Implementation with Evidence

Navin Kartik, Olivier Tercieux

Theoretical Economics 7 (2012), 323–355 1555-7561/20120323

Implementation with evidence

Navin Kartik
Department of Economics, Columbia University

Olivier Tercieux
Paris School of Economics

We generalize the canonical problem of Nash implementation by allowing agents to voluntarily provide discriminatory signals, i.e., *evidence*. Evidence can either take the form of hard information or, more generally, have differential but non-prohibitive costs in different states. In such environments, social choice functions that are not Maskin-monotonic can be implemented. We formulate a more general property, *evidence monotonicity*, and show that this is a necessary condition for implementation. Evidence monotonicity is also sufficient for implementation in economic environments. In some settings, such as when agents have small preferences for honesty, any social choice function is evidence-monotonic. Additional characterizations are obtained for hard evidence. We discuss the relationship between the implementation problem where evidence provision is voluntary and a hypothetical problem where evidence can be chosen by the planner as part of an extended outcome space.

Keywords. Mechanism design, costly signaling, verifiable information, Nash implementation.

JEL classification. C72, D02, D71.

1. Introduction

A classic issue in mechanism design is that of (full) Nash implementation. The goal is to design a mechanism such that in every state of the world, every Nash equilibrium outcome of the game induced by the mechanism is desirable. A maintained assumption in almost all of the literature following Maskin (1999; circulated in 1977) is that agents can manipulate their information without restraint. Specifically, the set of messages that is

Navin Kartik: nkartik@gmail.com
Olivier Tercieux: tercieux@pse.ens.fr
For helpful comments, we thank David Ahn, Vince Crawford, Geoffroy de Clippel, Eddie Dekel, Bob Evans, Jayant Vivek Ganguli, Vijay Krishna, Laurent Lamy, Eric Maskin, François Maniquet, Stephen Morris, Daijiro Okada, Mike Riordan, Hamid Sabourian, Roberto Serrano, Joel Sobel, Joel Watson, and a number of conference and seminar audiences. We also thank anonymous referees and the co-editor, Bart Lipman, for many useful suggestions. This research began when both authors were members of the Institute for Advanced Study at Princeton. We gratefully acknowledge the Institute's hospitality and financial support through the Roger W. Ferguson, Jr. and Annette L. Nazareth membership (Kartik) and the Deutsche Bank membership (Tercieux). Kartik is also grateful to the Alfred P. Sloan Foundation for financial support.

DOI: 10.3982/TE723

available to an agent in a given mechanism is assumed to be state-independent; furthermore, all messages are assumed to be costless. In this sense, all messages are "cheap talk": they do not affect an agent's payoff directly and matter only indirectly insofar as they affect the outcome chosen by the mechanism.

Our goal is to focus attention on why this aspect of the implementation problem is restrictive, and to generalize the set of environments to which the theory can be applied. To motivate our treatment, here are three examples.

1. A principal wishes to divide a fixed sum of money between agents as some function of their individual output. If asked only to send cheap-talk messages about their output, agents could claim anything they want. But agents may also be able to provide physical verification or some other kind of certification of their output. An agent would be unable to certify that his output is greater than it in fact is, but he could certify less, for example by simply not furnishing all of it. If it is costless to provide such certification, the setting is one of hard or verifiable information. If instead agents bear costs as a function of how much output they carry to the principal's court (so to speak), but not how much they actually produced, then we have a costly signaling instrument that combines hard information with "burning money." If the cost of certification also depends on how much they actually produced, then a more complex signaling instrument is at hand.

2. A principal wants to hire the agent who has the highest ability and pay him a wage equal to his marginal product. In addition to sending messages as requested by the mechanism, agents have the choice to voluntarily acquire any amount of education. Education is intrinsically useless, but the marginal cost depends on an agent's ability. This is an implementation version of the classical education-signaling problem (Spence 1973).

3. When asked to report a direct message about the state, some agents may have a (possibly small) degree of aversion to lying: they prefer to send a truthful message about the state if it results in an outcome that is not much worse for them than what could be obtained by lying. The extent of this aversion may be heterogenous across agents.

Common to all these examples is that some messages or actions are only feasible for an agent in some states of the world or have differential costs in different states. As this naturally arises in numerous settings, it is important to study implementation in a framework that accommodates this feature. While the issue has received some attention in the context of partial or weak implementation,[1] it has received almost none in treatments of full implementation, with exceptions that we discuss subsequently.

Accordingly, this paper adds *evidence* to an otherwise standard Nash-implementation environment. (Hereafter, we use the term "implementation" without qualification to mean full implementation in Nash equilibrium.) The defining feature of a piece

[1]This means that one is only concerned with ensuring that *some* equilibrium outcome of the mechanism is desirable, rather than *all* equilibrium outcomes.

of evidence is that it is a *discriminatory signal* about the state of the world, as opposed to a cheap-talk message. A mechanism not only relies on cheap-talk messages as usual, but also on the profile of evidence submitted by the players. Given the ability to commit to a mechanism, a planner cannot do worse when evidence is available than when it is not; our interest is to understand exactly when there is a strict benefit and precisely how much so. In particular, which social objectives are implementable given some evidentiary structure, and what evidentiary structure is needed to make a particular social objective implementable?

In Section 2, we formulate a fairly general problem of complete-information implementation with evidence.[2] Each player i chooses which evidence to provide from some feasible set, E_i. While our formal treatment is broader, assume for this Introduction that at each state θ, player i has preferences that are separable between the outcome chosen by the planner—his outcome preference—and the evidence he submits. We posit that submitting any evidence imposes a nonnegative cost on a player, but the magnitude of the cost can depend on the evidence and the state of the world. Crucially, a player's choice of evidence is *inalienable*: a mechanism cannot force any player to submit any particular piece of evidence. This renders a fundamental distinction between the profile of submitted evidence and the outcome chosen by the planner.

We investigate when a social choice function (SCF) is Nash-implementable in this setting, where the notion of implementation requires that no evidentiary costs be incurred on the equilibrium path.[3] Without evidence, a SCF is implementable only if it is *Maskin-monotonic* with respect to players' outcome preferences (Maskin 1999). A simple but significant observation is that this is no longer true once evidence is in the picture. Rather, what matters is preferences over the joint space of outcomes and evidence. In Section 3, we identify a necessary condition for implementablity that we call *evidence monotonicity*. This condition is weaker than Maskin monotonicity, and the two concepts coincide if and only if there is no evidence. Our condition can be roughly described by considering a hypothetical problem where instead of choosing an outcome after players voluntarily submit evidence and cheap-talk messages, the planner instead chooses both an outcome and an evidence profile after players submit only cheap-talk messages. Loosely speaking, evidence monotonicity requires that one finds an augmented SCF on this joint outcome-plus-evidence space that uses only costless evidence and is Maskin-monotonic with respect to the players preferences on the joint space.

Viewed in this way, it is fairly intuitive why evidence monotonicity is necessary for implementation. We prove that it is also almost sufficient in the sense that any evidence-monotonic SCF can be implemented when there are three or more players and the environment is *economic*. An economic environment is one where in any state, given any outcome, there are at least two players for whom this outcome is not top-ranked.

