $k \notin P_j(\pi)$; no updating of beliefs regarding agents who have not moved previously within Stage 1.

(ii) For all $j, k \in S$, $\beta_j^i = \beta_k^i$ for all $i \in P_j(\pi) \cap P_k(\pi)$; agents have the same beliefs about a common predecessor.

Suppose agent i makes a proposal (S, y) when the status-quo is x, the protocol is π with i as the proposer and β_S is a collection of beliefs consistent with (S, π). We say that a nonempty event $E(\beta_S) = \prod_{j \in S-i} E_j(\beta_S) \times T_{-S}$, where $E_j(\beta_S) \subseteq T_j$ for all $j \in S - i$, is *a plausible event given* β_S if, for all $j \in S - i$,

$$U_j(y \mid t_j, \beta_j, E(\beta_S)) > U_j(x \mid t_j, \beta_j, E(\beta_S)) \text{ if and only if } t_j \in E_j(\beta_S). \quad (3)$$

cessor. In constructing an equilibrium we are, of course, permitted to choose these to be the prior beliefs.

We will refer to the types $E_j(\beta_S)$ as j's *plausible types given* β_S.

Given a status-quo x, a proposal (S, y) by agent $i \in S$ is a *sequential objection* to x, if for every collection of beliefs β_S consistent with (S, π) there exists a plausible event $E(\beta_S)$ satisfying (3) and there exists $t_i(\beta_S) \in T_i$ such that

$$U_i(y \mid t_i(\beta_S), E(\beta_S)) > U_i(x \mid t_i(\beta_S), E(\beta_S)). \tag{4}$$

The *sequential core* consists of all allocations $x \in A$ to which there does not exist a sequential objection.

Since any sequential objection is a weak sequential objection, the sequential core is a superset of the strong sequential core.

We can now proceed to construct additional PBEs of the mechanism, and strengthen Proposition 1.

PROPOSITION **2.** *Every allocation in the sequential core is an outcome of a PBE of the mechanism.*

Proof. Let x be an allocation in the sequential core. We shall construct PBE strategies and beliefs that support x as the equilibrium outcome. As in the proof of Proposition 1, consider strategies such that at Stage 0, x is unanimously agreed to and at Stage 1 the winner of the integer game proposes (N, x), which is rejected by all, so the outcome is x.

Following the proof of Proposition 1, we know that the only possibility for a profitable deviation is that agent i deviates to be the proposer and makes a proposal $(S, y) \neq (N, x)$. We will now describe the off-path beliefs and actions that will make such a deviation unprofitable. In what follows, we will denote by π the protocol induced by π_i over S, with i as the proposer.

Since x is in the sequential core, given any (S, y, π), there exists β_S, consistent with (S, π), for which there does not exist an event $E(\beta_S)$ and $t_i(\beta_S)$ such that (3) and (4) both hold. Fix any such β_S corresponding to every proposal (S, y) and let β^i denote the responders' beliefs about agent i. To describe the actions and beliefs of the responders, there are two cases we consider:

Case (a): There exists a plausible event $E(\beta_S)$ given β_S, with plausible types $E_j(\beta_S)$, such that (3) is satisfied for each $j \in S - i$. Let each $j \in S - i$ accept if and only if $t_j \in E_j(\beta_S)$.[18]

Case (b): There does not exist a set of plausible types satisfying (3). Then every responder of every type rejects.

In case (a), for both the responders and the proposer, the arguments used in the proof of Proposition 1 apply virtually unchanged to show that because

[18]Note that in this case, since for every responder there exists a set of types that accept, all responders must make use of Bayes' rule to derive updated beliefs about other responders. Thus, (3) in this case is the same as (1).

x is in the sequential core, (4) cannot hold and i cannot have a profitable deviation.

In case (b), the proposer's payoff is unchanged as the proposal is rejected. For responders' deviations in this continuation game, let the beliefs of the responders be β_S. Such a deviation can change the outcome only if there is a positive probability event, say $E(\beta_S)$, over which all the responders accept. Given that the beliefs are β_S, this means that (3) is satisfied, a contradiction to the fact that there does not exist a set of plausible types given β_S. □

Note that, in all PBEs constructed so far (in the proofs of Propositions 1 and 2), the equilibrium outcome coincides with the status-quo. Although Proposition 2, compared to Proposition 1, has expanded the outcomes that can be supported by PBEs of our mechanism, it is possible that there are even more complex PBEs that correspond to outcomes that are not in the sequential core. In the next Subsection, we provide an upper bound to the entire set of PBE outcomes in which the outcome is the same as the status-quo from Stage 0. Fortunately, in some cases, for example, if effective coalitions have no more than two agents, the lower bound identified in Proposition 2 and the upper bound identified in Proposition 3 of the next Subsection coincide. Yet additional nonstatus-quo PBE outcomes can also be found; these display a lack of robustness, as the next subsections describe.

Another property of the PBEs we constructed in Propositions 1 and 2 is that in certain continuation games if there was a set of plausible types for each responder, then all non-plausible types reject the proposal. This allows for the possibility that a non-plausible type is indifferent between accepting and rejecting the proposal, but the equilibrium strategies we constructed had them reject. In the next Subsection, when we work from the other direction, given an arbitrary PBE, we may have to confront the possibility that an indifferent type may accept in equilibrium. Fortunately, this issue is easily resolved given our assumption that in each state the utility functions are monotonic and the consumption set excludes the zero bundle. Under these assumptions, we can break any such indifference by reducing the commodity bundle of each non-plausible type so that rejection is strictly better than acceptance. More precisely, suppose (S, y) is a weak sequential objection to x with $E_j(\beta^i)$ the set of plausible types of responder j when the beliefs about the proposer are β^i. Then there is a weak sequential objection (S, y') such that (1) can be strengthened to:

$$U_j(y' \mid t_j, \beta^i, E(\beta^i)) > U_j(x \mid t_j, \beta^i, E(\beta^i)) \text{ for all } t_j \in E_j(\beta^i) \tag{1.1}$$

and

$$U_j(y' \mid t_j, \beta^i, E(\beta^i)) < U_j(x \mid t_j, \beta^i, E(\beta^i)) \text{ for all } t_j \notin E_j(\beta^i). \tag{1.2}$$

In other words, we have

Remark 1. *Suppose (S, y) is a weak sequential objection to x with $E_j(\beta^i)$ as the set of plausible types of responder j when the beliefs about the proposer are β^i. Then there is no loss of generality in assuming that all non-plausible types of responders strictly prefer x to y. A similar statement applies to a sequential objection.*

5.3 An Upper Bound on the Set of PBE Outcomes

To formulate a notion of an objection that is based on an individual proposer that is sure to be accepted by others in the coalition, we will allow responders to hold arbitrary beliefs about each other. Hence, we shall say that responder j of type t_j *uniformly prefers* allocation z to allocation z' in coalition S given a belief β^i about the proposer i if she prefers z to z' for *any* (updated) beliefs she may hold about the types of other responders in S. In other words,

$$\begin{aligned} \text{for all } t_{S-ij} \in T_{S-ij}, \quad & \textstyle\sum_{t_i \in T_i, t_{-S} \in T_{-S}} \beta^i(t_i) q(t_{-S}) u_j(z_j(t_j, t_{S-ij}, t_i, t_{-S}), t)) \\ > \quad & \textstyle\sum_{t_i \in T_i, t_{-S} \in T_{-S}} \beta^i(t_i) q(t_{-S}) u_j(z'_j(t_j, t_{S-ij}, t_i, t_{-S}), t)) \end{aligned}$$

While prior beliefs are maintained about agents outside S, and β^i about the proposer, each responder is permitted to entertain (arbitrary) alternative beliefs about all other responders in S.