[2]By complete information, we mean that the state that is unknown to the planner is common knowledge among the agents, so that a mechanism induces a complete-information game in each state of the world.

[3]While a natural starting point, this is a substantive assumption. If one is willing to incur evidentiary costs at equilibrium, then the scope for implementation generally is greater. Our sufficiency results may be viewed as identifying conditions under which a SCF can be implemented without having to incur costly evidence production at equilibrium.

This sufficiency result is unexpected because the choice of what evidence to furnish is inalienable or voluntary, whereas the intuition described above behind the evidence-monotonicity condition assumes the planner has the power to choose the evidence profile. Section 4 develops a "bridge" between the two problems.

The characterization of implementability in terms of evidence monotonicity has a number of applications. In some problems, such as the benchmark education signaling described earlier, natural SCF's cannot be implemented because they are not evidence-monotonic. Alternatively, under some evidentiary cost structures, *every* SCF becomes evidence-monotonic. A striking case is when at least one player has a small preference for honesty. Formally, this is captured by setting each player's feasible set of evidence to be $E_i = \Theta$, where Θ is the set of all possible states of the world. The assumption is that an honest player suffers an arbitrarily small cost of submitting nontruthful evidence. Our results imply that in such a setting with three or more players, any SCF can be implemented in economic environments. Dutta and Sen (2011) and Matsushima (2008a, 2008b) find related results, focussing specifically on preferences for honesty and with some differences in formalization.

Section 5 specializes our general model to settings of *hard* or *nonmanipulable evidence*: it is prohibitively costly for an agent to produce evidence that he does not in fact possess. Formally, in each state θ, each agent i has a set of evidence, $E_i^\ell(\theta) \subseteq E_i$, such that he can costlessly submit any $e_i \in E_i^\ell(\theta)$, but incurs such a large cost of submitting any $e_i \notin E_i^\ell(\theta)$ that the latter is strictly dominated. As we place no restriction on the evidence structure, $\{E_i^\ell(\theta)\}$, the standard environment without evidence is a special case where for all θ, $E_i^\ell(\theta) = E_i$ for any player i. We deduce the implications of evidence monotonicity in this setting of hard evidence. Of particular interest, we find that (i) when there are no outcome-preference reversals to exploit, evidence needs to distinguish in an appropriate sense not only particular pairs of states, but moreover certain states from other sets of states or events, and (ii) some ability to reward agents for providing evidence is necessary. Further insights are developed for the subclass of hard-evidence problems that satisfy *normality* or *full reports* (Bull and Watson 2007, Lipman and Seppi 1995), which can be interpreted as a "no time constraints" assumption on the provision of evidence.

Before turning to a discussion of related literature, let us address one potential concern that some readers may have: why study Nash implementation in a setting with evidence when earlier work has already shown that quite permissive results can be obtained without evidence either by using refinements of Nash equilibrium (e.g., Moore and Repullo 1988, Jackson et al. 1994) or focussing on approximate or virtual implementation (Abreu and Sen 1991, Matsushima 1988)? There are at least three reasons. First, our motivation is not merely to broaden the scope of what is implementable, but rather to understand the role that evidence can play in implementation by studying such environments directly. It is natural to begin with the Nash benchmark, and indeed our necessary conditions identify constraints on how evidence can be used. Second, the aforementioned permissive results without evidence are not without limits,[4] and, third, these results have been questioned from various perspectives.[5]

This paper contributes to the burgeoning literature on mechanism design with evidence.[6] Most of this literature concerns partial implementation with hard evidence. An early reference is Green and Laffont (1986), and a sample of more recent work is Bull and Watson (2004, 2007), Deneckere and Severinov (2008), Glazer and Rubinstein (2004, 2006), Sher (2010), and Singh and Wittman (2001). Bull (2008) and Deneckere and Severinov (2007) study partial implementation with costly evidence production.

In the full-implementation literature, there is a small set of papers that study feasible implementation, where the set of feasible allocations is unknown to the planner. It is typically assumed that the planner can partially verify players' claims in particular ways. For example, in a Walrasian economy setting, Hurwicz et al. (1995) and Postlewaite and Wettstein (1989) assume that a player can claim to have any subset of his true endowment but not exaggerate; in a taxation problem with unknown incomes, Dagan et al. (1999) make a similar assumption. In our model, the set of allocations is constant and known to the planner; instead, it is the set of messages for players that either varies with or has varying costs with the state.

Closest to our work is a recent paper by Ben-Porath and Lipman (2011), who also tackle complete-information full implementation with evidence. While our results are derived independently, we have benefitted from reading their treatment. The motivations for their work and ours are similar—particularly with respect to advancing the prior literature—but the analytical focus is quite different. The two most important differences are that (i) our paper provides a treatment of a general costly evidence provision setting, whereas they focus entirely on hard evidence, and (ii) we study Nash implementation throughout, whereas they focus on subgame-perfect implementation.[7] Moreover, their main results require that the planner can augment monetary transfers off the equilibrium path; as mentioned earlier, we show that some ability to reward players is in fact necessary to exploit hard evidence.

[4]For instance, none of them has bite when players' outcome preferences do not vary across states, whereas evidence can be extremely useful in this regard.

[5]A well known weakness of virtual implementation is that the mechanism may provide an outcome that is arbitrarily inefficient, unfair, or "far" from the desired outcome, even if this occurs only with small ex ante probability. Implementation with refinements of Nash equilibrium has recently been critiqued in terms of robustness to the introduction of small amounts of incomplete information. In particular, if one requires these mechanisms to implement in environments with "almost" complete information, Maskin monotonicity is again a necessary condition (Chung and Ely 2003, Aghion et al. 2009).

[6]Beyond mechanism design, there are other literatures where evidence plays an important role. The introduction of hard evidence into implementation may be considered to be analogous to moving from communication games of cheap talk (Crawford and Sobel 1982) to those of verifiable information (Grossman 1981, Milgrom 1981). Costly evidence production is studied in communication games by Kartik et al. (2007) and Kartik (2009), in contract settings by, for example, Maggi and Rodriguez-Clare (1995), and in legal settings by, for example, Emons and Fluet (2009).

[7]Ben-Porath and Lipman's Theorem 2, derived contemporaneously with our work, provides sufficient conditions for one-stage subgame-perfect implementation with hard evidence (hence, Nash implementation). Remark 3 in Section 5 provides a detailed comparison.

2. The model

There is a nonempty set of *agents* or players, $I = \{1, \dots, n\}$, a set of allocations or *outcomes*, A, and a set of *states* of the world, Θ. To avoid trivialities, $|A| > 1$ and $|\Theta| > 1$. The state is common knowledge to the agents, but unknown to the planner. The planner's objectives are given by a *social choice function* (SCF), which is a function $f : \Theta \to A$.[8] In any state, agent i can produce a piece of evidence, $e_i \in E_i$, where $E_i \neq \varnothing$ is i's feasible set of evidence. Let $E := E_1 \times \cdots \times E_n$. Throughout, we use a subscript $-i$ to denote all players excluding i, so that, for example, $E_{-i} := \times_{j \neq i} E_j$.