Suppose agent i makes a proposal (S, y) when the status-quo is x, and β^i represents the beliefs held by $j \in S - i$ about i's types. We say that a nonempty event $E^s(\beta^i) = \prod_{j \in S-i} E^s_j(\beta^i) \times T_{-S}$, where $E^s_j(\beta^i) \subseteq T_j$ for all $j \in S - i$, is *a strongly plausible event given* β^i if, for all $j \in S - i$,

$$E^s_j(\beta^i) = \{t_j \in T_j | t_j \text{ uniformly prefers } y \text{ to } x \text{ in coalition } S \text{ given } \beta^i\} \quad (5)$$

and

$$T_j - E^s_j(\beta^i) = \{t_j \in T_j | t_j \text{ uniformly prefers } x \text{ to } y \text{ in coalition } S \text{ given } \beta^i\}. \quad (6)$$

We will refer to the types in $E^s_j(\beta^i)$ as j's *strongly plausible types given* β^i. Note the strong separation of incentives in the types of all agents $j \in S - i$: those who accept the objection have a uniform (strict) preference for doing so, while the ones rejecting have a uniform (strict) preference for the status-quo.

Given a status-quo x, a proposal (S, y) by agent $i \in S$ is a *strong sequential objection* to x, if for every belief β^i there exists a strongly plausible event $E^s(\beta^i)$ satisfying (5) and (6) and there exists $t_i(\beta^i) \in T_i$ such that

$$U_i(y \mid t_i(\beta^i), E^s(\beta^i)) > U_i(x \mid t_i(\beta^i), E^s(\beta^i)). \quad (7)$$

The *weak sequential core* consists of all allocations $x \in A$ to which there does not exist a strong sequential objection.

In general, the sequential core, which contains the strong sequential core, is itself contained in the weak sequential core.

In a two-agent coalition, with one proposer, i, and one responder, j, (5) reduces to:

$$E_j^s(\beta^i) = \{t_j \in T_j | U_j(y \mid t_j, \beta^i) > U_j(x \mid t_j, \beta^i)\}$$

and (6) to:

$$T_j - E_j^s(\beta^i) = \{t_j \in T_j | U_j(y \mid t_j, \beta^i) < U_j(x \mid t_j, \beta^i)\}$$

By Remark 1, this yields the following observation:

REMARK **2.** *If coalitions cannot have more than two members (as in pairwise stability), there is no difference between the weak sequential core, sequential core and strong sequential core.*

Our next result shows that the weak sequential core is an upper bound to all PBE outcomes that coincide with the status-quo.

PROPOSITION **3.** *Every status-quo PBE outcome of the mechanism is in the weak sequential core.*

Proof. Suppose there is a PBE in which the outcome is x. First, note that at Stage 0 the equilibrium strategies must be such that all agents announce the same allocation. Otherwise, the highest indexed agent to announce the lowest integer, who gets penalized, could have done better by announcing a higher integer. Let x' be the common allocation announced at Stage 0. By assumption, $x = x'$.

Suppose then that x is not in the weak sequential core. Let (S, y) be a strong sequential objection to x by agent i. Given the PBE, consider the deviation in which agent i chooses the highest integer, and in Stage 1 she proposes (S, y). Let β^i be the off-path beliefs prescribed by the PBE. Since (S, y) is a strong sequential objection, there exists a type $t_i(\beta^i)$ and strongly plausible types of responders, $E_i^s(\beta^i)$, such that (5), (6) and (7) are satisfied. Whatever beliefs the responders have about each other following this deviation, by sequential rationality, each type in $E_i^s(\beta^i)$ will accept the proposal, since they uniformly prefer y over x given β^i, while types not in $E_i^s(\beta^i)$ will reject it, since they uniformly prefer x over y. Hence, this is a profitable deviation for type $t_i(\beta^i)$, which is a contradiction. □

We summarize the best bounds provided in Propositions 2 and 3 in a single statement, our first main result, which also uses Remark 2:

Theorem 1.

(i) Every allocation in the sequential core is a status-quo PBE outcome of the mechanism.

(ii) Every status-quo PBE outcome of the mechanism is in the weak sequential core.

(iii) When effective coalitions consist of at most two agents, the sequential core coincides with the set of status-quo PBE outcomes of the mechanism.

We conclude this Section with the observation that the weak sequential core is not an upper bound to the *entire* set of PBE outcomes. The next example shows that one can construct additional PBEs, in which the outcome is no longer the status-quo, and such an outcome is not in the weak sequential core. The example presents a knife-edge construction, which makes it somewhat unappealing. This suggests that there should be a suitable refinement of the PBE concept that would eliminate such possibilities, and our conjecture is that such a refinement should imply the weak sequential core property:

Example **4.** There are two agents and two commodities. Each agent has two types. Suppose (t_1^*, t_2^*), (t_1^*, t_2') and (t_1', t_2^*) are all equally likely states, and (t_1', t_2') has 0 probability. For both agents, the commodities are perfect complements in state (t_1^*, t_2^*) (i.e., $u_i(x_i; (t_1^*, t_2^*)) = \min\{x_{i1}, x_{i2}\}$ for $i = 1, 2$). In state (t_1^*, t_2'), agents' utilities are $u_1(x_1; (t_1^*, t_2')) = x_{11}$ and $u_2(x_2; (t_1^*, t_2')) = x_{21} + x_{22}$. In state (t_1', t_2^*), agents' utilities are $u_1(x_1; (t_1', t_2^*)) = x_{11} + x_{12}$ and $u_2(x_2; (t_1', t_2^*)) = x_{22}$.

The next table shows the endowment along with allocation x.

	t_2^*	t_2'
t_1^*	$(2, 0), (0, 2)$	$(2, 0), (0, 1)$
t_1'	$(1, 0), (0, 2)$	

(a) Endowment ω

	t_2^*	t_2'
t_1^*	$(0.5, 0.5), (0.5, 0.5)$	$(1.5, 0), (0.5, 1)$
t_1'	$(1, 0.5), (0, 1.5)$	

(b) Allocation x

There is a PBE of the mechanism in which the status-quo is ω and the equilibrium outcome is x. These are the relevant strategies and beliefs:

In Stage 0, every type of every agent announces the endowment, the integer 0, and any permutation. Therefore, the status-quo is the endowment allocation. In Stage 1: every proposer makes the offer (N, x) on the equilibrium path, and it is accepted by all types. Following any deviation proposal, the informed types accept if and only if her utility is at least 1, while the uninformed types reject it. In any such off-path, the belief of the uninformed type is that the proposal comes from the informed type with probability 1 (pessimistic beliefs).