Agents are expected utility maximizers, and an agent i's preferences are represented by a von Neumann–Morgenstern utility function, $U_i : A \times E_i \times \Theta \to \mathbb{R}$. Here, $U_i(a, e_i, \theta)$ is agent i's utility in state θ when the outcome is a and he submits evidence e_i.[9] We assume that utilities are bounded in each state: for all i and θ, $\sup_{a, e_i} U_i(a, e_i, \theta) < \infty$ and $\inf_{a, e_i} U_i(a, e_i, \theta) > -\infty$. We say that preferences are *separable* (between outcomes and evidence) if for all i there is a decomposition $U_i(a, e_i, \theta) = u_i(a, \theta) - c_i(e_i, \theta)$. Under separability, $u_i(a, \theta)$ represents agent i's preferences over outcomes and $c_i(e_i, \theta)$ represents the cost to agent i of evidence provision.

We wish to capture situations in which evidence submission is not intrinsically valued by the agents or the planner. Let $E_i^\ell(\theta, a) := \arg\max_{e_i} U_i(a, e_i, \theta)$ be the set of *least-cost* evidence for a player i given outcome a and state θ. We assume that for each player i, outcome a, and state θ, $E_i^\ell(\theta, a) \neq \varnothing$. Let $E^\ell(\theta, a) := E_1^\ell(\theta, a) \times \cdots \times E_n^\ell(\theta, a)$, so that given outcome a and state θ, any profile of evidence in $E^\ell(\theta, a)$ consists of each player submitting some least-cost evidence. For short, we call $e \in E^\ell(\theta, a)$ a *costless* evidence profile. We say that e_i is *cheap-talk evidence* if $e_i \in \bigcap_\theta \bigcap_a E_i^\ell(\theta, a)$, because such an e_i is a least-cost evidence for i no matter the state or the outcome. If preferences are separable, we write $E_i^\ell(\theta) := \arg\min_{e_i} c_i(e_i, \theta)$.

In standard Nash-implementation theory, a mechanism consists of a (cheap-talk) message space and an outcome function that specifies an outcome for every profile of messages. In the current setting, a mechanism can also take advantage of the evidence that players submit. Formally, a *mechanism* is a pair (M, g), where $M = M_1 \times \cdots \times M_n$ is a *message space* and $g : M \times E \to A$ is an *outcome function* that specifies an outcome for every profile of messages and evidence.

A mechanism (M, g) induces a strategic-form game in each state of the world, θ, where a pure strategy for player i is $(m_i, e_i) \in M_i \times E_i$ and a pure-strategy profile $(m, e) := (m_i, e_i)_{i=1}^n$ yields a payoff $U_i(g(m, e), e_i, \theta)$ to player i. Let $\mathrm{NE}(M, g, \theta)$ be the set of pure-strategy Nash equilibria (NE, hereafter) of the mechanism (M, g) in state θ. For expositional simplicity, we restrict attention to pure-strategy equilibria; our results can be extended to mixed-strategy equilibria as discussed in Remark 1 following Theorem 2.

[8]Our results readily generalize to social choice correspondences at the cost of additional notation.

[9]It is common to focus on just ordinal preferences in each state. This approach suffices insofar as only pure-strategy Nash equilibria are considered, but our formulation allows us to subsume mixed Nash equilibria as well (see Remark 1). In addition, our formulation also allows for the view that utility functions contain cardinal information, and hence, allows for cardinally based social choice functions (such as utilitarianism and egalitarianism).

Since the planner's objective, represented by the SCF, specifies only an outcome for each state, a definition of implementation must take a stance on what profiles of evidence are acceptable to the planner. We adopt the following notion.

Definition 1 (Implementation). A mechanism (M, g) implements the SCF f if

(i) $\forall\theta : f(\theta) = \{a : a = g(m, e)$ for some $(m, e) \in \mathrm{NE}(M, g, \theta)\}$ and

(ii) $(m, e) \in \mathrm{NE}(M, g, \theta) \Longrightarrow e \in E^{\ell}(\theta, f(\theta))$.

A SCF is implementable if there is a mechanism that implements it.

We now comment on a number of aspects of the model.

1. One can think of any e_i as a document, physical object, verbal claim, or action that agent i can submit, provide, or take. What is crucial is that the decision of which evidence to submit is a player's private decision as in Myerson (1982) and cannot be coerced by the planner at any point. In this sense, following Bull and Watson's (2007) terminology, we view evidence as *inalienable*. This renders a fundamental distinction between the outcome space, A, which falls under the planner's purview, and the evidence profile space, E, which does not.

2. The present framework nests the standard model without evidence as a special case: it arises when all evidence for every player is cheap-talk evidence, in which case preferences are separable and one can set $c_i(e_i, \theta) = 0$ for all i, e_i, and θ. Note that the second part of Definition 1 is trivially satisfied in this case, because for any a and θ, $E^{\ell}(\theta, a) = E$. Hence, without evidence, our notion of implementation reduces to the standard notion.

3. More generally, part (ii) of Definition 1 requires that only costless evidence profiles must be sent in any equilibrium of an implementing mechanism. Given our motivation that evidence is not intrinsically valued by the planner or the players, the interpretation is that an implementing mechanism should not lead to any (Pareto) inefficient evidence production. This is a natural benchmark, although not the only reasonable one, as discussed in the conclusion.

4. We assume that a player's preferences depend only on the evidence he provides, but not on the evidence submitted by other players. While this is obviously appropriate in many situations, there may be some applications where it is restrictive, i.e., where a player's evidence submission has a direct externality on other players. We leave such cases to future research.

5. Consider settings with separable preferences. The framework allows for the possibility that two distinct states are identical in terms of all players' preferences over outcomes. In traditional implementation theory, it is common to equate states with profiles of preferences, because in the absence of evidence, it is impossible to implement different outcomes in two states that do not differ in players' (outcome)preferences. We shall see that this is no longer the case once evidence is

available. As emphasized by Ben-Porath and Lipman (2011), in many applications, such as contract and legal settings, a planner may wish to condition the outcome on the state even though players' preferences over outcomes are entirely *state-independent*.[10]

6. Each player's feasible set of evidence, E_i, is assumed to be nonempty. This is without loss of generality, because we can always endow a player who has no evidence with some cheap-talk evidence, since the planner can require submission of a cheap-talk message.

7. It is also without loss of generality that each player must submit exactly one piece of evidence. If one wants to allow a player to submit zero or no evidence, this just requires labeling a particular piece of evidence as "no evidence." If one wants to allow a player to submit multiple pieces of evidence, this just requires adding the appropriate conjunctions of underlying evidence. On the other hand, in particular applications, it may be reasonable that submitting no evidence is either not allowed or at least is not costless, and similarly that submitting multiple pieces of evidence imposes higher (possibly prohibitive) costs; we provide some examples later.

8. Our formulation of a mechanism is inherently static since we are considering the strategic-form game it induces in each state. Given the focus on Nash equilibrium, our results would not change if we were to consider dynamic mechanisms.[11] Furthermore, while our formulation considers only deterministic mechanisms, stochastic mechanisms can be encompassed by viewing A as a lottery space.

9. Finally, we do not allow the planner to prohibit or forbid players from submitting some pieces of evidence. This squares well with the view that a player's choice of evidence is inalienable. In any case, our results would not change even if a planner could forbid some evidence.[12]

10. An important special case of our model is when any piece of evidence is either costless or prohibitively costly.

Definition 2 (Hard evidence). The setting is of hard evidence if preferences are separable and for all i, θ, and e_i, either $c_i(e_i, \theta) = 0$ or $c_i(e_i, \theta) > \sup_a u_i(a, \theta) - \inf_a u_i(a, \theta)$.

In a setting with hard evidence, submitting any $e'_i \in E_i^\ell(\theta)$ strictly dominates submitting any $e_i \notin E_i^\ell(\theta)$ for player i at state θ because the latter's cost strictly outweighs any

[10]Formally, in a separable setting, preferences over outcomes are state-independent if $\forall i \in I$, $\forall \theta, \theta' \in \Theta$, $\forall a, b \in A$, $u_i(a, \theta) \geq u_i(b, \theta) \Longrightarrow u_i(a, \theta') \geq u_i(b, \theta')$.

[11]More precisely, the sufficient conditions we provide for implementation would obviously also remain sufficient; the necessary condition remains necessary as long as one allows only dynamic mechanisms that do not indirectly change the evidence structure, such as by allowing multiple instances of evidence submission or randomizing over what evidence to request. See also footnote 14.

[12]The sufficient conditions for implementation obviously remain sufficient; one can show also that our necessary condition remains necessary. Alternatively, in cases where our sufficient conditions fail, implementation may be possible when the planner can forbid some evidence but not when he cannot; interested readers are referred to earlier versions of this paper for an example.

possible utility gain from inducing a preferred outcome. Thus, by submitting e_i, player i effectively proves that the state is in the set $\{\theta : e_i \in E_i^\ell(\theta)\}$.[13] This justifies why such a setting is one of *hard evidence*, which is also referred to sometimes as certifiability, verifiability, or partial provability.[14] Note that the standard environment without evidence is in fact a special case of a hard-evidence setting, where for any i, $E_i^\ell(\theta) = E_i$ for all θ.

3. General results

Maskin (1999) shows that in a setting without evidence—which can be represented in our framework as a setting where all evidence is cheap-talk evidence and preferences are separable—a SCF must satisfy a monotonicity condition with respect to players' preferences over outcomes to be implementable. We refer to his condition as *Maskin monotonicity*, which can be stated as follows whenever preferences in the current context are separable:

Definition 3 (Maskin monotonicity). Assume separable preferences. A SCF is Maskin-monotonic provided that for all θ and θ', if

$$\forall i, a : \big[u_i(f(\theta), \theta) \geq u_i(a, \theta) \Longrightarrow u_i(f(\theta), \theta') \geq u_i(a, \theta')\big], \tag{1}$$

then $f(\theta) = f(\theta')$.

When evidence is available, Maskin monotonicity is not necessary for implementation, as illustrated starkly in the following example.

Example 1. Assume separable preferences and suppose $E_1 = \Theta$ with

$$c_1(\theta, \theta') = \begin{cases} 0 & \text{if } \theta = \theta' \\ k & \text{if } \theta \neq \theta', \end{cases}$$

where $k > \sup_{a,\theta} u_1(a, \theta) - \inf_{a,\theta} u_1(a, \theta)$. This can be interpreted as player 1 never being willing to misrepresent the state of the world. Trivially then, regardless of the agents' preferences over outcomes, any SCF f can be implemented by a mechanism with an arbitrary message space, M, and outcome function $g(m, (e_1, \dots, e_n)) = f(e_1)$. ◊

[13] While Definition 2 assumes separability, there is essentially no loss of generality. Consider the following definition that does not assume separability: for all i, θ, and e_i, either (a) $e_i \in \bigcap_a E_i^\ell(\theta, a)$ or (b) $\exists e_i'$ such that $\forall a, b : U_i(a, e_i', \theta) > U_i(b, e_i, \theta)$. This definition clearly subsumes Definition 2. Moreover, if preferences satisfy this condition, then it is strictly dominated at state θ for player i to submit any $e_i \notin \bigcap_a E_i^\ell(\theta, a)$. The setting is then effectively identical to one where preferences are separable, and at any state θ, player i has a cost $c_i(e_i, \theta) = 0$ if $e_i \in \bigcap_a E_i^\ell(\theta, a)$ and a cost $c_i(e_i, \theta) > \sup_a u_i(a, \theta) - \inf_a u_i(a, \theta)$ otherwise.

[14] Contrary to our treatment, models of hard evidence often assume that noncostless evidence is actually unavailable, rather than feasible but prohibitively costly to produce. Given our focus in this paper on static mechanisms, the two approaches are equivalent. More generally, however, there are some differences: for example, dynamic mechanisms can sometimes help (even without using randomization and requesting only that each player submit evidence once) when some evidence is infeasible rather than just prohibitively costly to produce. Interested readers should consult previous versions of this paper for details; see also Bull and Watson (2007) for a partial-implementation context.

The key to our analysis is identifying the appropriate notion of monotonicity in the present setting, which we call *evidence monotonicity*.

Definition 4 (Evidence monotonicity). A SCF f is evidence-monotonic if there exists $e^*:\Theta \to E$ such that

(i) for all θ, $e^*(\theta) \in E^\ell(\theta, f(\theta))$ and

(ii) for all θ and θ', if

$$\forall i, a, e_i' : \left[U_i(f(\theta), e_i^*(\theta), \theta) \geq U_i(a, e_i', \theta) \Rightarrow U_i(f(\theta), e_i^*(\theta), \theta') \geq U_i(a, e_i', \theta')\right], \tag{2}$$

then $f(\theta) = f(\theta')$.

In words, a SCF is evidence-monotonic if there is a function e^* that assigns a costless profile of evidence to each state such that if no player has a "preference reversal" with respect to the outcome and his component of e^* when the state changes from θ to θ', then $f(\theta) = f(\theta')$. Intuitively, one should think of $e^*(\cdot)$ as the evidence profile that is submitted to an implementing mechanism. The existential quantifier on $e^*(\cdot)$ is unavoidable: it stems from the fact that the planner does not intrinsically care about which evidence profile is submitted. We show later that in a special but important class of problems, verifying the definition can be simplified. Notice that the second part of Definition 4 bears a resemblance to how one would view Maskin monotonicity on an extended outcome space, $A \times E$, if the planner could somehow choose evidence profiles in addition to choosing outcomes. We clarify the connection in Section 4.

In settings with separable preferences, evidence monotonicity is a weaker requirement than Maskin monotonicity, with the two concepts being equivalent when all evidence is cheap talk. A formal statement follows.