Note how given these pessimistic beliefs, the rejection is a best response in that continuation game, even if the deviation proposal were to offer the uninformed agent the entire endowment in the off-diagonal state (in this case, both acceptance and rejection are best responses). It follows that if we perturb the status-quo slightly to be anything different from the endowment, the resulting configuration is no longer a PBE, as such a deviation must be accepted by the uninformed types, even under pessimistic beliefs.

Finally, the allocation x is not in the weak sequential core. Indeed, consider an objection $(\{1,2\}, z)$, where z differs from ω only in state (t_1^*, t_2^*), where it assigns the bundle $(1,1)$ to each agent. Then, with $t_1(\beta^1) = t_1^*$, $E_2(\beta^1) = \{t_2^*\}$ for every β^1, we have a strong sequential objection to x. Thus, there are nonstatus-quo PBE outcomes that lie outside of the weak sequential core. ◇

6 Relationship with Other Core Notions

One should perhaps expect that all the complexities in the revelation of information built into the notion of a PBE would result in difficulties connecting with previous core concepts in this literature, and indeed this is what we find. As a basic desideratum, the sequential core does satisfy interim individual rationality. However, it is not directly comparable to the coarse core; there may be allocations in the coarse core that are not in the sequential core, and vice versa, as shown in Examples 1, 5, and 6; see also Proposition 4 below for a refinement that results in all equilibrium outcomes being in the coarse core. Similarly, the fine core may not be included in the sequential core; see Example 7 below. This example and Proposition 5 detail this complex relationship.

6.1 The Coarse Core

We have already seen in Example 1 that information transmission through signaling can destabilize a coarse core allocation; an allocation in the coarse core may not be a PBE of a sequential game. However, that example is not totally convincing because it relies on defining objections as *strict* improvements for each type. Note that the endowment in Example 1 is *not* classically interim efficient. It is Pareto dominated by the allocation that assigns ω in state t_1 and y in state t_1'. Can we find a more robust example, not relying on strict improvement for all types, to show that a coarse core allocation may not be in the sequential core? The following three-agent example shows that the answer to this question is in the affirmative.

EXAMPLE **5.** There are three agents, three commodities and two equally likely information states, a and b. Agents 1 and 2 are fully informed while agent 3 is uninformed.[19] Each agent i has a state-independent endowment of 2 units of commodity i. The utility functions of the three agents are shown in the following table, along with two feasible allocations, x and y.

Agent	Utility function		Endowment	Allocation x		Allocation y	
	a	b		a	b	a	b
1	$\min\{x_1, x_2\}$	$\min\{x_1, x_3\}$	$(2, 0, 0)$	$(1, 1, 0)$	$(0, 1, 0)$	$(1, 0, 1)$	$(1, 0, 1)$
2	$\min\{x_1, x_2\}$	$x_1 + x_2$	$(0, 2, 0)$	$(1, 1, 0)$	$(2, 1, 0)$	$(0, 2, 0)$	$(0, 2, 0)$
3	$\min\{x_1, x_3\}$	$\min\{x_1, x_3\}$	$(0, 0, 2)$	$(0, 0, 2)$	$(0, 0, 2)$	$(1, 0, 1)$	$(1, 0, 1)$

A coarse core allocation x that is not in the sequential core

We claim that x is in the coarse core even if we allow for coarse objections where no type loses and at least one type gains. Allocation x has the property that increasing the expected utility of agent 3 would require the utility of at least one of the other agents to become *strictly* lower than it is at x. Since 1 and 2 are fully informed, this means that x is classically interim efficient in the sense of HM. As agents 2 and 3 have no interest in each other's endowment, and x is interim individually rational, the only possible coarse objection to x must come from coalition $\{1, 3\}$. But in this coalition, agent 1's utility in state a is 0 (strictly less than at x), so there is no coarse objection to x.

Although x is in the coarse core, if the true state is b agent 1 would prefer to replace x with y. Note that y is feasible for coalition $\{1, 3\}$. Because agent 3 prefers y to x regardless of the state, this means that y is a sequential objection to x, led by agent 1.The instability of x in this example is reminiscent of the idea that it is not durable in the sense of HM, or not resilient in the sense of SV (2007). It is not precisely the same phenomenon, because we are not imposing incentive compatibility.[20] $\diamond$

On the other hand, there is a great deal of flexibility in assigning off-path beliefs in a PBE, so there may exist PBE outcomes that are not in the coarse

[19] This information structure corresponds to one in which agents 1 and 2 are of two possible types that are perfectly correlated.

[20] See also de Clippel and Minelli (2002), who analyze a two-person bargaining problem with verifiable types, where agent 1 proposes an allocation rule and agent 2 chooses whether to accept agent 1's proposal, and provide an example (Example 11) to illustrate that the coarse core should not be expected as a reasonable outcome of a bargaining process.

core, as the next Example illustrates. It is therefore possible that a sequential core allocation is not in the coarse core.

Example **6.** Consider the economy in Example 4 again. That is, there are two agents and two commodities. Each agent has two types. Suppose (t_1^*, t_2^*), (t_1^*, t_2') and (t_1', t_2^*) are all equally likely states, and (t_1', t_2') has 0 probability. For both agents, the commodities are perfect complements in state (t_1^*, t_2^*) (i.e., $u_i(x_i; (t_1^*, t_2^*)) = \min\{x_{i1}, x_{i2}\}$ for $i = 1, 2$). In state (t_1^*, t_2'), agents' utilities are $u_1(x_1; (t_1^*, t_2')) = x_{11}$ and $u_2(x_2; (t_1^*, t_2')) = x_{21} + x_{22}$. In state (t_1', t_2^*), agents' utilities are $u_1(x_1; (t_1', t_2^*)) = x_{11} + x_{12}$ and $u_2(x_2; (t_1', t_2^*)) = x_{22}$.

The next table reminds us of the endowment and also shows a new allocation, y:

	t_2^*	t_2'
t_1^*	$(2, 0), (0, 2)$	$(2, 0), (0, 1)$
t_1'	$(1, 0), (0, 2)$	

(a) Endowment ω

	t_2^*	t_2'
t_1^*	$(1, 1), (1, 1)$	$(1.5, 0), (0.5, 1)$
t_1'	$(1, 0.5), (0, 1.5)$	

(b) Allocation y

Clearly, y is a coarse objection to ω, so ω is not in the coarse core. However, we claim that ω is in the sequential core. No singleton coalition can have a sequential objection to the initial endowment, so a potential objection must involve a proposal from one agent to the other. Without loss of generality, suppose agent 1 makes a proposal to replace ω with a different allocation. If agent 2 of type t_2^* believes that agent 1 is of type t_1' with probability 1, she cannot possibly gain.