Proposition 1. *Assume separable preferences. (i) Any Maskin-monotonic SCF is evidence-monotonic. (ii) If all evidence is cheap-talk evidence, then any evidence-monotonic SCF is Maskin-monotonic.*

Proof. For the first statement, assume f is Maskin-monotonic. Let e^* be any function such that for all θ, $e^*(\theta) \in E^\ell(\theta, f(\theta))$. Fix any θ and θ', and assume (2). We must show that $f(\theta) = f(\theta')$. Separability and (2) imply that for all i, a, and e_i',

$$u_i(f(\theta), \theta) \geq u_i(a, \theta) - \left(c_i(e_i', \theta) - c_i(e_i^*(\theta), \theta)\right)$$
$$\Rightarrow u_i(f(\theta), \theta') \geq u_i(a, \theta') - \left(c_i(e_i', \theta') - c_i(e_i^*(\theta), \theta')\right).$$

By taking $e_i' = e_i^*(\theta)$ for all i above and applying Maskin monotonicity, it follows that $f(\theta) = f(\theta')$. Thus, f is evidence-monotonic.

For the second statement, notice that when all evidence is cheap-talk evidence and preferences are separable, (2) reduces to (1). □

A striking observation is that even small evidentiary costs can create a substantial wedge between evidence monotonicity and Maskin monotonicity. In particular, this arises when players have small preferences for honesty, as mentioned in the third motivating example of the Introduction. The following generalization of Example 1 demonstrates the point.

Example 2. Consider a setting where players have a small preference for honesty when asked to report a direct message about the state. Formally, assume separable preferences and suppose that for each i, $E_i = \Theta$ and the cost function is given by

$$c_i(\theta, \theta') = \begin{cases} 0 & \text{if } \theta = \theta' \\ \varepsilon & \text{if } \theta \neq \theta', \end{cases}$$

where $\varepsilon > 0$ can be arbitrarily small.

This structure implies that for any i, θ, and a, $E_i^\ell(\theta, a) = \{\theta\}$. Hence, for any a and θ, $E^\ell(\theta, a) = \{(\theta, \dots, \theta)\}$. Let $e^*(\theta) = (\theta, \dots, \theta)$ for all θ. Observe that for any θ and $\theta' \neq \theta$, (2) is false: consider $a = f(\theta)$ and $e_i' = \theta'$. Thus, any SCF is evidence-monotonic.

In fact, it is not necessary that all players have such a preference for honesty, only that in each state there be some player who does (the identity of the player could vary with the state).[15] ◊

Our first main result is that only evidence-monotonic SCFs are implementable.

Theorem 1. *If f is implementable, then f is evidence-monotonic.*

Proof. Assume f is implementable and pick any mechanism (M, g) that implements f. For each θ, there exists $(m(\theta), e(\theta)) \in M \times E^\ell(\theta, f(\theta))$ that is a Nash equilibrium at θ such that $g(m(\theta), e(\theta)) = f(\theta)$. For each θ, set $e^*(\theta) := e(\theta)$. We show that this choice verifies Definition 4. Part (i) of the definition is obviously satisfied, so consider part (ii). Pick any θ and θ', and assume that (2) is satisfied. Let m be such that $(m, e^*(\theta)) \in \mathrm{NE}(M, g, \theta)$ and fix any i. By the optimality of i's strategy,

$$U_i(f(\theta), e_i^*(\theta), \theta) \geq U_i(g(m_i', e_i', m_{-i}, e_{-i}^*(\theta)), e_i', \theta)$$

for any $(m_i', e_i') \in M_i \times E_i$. By (2),

$$U_i(f(\theta), e_i^*(\theta), \theta') \geq U_i(g(m_i', e_i', m_{-i}, e_{-i}^*(\theta)), e_i', \theta')$$

for any $(m_i', e_i') \in M_i \times E_i$.

Consequently, $(m, e^*(\theta)) \in \mathrm{NE}(M, g, \theta')$. Hence, $f(\theta) = g(m, e^*(\theta)) = f(\theta')$, as required. □

To illustrate how Theorem 1 has bite, we return to the second motivating example in the Introduction about education signaling.

[15] It is without loss of generality to assume that the planner knows the identity of the player with preferences for honesty in any state: if the planner has uncertainty about which player it is, this requires only extending the state space.

Example 3. There are n workers, each with an ability level that measures his marginal productivity. The state of the world is a vector of abilities. There is one job that must be allocated to a single worker along with a wage, so that $A = \{1, \ldots, n\} \times \mathbb{R}_+$. The SCF is $f(\theta) = (i^*(\theta), \theta_{i^*(\theta)})$, where $i^*(\theta) = \max(\arg\max_i \theta_i)$, i.e., the goal is to allocate the job to the most able worker and pay him his marginal product (ties in ability are broken in favor of workers with higher indices, which is convenient but inessential). Suppose that workers can signal their ability through a choice e_i of education, so that the evidence for i is $e_i \in \mathbb{R}_+$. Workers' preferences are separable, and any worker i's utility from an outcome $a = (a_1, a_2)$ is state-independent, while his cost of education depends only on his own ability; hence, we can write $U_i(a, e_i, \theta) = u_i(a) - c_i(e_i, \theta_i)$. Assume the cost of education satisfies two reasonable properties: for all i and θ_i, $c_i(e_i, \theta_i) = 0$ if $e_i = 0$; for all i, $c_i(e_i, \theta_i)$ is strictly increasing in e_i and strictly decreasing in θ_i.

Then for any i and θ, $E_i^\ell(\theta) = \{0\}$, and it follows that the only candidate to verify evidence monotonicity according to Definition 4 is $e^*(\theta) = (0, \ldots, 0)$. It is easily checked that (2) is satisfied for any $\theta' \leq \theta$, in the sense of usual vector order.[16] Since $\theta' \leq \theta$ does not imply $f(\theta') = f(\theta)$, f is not evidence-monotonic and, by Theorem 1, is not implementable. ♢

While evidence monotonicity is necessary for implementation, it is not sufficient. Rather than pursuing an exhaustive characterization, we first tackle sufficiency in *economic environments*.

Definition 5 (Economic environment). The environment is economic if there is no state θ, outcome a, and evidence profile e such that $|\{i : (a, e_i) \in \arg\max_{b, e_i'} U_i(b, e_i', \theta)\}| \geq n - 1$.

In other words, an environment is economic if in any state, given any outcome and evidence–profile pair (a, e), there are at least two players for each of whom (a, e_i) is not top-ranked. Various versions of such a domain restriction are used in the implementation literature. Definition 5 is identical to a condition in Bergemann and Morris (2008) that has the same name, as long as one views their condition on an extended outcome space $A \times E$; under separable preferences, it is equivalent to Bergemann and Morris' (2008) condition viewed on the outcome space A alone, because in this case our condition simplifies to requiring that for any θ and a, $|\{i : a \in \arg\max_b u_i(b, \theta)\}| < n - 1$.

To understand the scope of economic environments, focus on settings with three or more agents. An environment is economic if there is a divisible private good that is positively valued by all agents. In particular, the environment is economic if the planner

[16] In this example, using $e^*(\theta) = (0, \ldots, 0)$ for each θ, (2) reduces to

$$\forall i, a, e_i' : \left[u_i(f(\theta)) \geq u_i(a) - c_i(e_i', \theta_i) \Longrightarrow u_i(f(\theta)) \geq u_i(a) - c_i(e_i', \theta_i')\right],$$

which is equivalent to

$$\forall i, a, e_i' : \left[c_i(e_i', \theta_i) \geq u_i(a) - u_i(f(\theta)) \Longrightarrow c_i(e_i', \theta_i') \geq u_i(a) - u_i(f(\theta))\right],$$

which is true for any $\theta' \leq \theta$ because then $c_i(e_i', \theta_i') \geq c_i(e_i', \theta_i)$ for any i, e_i'.

can augment an underlying outcome space with arbitrarily small transfers, even with a requirement of budget balance (cf. Benoît and Ok 2008, Ben-Porath and Lipman 2011, Sanver 2006).[17] Even without private goods or transfers, an environment is economic as long as there is enough disagreement among agents about their most preferred outcome in any state.

Theorem 2. *Assume $n \geq 3$ and an economic environment. If f is evidence-monotonic, then f is implementable.*

The proof of Theorem 2 is by construction of a canonical mechanism that is familiar from existing mechanisms in the literature, but is modified appropriately to deal with evidence.

Proof of Theorem 2. Since f is evidence-monotonic, let e^* be the function that verifies Definition 4. For all i, set $M_i = \Theta \times A \times \mathbb{N}$. Define $g(m, e)$ according to the following rules.

Rule 1. If $m_1 = \cdots = m_n = (\theta, f(\theta), k)$ and $e = e^*(\theta)$, then $g(m, e) = f(\theta)$.

Rule 2. If $\exists i$ such that (i) for all $j \neq i$, $m_j = (\theta, f(\theta), k)$ and $e_j = e_j^*(\theta)$, and (ii) either $m_i = (\tilde{\theta}, a, l) \neq (\theta, f(\theta), k)$ or $e_i \neq e_i^*(\theta)$, then there are two alternatives.

(a) If $U_i(f(\theta), e_i^*(\theta), \theta) \geq U_i(a, e_i, \theta)$, then $g(m, e) = a$.

(b) If $U_i(f(\theta), e_i^*(\theta), \theta) < U_i(a, e_i, \theta)$, then $g(m, e) = f(\theta)$.

Rule 3. For any other (m, e), letting $m_i = (\theta_i, a_i, k_i)$ and $i^* = \min \arg\max_{i \in I} k_i$, then $g(m, e) = a_{i^*}$.

Step 1. It is routine to verify that for any θ, there is a "truthful" NE, where for some $k \in \mathbb{N}$, each agent i plays $m_i = (\theta, f(\theta), k)$ and $e_i = e_i^*(\theta)$. This NE results in outcome $f(\theta)$ and, moreover, e clearly belongs to $E^\ell(\theta, f(\theta))$.

For the remainder of the proof, assume that the true state is θ' and that (m, e) is a NE.

Step 2. We show that (m, e) cannot fall into Rule 2. Suppose, to the contrary, that (m, e) is such an equilibrium. Then it must be that for all $j \neq i$ (where i is defined in Rule 2),

[17]To be more precise, assume $n \geq 3$, assume separable preferences (for simplicity), and consider an underlying outcome space, $\tilde{A}$, with each agent having a utility function $\tilde{u}_i : \tilde{A} \times \Theta \to \mathbb{R}$. Note that $\tilde{A}$ itself may include transfers or private goods, but need not. Fix some SCF $\tilde{f} : \Theta \to \tilde{A}$. Now suppose the planner can impose an additional vector of transfers $(t_1, \ldots, t_n) \in X \subseteq \mathbb{R}^n$ and that each agent values his personal transfer quasilinearly. Assume the space of possible transfers satisfies two mild properties: $(0, \ldots, 0) \in X$ and, for all $(t_1, \ldots, t_n) \in X$, there exists $(\tilde{t}_1, \ldots, \tilde{t}_n) \in X$ and $i \neq j$ such that $\tilde{t}_i > t_i$ and $\tilde{t}_j > t_j$. An obvious example is $X = \{(t_1, \ldots, t_i, \ldots, t_n) \in \mathbb{R}^n : \sum_j t_j = 0, |t_i| \leq k\}$ for some $k > 0$, i.e., the planner must balance his budget and cannot reward or punish any player by more than k utility units. We can then define an augmented outcome space $A = \tilde{A} \times X$, an augmented utility function for each agent $u_i : A \times \Theta \to \mathbb{R}$, where $u_i(\tilde{a}, t_1, \ldots, t_n, \theta) = u_i(\tilde{a}, \theta) + t_i$, and an augmented SCF $f : \Theta \to A$ derived from $\tilde{f}$ by setting $f(\theta) = (\tilde{f}(\theta), 0, \ldots, 0)$. This augmented environment satisfies Definition 5.

$(g(m,e), e_j) \in \arg\max_{b,e'_j} U_j(b, e'_j, \theta')$: otherwise, one of these $n-1$ players, say j^*, can profitably deviate into Rule 3, submitting some e'_{j^*} and requesting and receiving some b such that $U_{j^*}(b, e'_{j^*}, \theta') > U_{j^*}(g(m,e), e_{j^*}, \theta')$. But this contradicts the environment being economic. A similar argument applies to show that no equilibrium (m,e) can fall into Rule 3.

Step 3. It remains to consider the case where (m,e) falls into Rule 1, so that $e = e^*(\theta)$ for some θ. Here, $g(m,e) = f(\theta)$. Rule 2 must hold since a player i can always deviate into Rule 2(a) by producing evidence e'_i and get any outcome a such that $U_i(f(\theta), e^*_i(\theta), \theta) \geq U_i(a, e'_i, \theta)$. Thus, evidence monotonicity implies that $f(\theta) = f(\theta')$. Finally, since any player can deviate to Rule 2(a) and get the same outcome $f(\theta')$ while submitting some evidence in $E^\ell_i(\theta', f(\theta'))$, the hypothesis that (m,e) is a Nash equilibrium implies that $e \in E^\ell(\theta', f(\theta'))$. □

Remark 1. The mechanism used in the proof of Theorem 2 does not work when mixed Nash equilibria are considered. Arguments analogous to those that deal with mixed strategies in standard settings without evidence (e.g., Kartik and Tercieux 2012, Maskin and Sjöström 2002, Section 4.3) can be adapted to extend Theorem 2 to mixed Nash equilibria.

Remark 2. Theorem 2 concerns three or more agents. When there are only two agents, the economic environment condition is rather stringent. For $n = 2$, evidence monotonicity can be shown to be sufficient for implementation under a less demanding version of economic environments in conjunction with Moore and Repullo's (1990) *bad outcome* condition.

To illustrate how Theorem 2 is useful, we apply it to a setting where players have (possibly small) preferences for honesty.