Thus, pessimistic beliefs on the part of the responder make it impossible to construct a sequential objection to ω. In fact, as we know from Proposition 1, there is a PBE in which the outcome is ω. This is one in which both agents announce ω and neither one proposes another allocation. This is supported by off-path beliefs in which an uninformed agent has pessimistic beliefs. In particular, if agent 1 of type t_1^* proposes a different allocation, agent 2 of type t_2^* believes that agent 1 is of type t_1' with probability 1 and therefore rejects the proposal.

We remark that the initial endowment allocation, ω, while being in the sequential core, is not interim efficient. We thus find that the same economic institution –our mechanism– that yields the core under complete information is not capable of eliminating inefficiencies if information is asymmetric (see also Example 4 on this point). ◇

Example 6 is also relevant for the type-agent core proposed by de Clippel (2007). This concept is defined as the subgame-perfect equilibrium outcomes of a competitive screening model in the style of Rothschild-Stiglitz. In the game, at least two uninformed intermediaries make offers to the agents, trying to screen out different types.

It is shown in de Clippel (2007) that the type-agent core is a subset of the coarse core. Therefore, Example 6 above shows that there might be allocations in the sequential core that are not in the type-agent core. In general, the two sets are nested, i.e., the type-agent core is always a subset of the weak sequential core. This follows from the observation that if there exists a strong sequential objection, one of the intermediaries in the de Clippel model can deviate by proposing that objection and reaping some of its benefits, so that the initial allocation is not a subgame-perfect equilibrium outcome of the competitive screening game.

One could argue that in Example 6 it seems odd that the agents would resort to using such pessimistic beliefs off the equilibrium path. What would happen if such pessimism is abandoned? For instance, one possibility is to impose the restriction that off-path beliefs continue to be the prior beliefs. Then, we can state the following result:

Proposition **4.** *Suppose x is a status-quo PBE outcome of the mechanism in which agents hold their prior beliefs off the equilibrium path. Then, x belongs to the coarse core.*

Proof. Let x be a status-quo PBE outcome supported by off-path beliefs equal to the prior. (As we have argued above, in equilibrium, all agents must make a common announcement $x^* = x$ at Stage 0).

Suppose that the status-quo PBE outcome $x^* = x$ is not in the coarse core. Then, there must exist a coarse objection (S, y) over a common knowledge event E. Let one of the agents $i \in S$ of type $t_i \in E_i$ deviate to become the first in line and propose (S, y). Since E is a common knowledge event, given $t_i \in E_i$ there is zero probability that $t_j \notin E_j$ for any responder j. With the specified off-path beliefs, this deviation will be accepted by all responders j, and the deviating agent would increase her interim payoff, contradicting that the specified strategies were part of a PBE. □

6.2 Forward Induction and the Fine Core

An alternative idea is to make use of forward induction to further constrain off-path beliefs. The idea is that, when an alternative to the status-quo is suggested, agents should make use of the proposal itself to rule out types of other agents who would not gain compared to the status-quo. The prior beliefs can then be updated using Bayes' rule over the types that cannot be ruled out.

Suppose (S, y) is an off-path proposal following a status-quo x^* reached in Stage 0. Off-path beliefs β_j are said to satisfy *credible updating* if whenever there is a unique event $E = \prod_{i \in S} E_i \times T_{-S}$ such that $U_i(y|t_i, E) > U_i(x^*|t_i, E)$ for every $t_i \in E_i$ for every $i \in S$; and $U_i(y|t_i, E) \leq U_i(x^*|t_i, E)$ for every $t_i \notin E_i$ for every $i \in S$; then $\beta_j(\iota_{js})$ at that information set for type $t_j \in E_j$ is derived from Bayes' rule applied to the prior beliefs on E. If no such event E exists, then off-equilibrium beliefs are restricted only by the general PBE conditions.

This is the credible updating rule that Grossman and Perry (1986) use in defining a perfect sequential equilibrium. It is also the basis of the *credible core* of Dutta and Vohra (2005). The intuitive criterion of Cho and Kreps (1987) relies on a similar argument for concentrating on a subset of types $E \subseteq T$, but allows arbitrary beliefs over E. The credible updating rule, on the other hand, makes use of new information to update the prior beliefs by applying Bayes' rule over E. The idea of relying on prior beliefs over an informational event that has not been ruled out is quite natural and has been used in several other papers, e.g., Okada (2012) and Liu (2020).

One can then state the next result, parallel to Proposition 4, but for the fine core:

Proposition **5.** *Suppose x is a status-quo PBE outcome of the mechanism in which agents use the credible updating rule off the equilibrium path. Then, x belongs to the fine core.*

Proof. Let x be a status-quo PBE outcome supported by off-path beliefs that use the credible updating rule. (As we have argued above, in equilibrium, all agents must make a common announcement $x^* = x$ at Stage 0).

Suppose that the status-quo PBE outcome $x^* = x$ is not in the fine core. Then, there must exist a fine objection to x. Let (S, y') be a fine objection over a minimal admissible event E in the sense that there does not exist another fine objection over a smaller admissible event $E' \subset E$.

For every t let $z \in A_S$ be a feasible allocation for S with the additional property that for all $i \in$S and $t \in T$,

$$u_i(z_i(t), t) < u_i(x_i(t), t). \tag{8}$$

The existence of such an allocation is assured by the assumption that the utility functions are continuous and strictly monotonic and $x_i(t) > 0$ for all t.

Now consider the fine objection (S, y) over E where

$$y(t) = \begin{cases} y'(t) & \text{if } t \in E \\ z(t) & \text{otherwise} \end{cases}$$

Let $i \in S$ of type $t_i \in E_i$ deviate to become the proposer and propose (S, y). Since (S, y) is a fine objection, we know that

$$U_i(y|t_i, E) > U_i(x|t_i, E) \text{ for every } t_i \in E_i \text{ for every } i \in S. \tag{9}$$

From (8) it follows that

$$U_i(y|t_i, E) < U_i(x|t_i, E) \text{ for every } t_i \notin E_i \text{ for every } i \in S. \tag{10}$$

Condition (10) implies that there cannot be an admissible event E' with $E'_j - E_j \neq \emptyset$ for some $j \in S$ such that (S, y) is a fine objection over E'. By construction, there is not a smaller event over which (S, y) is a fine objection. Thus E is a unique admissible event over which (9) and (10) hold. Since the PBE applies the credible updating rule, responders' off-path beliefs following the deviation (S, y) are derived from Bayes' rule applied to the prior beliefs on E. Clearly, every responder of type $t_j \in E_j$ will accept the proposal and all types not in E_j reject it. Since the proposer also gains, she has a profitable deviation, and this contradicts the hypothesis that x is a status-quo PBE outcome. □

REMARK **3.** *Proposition 5, and its use of the credible updating rule, is related to other results on the fine core. When types are ex post verifiable, the fine core is also the credible core of Dutta and Vohra (2005) and the set of resilient allocations of SV (2007).*

It does not appear to be possible to show that the fine core is contained in the sequential core. While we have not been able to construct a counterexample to settle this question, we come very close by proving that the fine core is not contained in a robust version of the sequential core, as the next example illustrates:

EXAMPLE **7.** *There are two agents. Each agent has two equally-likely types. Each agents owns 1 unit of the commodity in each state and $u_i(x) = x$ for each i. The endowment is clearly in the fine core. Consider the allocation y shown in the following table*

	t_2^*	t_2'
t_1^*	$(1, 1)$	$(1, 1)$
t_1'	$(1, 1)$	$(1, 1)$

(a) Endowment ω

	t_2^*	t_2'
t_1^*	$(0, 2)$	$(2, 0)$
t_1'	$(2, 0)$	$(0, 2)$

(b) Allocation y

Consider the coalition $S = \{1, 2\}$ and observe that it almost has a sequential objection, in the sense that for almost *every belief β_S, the inequalities*

(3) and (4) can be met. Specifically, if $\beta_S(t_1^*) > 0.5$*, let* $E_2(\beta_S) = \{t_2^*\}$ *and* $t_1(\beta_S) = t_1'$*; and if* $\beta_S(t_1^*) < 0.5$*, let* $E_2(\beta_S) = \{t_2'\}$ *and* $t_1(\beta_s) = t_1^*$*.*

Note that the reason this objection is not a sequential objection is that it does not work only *for* $\beta_S(t_1^*) = 0.5$*. That is, the endowment is still in the sequential core, but "not in a robust way".*

6.3 The Signaling Core

Okada (2012) considers stationary sequential equilibria of a bargaining game to identify core stability with incomplete information. The bargaining game includes a rule, as in Okada and Winter (2002), that restarts the game if successive proposals are rejected a certain number of times. He shows that the equilibria of this bargaining game have a close connection to a notion of the core based on defining objections that allow a proposer to signal her private information.

Given $x \in A$, coalition S is said to have a *signaling objection*, (S, y), to x if $y \in A_S$, there exists $i \in S$ and $E_i \subseteq T_i$ such that

$$U_i(y|t_i) > U_i(x|t_i) \text{ if and only if } t_i \in E_i$$

and

$$U_j(y|t_j, E_i) > U_j(x|t_j, E_i) \text{ for all } j \in S,\ j \neq i, t_j \in T_j.$$

Note that all agents in S other than the proposer accept that the proposer's type lies in E_i and, conditional on this, are required to gain for each of their types.

The *signaling core* consists of all allocations in A to which no coalition has a signaling objection.

Since the signaling core is contained in the coarse core, it follows from Example 6 that a sequential core allocation need not be in the signaling core. Our next example shows that an allocation in the signaling core may not be in the sequential core. This is due to the fact that the signaling core does not allow for screening.[21] Thus, there is no logical inclusion relationship between the sequential core and the signaling core.

Example **8.** There are two agents and two commodities.[22] Each agent has two types. Suppose (t_1^*, t_2^*), (t_1^*, t_2') and (t_1', t_2^*) are all equally likely states, and (t_1', t_2') has 0 probability. Agents' utility functions are $u_i(x_i; s) = \min\{x_{i1}, x_{i2}\}$ for all i and all s.

The endowments are shown in the next table, along with another allocation, y.

[21] Example 3 is not adequate for this purpose because it does not consider the possibility of the informed agent becoming the proposer.

[22] This example is similar to one in Corollary 3.1 in Okada (2012).

	t_2^*	t_2'
t_1^*	$(2,0),(0,2)$	$(2,0),(1,1)$
t_1'	$(1,1),(0,2)$	

(a) Endowment ω

	t_2^*	t_2'
t_1^*	$(1,1),(1,1)$	$(2,1),(1,0)$
t_1'	$(0,1),(1,2)$	

(b) Allocation y

We argue next that ω is in the signaling core. Since the endowment is interim individually rational, a signaling objection cannot come from a singleton coalition. It must be from coalition $\{1,2\}$. Suppose there is such an objection with agent 1 as the proposer. If $E_1 = \{t_1'\}$, agent 1 cannot gain because ω provides agent 1 the highest possible utility level of 1. For the same reason, it cannot be the case that $E_1 = \{t_1^*, t_1'\}$. The remaining possibility is that $E_1 = \{t_1^*\}$. In this case, a signaling objection requires that *both* types of agent 2 prefer the proposed allocation to ω. But this is impossible because ω provides agent 2 of type t_2' the highest possible utility level of 1. Since the example is symmetric, the same argument applies if agent 2 is the proposer. We conclude that there is no signaling objection to ω.

However, ω is not in the sequential core. Although y is not a signaling objection to ω, agent 2 of type t_2^* prefers y to ω (no matter what his belief is), while type t_2' does not. Similarly, agent 1 of type t_1^* also prefers y to ω (no matter what his belief is), while t_1' does not. Thus y is a sequential objection to ω by $\{1,2\}$ with $t_1(\beta_{\{1,2\}}) = t_1^*, E_2(\beta_{\{1,2\}}) = \{t_2^*\}$, and ω is not in the sequential core (which coincides with the weak sequential core in this example). It follows from Proposition 3 that ω is not a status-quo PBE outcome. This is also easy to check directly: if the outcome is ω, agent 1 of type t_1^* can deviate profitably by offering y, which will be accepted by agent 2 if and only if he is of type t_2^*. $\diamond$

REMARK 4. *The arguments used above also show that in this example ω is in the individualistic core, thereby demonstrating that the sequential core may be a strict subset of the individualistic core.*

As we saw in Example 2, a PBE or a sequential equilibrium may impose no restrictions on off-path beliefs and may therefore lead to a coordination failure; see also Okada (2012), Example 4.2. For this reason, Okada (2012) imposes a forward induction refinement in order to show that equilibrium outcomes belong to the signaling core.

A sequential equilibrium of Okada's bargaining game satisfies *self-selection* if whenever $E_i = \{t_i \in T_i \mid U_i(y|t_i) > U_i(x|t_i)\} \neq \emptyset$ for a proposer i, the posterior beliefs of the responders are conditioned on E_i.

Okada's (2012) Theorem 4.1 shows that in his general game, if a status-quo x is accepted with probability 1 in a sequential equilibrium satisfying self-selection, then x is in the signaling core. However, it is possible that there is an allocation in the signaling core such that, in a sequential equilibrium satisfying self-selection, it is rejected with positive probability. To ensure that every 'core' allocation is accepted in equilibrium, Okada (2012) refines the signaling core.

Given $x \in A$, coalition S is said to have a *weak signaling objection*, (S, y), to x if $y \in A_S$, there exists $i \in S$ and $E_i \subseteq T_i$ such that

$$U_i(y|t_i) > U_i(x|t_i) \text{ if and only if } t_i \in E_i$$

and

$$U_j(y|t_j, E_i) > U_j(x|t_j, E_i) \text{ for all } j \in S,\ j \neq i \text{ and } \textit{some } t_j \in T_j.$$

The *strong signaling core* consists of all allocations to which there does not exist a weak signaling objection.