Corollary 1. *Assume $n \geq 3$ and that the environment is economic. If in each state, at least one player has a preference for honesty as formalized in Example 2, then any SCF is implementable.*

Proof. As shown in Example 2, any SCF is evidence-monotonic when at least one player has a preference for honesty. Thus, Corollary 1 is a direct implication of Theorem 2. □

There is a growing literature on implementation when players have preferences for honesty. Matsushima (2008a, 2008b) was the first to investigate such a question and obtain permissive results in related, but not identical, settings. More similar to Corollary 1 is a contemporaneous finding of Dutta and Sen (2011). Because their paper is entirely about implementation with preferences for honesty, their main result is slightly stronger than Corollary 1 and their proof uses a remarkably simple implementing mechanism. Our approach has the benefit of identifying that the fundamental reason why a preference for honesty produces permissive results is that it renders any SCF evidence-monotonic.

We now discuss the role of the economic environment assumption in Theorem 2. In an economic environment, any SCF trivially satisfies the following version of Maskin's (1999) *no veto power* condition.

Definition 6 (No veto power). A SCF f satisfies no veto power provided that for all θ,

$$\text{if } (a, e) \text{ is such that } \left|\left\{i : (a, e_i) \in \arg\max_{a', e'_i} U_i(a', e'_i, \theta)\right\}\right| \geq n - 1, \text{ then } a = f(\theta).$$

In other words, if at any state, there is some outcome and evidence–profile pair that is top-ranked by at least $n - 1$ players, then the outcome must be chosen by the SCF at that state. Plainly, when preferences are separable—a fortiori, when all evidence is cheap-talk evidence—the above definition reduces to the standard no veto power condition. In a setting without evidence, Maskin (1999) shows that no veto power is sufficient to ensure that Maskin-monotonic SCF's are implementable, given $n \geq 3$. This might suggest that Theorem 2 could be strengthened by assuming only no veto power rather than an economic environment. The following counterexample proves otherwise.

Example 4. Suppose $n = 3$, $\Theta = \{X, Y\}$, and $A = \{b, c, d_1, d_2, d_3\}$. Only player 1 has non-cheap-talk evidence, so we ignore the evidence of player 2 and 3. Let $E_1 = E = \{x, y\}$. All players have separable preferences. Using the standard notation of $\succ$ for strict preference, player 3's preferences over outcomes are given by

$$\begin{array}{ccl} X & : & b \prec c \prec d_1 \prec d_2 \prec d_3 \\ \hline Y & : & c \succ b \succ d_1 \succ d_2 \succ d_3 \end{array}.$$

Player 2's preferences over outcomes are given by

$$\begin{array}{ccl} X & : & b \prec c \prec d_1 \prec d_3 \prec d_2 \\ \hline Y & : & c \succ b \succ d_1 \succ d_3 \succ d_2 \end{array}.$$

Player 1's preferences over outcome–evidence pairs are given by

$$\begin{array}{ccl} X & : & b, y \prec c, y \prec b, x \prec c, x \prec d_3, y \prec d_3, x \prec d_2, y \prec d_2, x \prec d_1, y \prec d_1, x \\ \hline Y & : & c, x \succ c, y \succ b, x \succ b, y \succ d_3, x \succ d_3, y \succ d_2, x \succ d_2, y \succ d_1, x \succ d_1, y \end{array}.$$

It can be checked that player 1's preferences are separable with $E_1^{\ell}(X) = E_1^{\ell}(Y) = \{x\}$. The important point to note is that in state X, the cost for player 1 to produce evidence y outweighs her preference for outcome c over b, whereas in state Y, the outcome preference for c over b outweighs the cost of producing evidence y.

Consider the SCF f, where $f(X) = b$ and $f(Y) = c$. The following observations hold.

(i) No veto power is satisfied, because in state X there is no outcome that is most preferred by any two players, whereas in state Y, all players unanimously prefer outcome c over any other and $f(Y) = c$. The latter point shows that the environment is not economic.

(ii) SCF f is evidence-monotonic; this can be verified by using $e_1^*(X) = e_1^*(Y) = x$ in Definition 4.

(iii) SCF f is not implementable. Suppose, to the contrary, that it is implementable by a mechanism (M, g). Then there is a message profile m such that $g(m, x) = b$ and (m, x) is a Nash equilibrium at X. Since b is bottom-ranked by players 2 and 3 at state X, any unilateral deviation by either player 2 or 3 must not change the outcome. Moreover, if player 1 deviates by sending some cheap-talk message together with evidence x, this cannot induce outcome c; otherwise, this would constitute a profitable deviation for him at state X. Since $f(Y) = c$, (m, x) cannot be a Nash equilibrium at Y, hence player 1 must have a unilateral deviation from (m, x) by submitting evidence y with some cheap-talk message to induce outcome c. Let the resulting message profile and evidence be (m', y). But then (m', y) is a Nash equilibrium at state Y, because outcome c is top-ranked by players 2 and 3, and no unilateral deviation of player 1 can induce (c, x). Since $y \notin E_1^\ell(Y)$, we contradict the assumption that f is implemented by (M, g). ◊

Therefore, in a general evidentiary setting, evidence monotonicity, no veto power, and $n \geq 3$ do not guarantee that a SCF is implementable (even if one assumes separable preferences). However, these conditions are sufficient in settings of hard evidence:

Theorem 3. *Assume $n \geq 3$ and a setting of hard evidence. If f is evidence-monotonic and satisfies no veto power, then f is implementable.*

Proof. Consider the proof of Theorem 2 and the mechanism constructed therein. The economic environment condition was used only in Step 2 of the argument, so it suffices here to deal with equilibria that fall into Rule 2 or Rule 3 of the mechanism. Suppose the true state is θ' and there is an equilibrium (m, e) that falls into Rule 2 of the mechanism. It must be that for all $j \neq i$ (where i is defined in Rule 2), $(g(m, e), e_j) \in \arg\max_{b, e'_j} U_j(b, e'_j, \theta')$: otherwise, one of these $n - 1$ players, say j^*, can profitably deviate into Rule 3, submitting some e'_{j^*} and requesting and receiving some b such that $U_{j^*}(b, e'_{j^*}, \theta') > U_{j^*}(g(m, e), e_{j^*}, \theta')$. No veto power now implies $g(m, e) = f(\theta')$. Moreover, for every player k, we must have $e_k \in E_k^\ell(\theta')$,[18] since in a hard-evidence setting it is strictly dominated for k to submit any $e_k \notin E_k^\ell(\theta')$. Therefore, $g(m, e) = f(\theta')$ and $e \in E^\ell(\theta')$, as required. A similar argument applies if (m, e) falls into Rule 3. □

Beyond its intrinsic interest, Theorem 3 also serves as a strict generalization of Maskin's (1999) classic sufficiency result, because the traditional environment without evidence is a special case of a hard-evidence setting where for any i, $E_i^\ell(\theta) = E_i$ for all θ.

[18]This notation uses the fact that a hard-evidence setting is separable.