Okada's (2012) Proposition 4.1 then shows that every allocation in the strong signaling core can be supported as a sequential equilibrium satisfying the self-selection property. While the strong signaling core allows information leakage from the proposer, it does not permit information to flow in the other direction. When some types of a responder accept a proposal and others reject it, the logic of the strong signaling core does not allow the proposer (and possibly other responders) to update their beliefs. Thus, it may be the case that there is a weak signaling objection in which a proposer expects to gain in terms of interim utility (based on prior beliefs), but the responders who accept are only of a certain type. It is then possible that the posterior belief of the proposer, conditional on this information, reveals that her utility is *lower* when she makes the proposal that gets accepted with positive probability. The sequential core, by allowing for information flows in both directions, does not have this issue. The following example illustrates the possible problem when the proposer fails to account for the information revealed by the actions of the responder.

Example **9.** The economy consists of three agents and three commodities. The endowments are

$$\omega_1 = (1, 0, 0),\ \omega_2 = (0, 1, 0),\ \omega_3 = (0, 0, 1).$$

Agent 3 is completely uninformed while $T_1 = \{t_1^*, t_1'\}, T_2 = \{t_2^*, t_2'\}$. Each type is independently drawn and equally likely. The utility functions are as follows:

$$\begin{aligned} u_1(x_1, t_1^*, t_2^*) &= \min\{x_{i1}, x_{i2}\} \\ u_2(x_2, t_1^*, t_2') &= \min\{x_{i1}, x_{i2}\} \\ u_i(x_i, t) &= \min\{x_{i1}, x_{i2}, x_{i3}\} \text{ for all other cases} \end{aligned}$$

Let x be the allocation in which all commodities are shared equally in each state. The utility of each agent is then $1/3$ in each state. Let $S = \{1, 2\}$ and $y \in A_S$ be the following allocation:

$$y_1(t_1, t_2^*) = (0.9, 0.9, 0),\ y_2(t_1, t_2^*) = (0.1, 0.1, 0) \text{ for all } t_1$$
$$y_1(t_1, t_2') = (0.1, 0.1, 0),\ y_2(t_1, t_2') = (0.9, 0.9, 0) \text{ for all } t_1$$

The utility to agents 1 and 2 corresponding to x and y are shown in the next table.

	t_2^*	t_2'
t_1^*	$1/3, 1/3$	$1/3, 1/3$
t_1'	$1/3, 1/3$	$1/3, 1/3$

(a) Utilities to 1 and 2 from x

	t_2^*	t_2'
t_1^*	$0.9, 0$	$0, 0.9$
t_1'	$0, 0$	$0, 0$

(b) Utilities to 1 and 2 from y

It is easy to see that y is a weak signaling objection to x. Agent 1 of type t_1^* gets higher expected utility from y compared to x. And given that y is proposed by t_1^*, only type t_2' of agent 2 will accept. However, this being the case, agent 1 should realize that he will actually do *worse* than x when y is accepted by agent 2. ◇

The next example is a variation on the previous one to show the converse problem; there may be a strong signaling core allocation that can only be supported with unreasonable beliefs.

EXAMPLE **10.** The endowments and information structure are the same as in the previous Example. The utility functions are as follows:

$$u_i(x_i, t_1^*, t_2^*) = \min\{x_{i1}, x_{i2}\} \text{ for all } i$$
$$u_i(x_i, s) = \min\{x_{i1}, x_{i2}, x_{i3}\} \text{ for } i \text{ and for all } s \neq (t_1^*, t_2^*)$$

As in the previous example, let x be the allocation in which all commodities are shared equally in each state. Let $S = \{1, 2\}$ and $y \in A_S$ be the allocation in which agents 1 and 2 share their endowments equally among themselves. This results in agents 1 and 2 receiving utility (0.5, 0.5) in state (t_1^*, t_2^*) and (0,0) in every other state. (Agent 3 gets 0 in every state). Suppose agent 1 proposes y to agent 2. Note that $U_1(x|t_1^*) = 1/3$ while $U_1(y|t_1^*) = 0.25$. Similarly, $U_1(x|t_1') > U_1(y|t_1')$. So y is not a weak signaling objection and it can be checked that x belongs to the strong signaling core.[23] So, there is a sequential equilibrium with self-selection in which agent 1 accepts x and does not propose y. Let us examine the off-path beliefs in this equilibrium.

What should agent 2 believe if agent 1 actually were to propose y? Although both types of agent 1 prefer x to y in terms of interim utility, there is a difference between the two types. Agent 1 of type t'_1 can never do better by proposing y rather than accepting x; by proposing y, he gets payoff 0 if agent 2 accepts and 1/3 if he rejects while the status-quo guarantees 1/3. In other words, for type t'_1, a deviation to y is equilibrium dominated. But type t^*_1 could conceivably gain. In fact, the only reasonable posterior for agent 2 is that when y is proposed, agent 1 must be of type t^*_1. But then, agent 2 will accept y if and only if she is of type t^*_2. This suggests that the only equilibrium satisfying *forward induction* is one in which type t^*_1 proposes y and agent 2 accepts the proposal only in state t^*.

This is not a signaling game, so we cannot directly apply the intuitive criterion but this seems to conform with the Cho's (1987) definition of a forward induction equilibrium. It seems reasonable to conjecture that, as in our mechanism, also in Okada's (2012) model, forward induction yields the fine core, although we have not been able to establish this result. ◇

6.4 The Ex-Post Nonverifiable Case

In the absence of ex-post verifiability of types, agents may be tempted to misrepresent their private information and incentive compatibility must be imposed explicitly. In this case, the full implementation problem becomes more challenging. Indeed, adding type reports as part of Stage 0, it is possible that there might be equilibria of the mechanism where some deception is used, i.e., nontruthful reports of types. Let $\alpha_i : T_i \mapsto T_i$ and let $\alpha = (\alpha_i)_{i \in N}$ be a deception. In Stage 0, we might have that $\alpha_i(t_i) \neq t_i$ is part of a PBE with a unanimous announcement of the allocation x^*. Then, if the status-quo is accepted by all, the outcome would be $x^* \circ \alpha$, where in each state t, the allocation of goods is $x^*(\alpha(t))$.

[23]Indeed, it is sufficient to show that there is no weak signaling objection to x by coalition $S = \{1, 2\}$. Suppose that $(\{1, 2\}, z)$ is a weak signaling objection to x. First, we note that in all states $t \neq t^*, t'$, both agents have zero utilities no matter how they share the commodities. Consider the case that agent 1 is the proposer and $E_1 = \{t^*_1\}$. Then one of the conditions in the weak signaling objection is $u_1(z, t^*_1, t^*_2) + u_1(z, t^*_1, t'_2) > 2/3$, and then $u_1(z, t^*_1, t^*_2)$ is greater than 2/3. Then, $u_2(z, t^*_1, t^*_2)$ would be less than or equal to 1/3. So there is no $t_2 \in T_2$ satisfying $U_2(z|t_2, E_1) > U_2(x|t_2, E_1)$. By taking similar steps, we can see that it is impossible to find an event E_i satisfying the conditions of weak signaling objections.

One would therefore need to restrict attention to those PBE with truthful reports in stage 0. Adding that restriction, the rest of the analysis would be similar to the one we have conducted here and the set of status-quo PBE outcomes in this class corresponds to the sequential core and weak sequential core. One can generate examples similar to the ones already presented to make the point, for instance, that there may be status-quo PBE outcomes of the mechanism in this class that are not even in the incentive-compatible coarse core of Vohra (1999), let alone any of its nested subsets, identified in the literature after different information transmission processes (Dutta and Vohra (2005), Myerson (2007), SV (2007), Kamishiro and Serrano (2011)).

7 Robustness to Changes in the Mechanism

7.1 Integer Games

One may wish to dispense with the integer game and also eliminate the announcement of the protocol in Stage 0 of our mechanism. Retaining all its remaining rules and fixing exogenously a protocol π, call the resulting mechanism Γ^π. One can then obtain a message equivalent to Theorem 1 by stating the following results: (i) Every allocation in the sequential core is a status-quo PBE outcome of Γ^π for all π; (ii) if an allocation x is a status-quo PBE outcome in every Γ^π, then x must be in the weak sequential core; and (iii) if effective coalitions consist of at most two agents, the sequential core coincides with the set of status-quo PBE outcomes of Γ^π for all π.

7.2 Infinite-Horizon Coalitional Bargaining and Stationary Strategies

As an important robustness check for our results, in this subsection we replace the simultaneous moves of Stage 0 with an infinite horizon model of coalitional bargaining and show that the sequential core continues to have the kind of connection to equilibrium outcomes shown in Theorem 1.

Even with complete information, a standard model of coalitional bargaining with discounting does not generally yield a coincidence between the core and stationary equilibrium outcomes; see for example Chatterjee *et al.* (1993). Additional assumptions or modifications to the game that do implement the core include the continuous time model of Perry and Reny (1994), order-independent equilibria in a model without discounting as in Moldovanu and Winter (1995), or a game with a restarting rule after a fixed number of rejections as in Okada and Winter (2002).[24] We follow the Moldovanu-Winter approach in focusing on order independent equilibria of a bargaining model similar to one used by them.[25]

The rules of the infinite-horizon coalitional bargaining game are as follows. Fix a protocol π, an ordering of the agents. Let $1(\pi)$ refer to the first agent according to π.

At time 0, agent $1(\pi)$ makes a proposal (S, y) to a coalition that includes her as a member, with $y \in A_S$. Then, the rest of the agents in S respond sequentially to the proposal, according to the protocol π restricted to S. If all of them accept the proposal, it is implemented for S, while agents not in S receive their endowments. The game ends, with utilities $u_j(y_j, t)$ in state t for every $j \in S$, and $u_j(\omega_j, t)$ for every $j \notin S$.[26] If one of the members of S rejects the proposal, the game moves to period 1, and the first rejector of the proposal makes a fresh proposal. Again, the game ends if there is unanimous acceptance. Otherwise it moves to the next time period with the first rejector becoming the new proposal, and so on. Perpetual disagreement results in a negative payoff to each agent in each state, which for every i is worse than any $x \in X_i$.

We consider stationary PBE of this bargaining game. Stationarity refers to PBE with strategies and beliefs that are independent of calendar time, i.e., they prescribe changes only after deviations within a given period (a proposal different from the one expected in equilibrium, or responses different from the ones expected in equilibrium).

Then, relying essentially on the same proofs as those in Propositions 2 and 3, one can show the following result:

Theorem 2.

(i) If x is in the sequential core, x can be supported as a stationary PBE outcome of the game for any protocol π.

(ii) If x can be supported as a stationary PBE outcome of the game for any protocol π, then x is in the weak sequential core.

(iii) If effective coalitions consist of at most two agents, the sequential core coincides with the set of stationary PBE outcomes of the game for any protocol π.

[24] Okada (2021) studies this issue in a many-to-one matching model.

[25] The main difference is that in our model, as in Selten (1981), the game ends as soon as one coalition is formed.

[26] If the proposal is made to a singleton coalition – the proposer does not invite anyone – the game ends and everyone receives their endowments.

Proof. (i) The strategies and beliefs would be a small variant, at each time period τ, of those written down in the proof of Proposition 2. Specifically:

In every period, every agent proposes (N, x), which is accepted by all, so the outcome is an immediate agreement on x. Suppose now that, at time τ, the proposing agent i deviates by making a proposal $(S, y) \neq (N, x)$. For any such deviation, the off-path period-τ beliefs and actions taken by each agent j are described next.

Since x is in the sequential core, given (S, y, π), there exists β_S, consistent with (S, π), for which there does not exist an event $E(\beta_S)$ such that (3) and (4) in the definition of a sequential objection are satisfied. Let the beliefs of the responders be given by any such β_S –these beliefs correspond to a single period-τ deviation by agent i as described, while previous responders within period τ are following the equilibrium strategies in the continuation game following i's deviation. Let the responders' strategies in such a period-τ continuation be defined as follows:

(a) If there exists a nonempty set of plausible types $E_j(\beta_S)$ for each $j \in S - i$ such that (4) in the definition of a sequential objection is satisfied, then each $j \in S - i$ accepts if and only if $t_j \in E_j(\beta_S)$.

(b) If there does not exist a set of plausible types, $E_j(\beta_S)$ for some $j \in S-i$ for whom condition (4) in the definition of a sequential objection holds, then each responder of every type rejects the proposal.

The proof that these strategy-belief configuration is a PBE follows from the fact that the stationarity –following a rejection, x is expected to be the outcome in the next period– implies that the relevant inequalities in the equilibrium become those in the definition of a sequential objection.

(ii) Suppose not. If x is not in the weak sequential core, for the fixed arbitrary protocol π, there exists a continuation game in which agent i whose type t_i features in the definition of a strong sequential objection is first in line. Suppose then that there is a stationary PBE of the game with that protocol in which the outcome is x. Let (S, y) be a strong sequential objection to x by agent i of type t_i. Given the PBE, consider the deviation in which agent i makes this proposal when it is her turn to offer. Denote by $E(\beta^i)$ the event over which the objection is accepted, where β^i is the off-path belief in the equilibrium after i's deviation proposal. Whatever beliefs the responders have about each other following this period-τ deviation, by sequential rationality, each type consistent with the event $E(\beta^i)$ will accept the proposal, since they uniformly prefer y over x given β^i, while types outside of the event will reject it, since they uniformly prefer x over y given β^i. Note how stationarity implies that the outcome of the rejection would be x in the next period. Hence, this is a profitable deviation, which is a contradiction.

Again, the proof is essentially the same as that in Proposition 3, given that the deviation inequalities are those in the strong sequential objection.

(iii) It follows from parts (i) and (ii), and from Remark 2. □

8 Conclusion

Mechanisms in the Nash program can be used as launching platforms to explore how a solution concept can be extended to larger domains.[27] This paper executes this agenda by using the mechanism in SV (1997). That mechanism implements the core of complete-information economies in subgame perfect equilibria, and the current paper comes close to characterizing the set of status-quo PBE outcomes of its extension to interim economies with incomplete information. Furthermore, the findings are confirmed in an infinite-horizon model of coalitional bargaining for stationary PBE outcomes.

The set of status-quo PBE outcomes of the mechanism is sandwiched between the sequential core and the weak sequential core (both coincide when at most two-agent coalitions are effective, as in the case of matching). These core notions are based on the sequentiality of moves and rely on objections that incorporate signaling –the proposer may signal some of her private information– and screening –the objection may be targeted to a strict subset of types of the responders. There is no inclusion relationship between the sequential core and Wilson's coarse core, Wilson's fine core, or Okada's signaling core, and several examples have been provided to explain these facts. On the other hand, the sequential core contains de Clippel's type-agent core, where the added outcomes crop up due to the large multiplicity of beliefs compatible with the PBE notion. In addition, this fact also explains why interim inefficient allocations may be contained in the sequential core.

Our aim in this paper has been to study the core of an economy at the interim stage, incorporating two-way information transmission among asymmetrically informed agents. We have done this by relying on a simple mechanism having only one round of proposal/responses. It will be interesting to explore the connection between the sequential core and the equilibria of mechanisms that involve more complex information revelation schemes based on multi-stage procedures.

[27]See Serrano (2021, Section 2) for an appraisal of this point. For instance, Hart and Mas-Colell (1996) and Serrano (1997) do this in extending solutions in the domain of transferable-utility games to games with nontransferable utility, and Maskin, de Clippel, and Serrano (2021) extends a solution from games in characteristic-function form to partition functions.

References

Chatterjee, K., Dutta, B., Ray, D., Sengupta, K., (1993). "A Non-Cooperative Theory of Coalitional Bargaining," *Review of Economic Studies*, **60**, 463–477.

Cho, I., (1987). "A Refinement of Sequential Equilibrium," *Econometrica*, **55**, 1367–1389.

Cho, I., Kreps, D., (1987). "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, **102**, 179–221.

de Clippel, G., (2007). "The Type-Agent Core of Exchange Economies with Asymmetric Information," *Journal of Economic Theory*, **135**, 144–158.

de Clippel, G., Minelli, E., (2002). "Two-Person Bargaining with Verifiable Information," *Journal of Mathematical Economics*, **40**, 799–813.

Dutta, B., Vohra, R., (2005). "Incomplete Information, Credibility and the Core," *Mathematical Social Sciences*, **50**, 148–165.

Fernandez, M. A., Rudov, K., Yariv, L., (2022). "Centralized Matching with Incomplete Information," *American Economic Review: Insights*, **4**, 18–33.

Forges, F., (1994). "Posterior Efficiency," *Games and Economic Behavior*, **6**, 238–261.

Forges, F., (2004). "The Ex Ante Incentive Compatible Core of the Assignment Game", *Mathematical Social Sciences*, **47**, 135–151.

Forges, F., Minelli, E., Vohra, R., (2002). "Incentive and the Core of an Exchange Economy: a Survey," *Journal of Mathematical Economics*, **38**, 1–41.

Forges, F., Serrano, R., (2013). "Cooperative Games with Incomplete Information: Some Open Problems," *International Game Theory Review*, **15**, 301–317.

Fudenberg, D., Tirole, J., (1991). *Game Theory*, M.I.T. Press.

Gale., D., (1978). "The Core of a Monetary Economy without Trust," *Journal of Economic Theory*, **19**, 456–491.

Grossman, S., Perry, M., (1986). "Perfect Sequential Equilibrium," *Journal of Economic Theory*, **39**, 97–119.

Habis, H., Herings, P. J-J., (2011). "Transferable Utility Games with Uncertainty," *Journal of Economic Theory*, **146**, 2126–2139.

Hart, S., Mas-Colell, A., (1996). "Bargaining and Value," *Econometrica*, **64**, 357–380.

Holmström, B., Myerson, R., (1983). "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, **51**, 1799–1819.

Kamishiro, Y., Serrano, R., (2011). "Equilibrium Blocking in Large Quasilinear Economies," *Mathematics of Operations Research*, **36**, 552–567.

Lee, D., (1998). "A Note on the Individualistic Foundations of the Core in Economies with Asymmetric Information," *Economics Letters*, **61**, 203–208.

Lee, D., Volij, O., (2002). "The Core of Economies with Asymmetric Information: An Axiomatic Approach," *Journal of Mathematical Economics*, **38**, 43–63.

Liu, Q., (2020). "Stability and Bayesian Consistency in Two-Sided Markets," *American Economic Review*, **110**, 2625–2666.

Liu, Q., (2022). "A Theory of Coalitional Games with Incomplete Information", Working Paper.

Liu, Q., Mailath, G., Postlewaite, A., Samuelson, L, (2014). "Stable Matching with Incomplete Information," *Econometrica*, **82**, 541–587.

Maskin, E., de Clippel, G., Serrano, R. (2021). "Sequential Bargaining, Coalitions, and Externalities," Mimeo.

Moldovanu, B., Winter, E., (1995). "Order Independent Equilibria," *Games and Economic Behavior*, **9**, 21–34.

Myerson, R., (2007), "Virtual Utility and the Core for Games with Incomplete Information," *Journal of Economic Theory*, **136**, 260–285.

Okada, A., (2012). "Non-Cooperative Bargaining and the Incomplete Informational Core," *Journal of Economic Theory*, **147**, 1165–1190.

Okada, A., (2021). "Stable Matching and Protocol-Free Equilibrium," *Games and Economic Behavior*, **128**, 193–201.

Okada, A., Winter, E. (2002). "A Noncooperative Axiomatization of the Core," *Theory and Decision*, **53**, 1–28.

Perry, M., Reny, P., (1994). "A Non-Cooperative View of Coalition Formation and the Core," *Econometrica*, **62**, 795–817.

Selten, R. (1981). "A Noncooperative Model of Characteristic Function Bargaining," in Bohm, V. and Nachtkamp, H. (eds.), *Essays in Game Theory and Mathematical Economics in Honour of Oskar Morgenstern* (Mannheim: Bibliographisches Institut).

Serrano, R., (1997). "Reinterpreting the Kernel," *Journal of Economic Theory*, **77**, 58–80.

Serrano, R. (2021). "Sixty-Seven Years of the Nash Program: Time for Retirement?" *SERIEs, Journal of the Spanish Economic Association* **12**, 35–48.

Serrano, R., Vohra, R., (1997). "Non-Cooperative Implementation of the Core," *Social Choice and Welfare*, **14**, 513–525.

Serrano, R., Vohra, R., (2007). "Information Transmission in Coalitional Voting Games," *Journal of Economic Theory*, **134**, 117–137.

Vohra, R., (1999). "Incomplete Information, Incentive Compatibility and the Core," *Journal of Economic Theory*, **86**, 123–147.

Wilson, R., (1978). “Information, Efficiency and the Core of an Economy,” *Econometrica*, **46**, 807–816.