4. Inalienable and alienable evidence

As already noted, an important feature of the current implementation exercise is that evidence is inalienable, i.e., a player's evidentiary choice is his private domain. It is of interest to understand how constrained the planner is by evidence inalienability. To this end, consider a hypothetical problem where the planner can in fact choose the profile of evidence along with the outcome, i.e., evidence becomes alienable. Formally, define an *extended outcome space* $\hat{A} := A \times E$ and say that a correspondence $\hat{f} : \Theta \rightrightarrows \hat{A}$ is an *extension of a SCF* $f : \Theta \to A$ if there exists a correspondence $\hat{e} : \Theta \rightrightarrows E$ such that $\hat{f} = (f, \hat{e})$, by which we mean that for all θ, $\hat{f}(\theta) = \{(a, e) : a = f(\theta), e \in \hat{e}(\theta)\}$. The reason to consider extensions of f that are correspondences even though f is a function is clarified shortly. Say that $\hat{f}$ is a *costless extension* of f if $\hat{f} = (f, \hat{e})$ for some correspondence $\hat{e}$ such that for all θ, $\hat{e}(\theta) \subseteq E^{\ell}(\theta, f(\theta))$.

If evidence were alienable, then the planner would face a standard implementation problem on the extended outcome space, hence Maskin monotonicity of an extended social choice rule would be a necessary and almost sufficient condition for its implementation.[19] Theorems 1 and 2 establish that with inalienable evidence, evidence monotonicity is necessary and, under some other conditions, also sufficient for implementation. We now derive a close relationship between evidence monotonicity in the underlying problem and Maskin monotonicity on the extended outcome space.

Theorem 4. *A SCF is evidence-monotonic if and only if it has a costless extension that is Maskin-monotonic on the extended outcome space.*

See the Appendix for the proof for Theorem 4, as well as proofs for Propositions 2 and 3 that appear below.

The equivalence in Theorem 4 requires allowing costless extensions to be correspondences; specifically, the "only if" direction would fail if one restricts attention to costless extensions that are single-valued.[20] Theorem 4 provides the foundation for a "bridge" between our implementation problem with inalienable evidence and a hypothetical implementation problem with alienable evidence. In particular, we can state the following corollary.

---

[19] An extended social choice rule $\hat{f} : \Theta \rightrightarrows \hat{A}$ is Maskin-monotonic provided that for all θ, $(a, e) \in \hat{f}(\theta)$, and θ', if

$$\forall i, b, e_i' : [U_i(a, e_i, \theta) \geq U_i(b, e_i', \theta) \Longrightarrow U_i(a, e_i, \theta') \geq U_i(b, e_i', \theta')],$$

then $(a, e) \in \hat{f}(\theta')$.

[20] Here is an example. Let $\Theta = \{\theta_1, \theta_2, \theta_3\}$, let $n = 1$, and let the setting be of separable preferences. Let $E_1 = \{x, y\}$, with $E_1^{\ell}(\theta_1) = \{x\}$, $E_1^{\ell}(\theta_2) = \{y\}$, and $E_1^{\ell}(\theta_3) = \{x, y\}$. The agent's outcome preferences are state-independent, with outcome a always being his most preferred outcome. The SCF f is given by $f(\theta) = a$ for all θ. This SCF is evidence-monotonic because Definition 4 is verified by any $e^*(\cdot)$ such that for all θ, $e^*(\theta) \in E_1^{\ell}(\theta)$. However, any single-valued costless extension of the SCF, say $\hat{f}$, is not Maskin-monotonic on the extended outcome space: since both $\hat{f}(\theta_1) = (a, x)$ and $\hat{f}(\theta_2) = (a, y)$ maximize the agent's utility in state θ_3, Maskin monotonicity requires that $\hat{f}(\theta_1) \in \hat{f}(\theta_3)$ and also $\hat{f}(\theta_2) \in \hat{f}(\theta_3)$, which is not possible unless $\hat{f}(\theta_3)$ is multivalued.

Corollary 2. *Assume $n \geq 3$ and an economic environment. Then any SCF f is implementable (with inalienable evidence) if and only if a costless extension of f is implementable with alienable evidence on the extended outcome space.*

Proof. Assume $n \geq 3$ and an economic environment. By Theorems 1 and 2, a SCF f is implementable if and only if it is evidence-monotonic. On the other hand, any extension of f trivially satisfies standard no veto power on the extended outcome space,[21] hence is implementable with alienable evidence if and only if it is Maskin-monotonic on the extended outcome space. The conclusion now follows from Theorem 4. □

For economic environments with at least three agents, we therefore have an equivalence between implementation with inalienable evidence and implementation with alienable evidence. The "only if" direction of Corollary 2 is intuitive, as inalienability can only make implementation harder than if evidence were alienable. Indeed, this direction of the equivalence does not depend on having $n \geq 3$ or an economic environment, because of the following logic: suppose (M, g) is a mechanism that implements f. Define $\hat{M}$ by $\hat{M}_i = M_i \times E_i$ for all i and define $\hat{g} : \hat{M} \to \hat{A}$ by $\hat{g}(m, e) = (g(m, e), e)$ for all (m, e). Then, on the extended outcome space, $(\hat{M}, \hat{g})$ achieves standard implementation of the costless extension of f given by $\hat{f} = (f, \hat{e})$, where for all θ, $\hat{e}(\theta) := \{e \in E : \exists m \text{ s.t. } (m, e) \in \mathrm{NE}(M, g, \theta)\}$.[22]

On the other hand, the "if" direction of Corollary 2 is quite surprising because it implies that given $n \geq 3$ and an economic environment, making evidence alienable is of no benefit to the planner. To understand why, assume f has a costless extension $\hat{f} = (f, \hat{e})$, where $\hat{e}(\cdot)$ is single-valued. Consider the canonical mechanism presented in Maskin (1999) that would be used to implement $\hat{f}$ on the extended outcome space, assuming $n \geq 3$ and that $\hat{f}$ satisfies standard no veto power on the extended outcome space (which is ensured by an economic environment). This mechanism gives each agent i a cheap-talk message space $\hat{M}_i := \Theta \times A \times E \times \mathbb{N}$ and has an outcome function $\hat{g} : \hat{M}_1 \times \cdots \times \hat{M}_n \to A \times E$. The key observation is that because agents' preferences do not depend on other agents' evidence, it is possible to reduce the message space for each agent i to $\hat{M}_i' := \Theta \times A \times E_i \times \mathbb{N}$ and use an outcome function $\hat{g}' : \hat{M}_1' \times \cdots \times \hat{M}_n' \to A \times E$ that is consistent with $\hat{g}$ except that in Rules 2 and 3, when "rewarding" a deviator i, $\hat{g}'$ chooses for any $j \neq i$ the evidence that he has announced, and similarly when not rewarding a deviator in Rule 2, $\hat{g}'$ can choose for him the evidence he has announced.[23] This modified mechanism always chooses an evidence for each agent that

[21] An extended social choice rule $\hat{f} : \Theta \rightrightarrows A \times E$ satisfies standard no veto power provided that for all θ,

$$\text{if } (a, e) \text{ is such that} \Big[\Big| \Big\{ i : (a, e_i) \in \arg\max_{a', e_i'} U_i(a', e_i', \theta) \Big\} \Big| \geq n - 1 \Big], \text{ then } (a, e) \in \hat{f}(\theta).$$

[22] This argument actually shows an alternate proof of Theorem 1 as a corollary of Theorem 4: if f is implementable, it must have a costless extension that is implementable on the extended outcome space, hence the costless extension must be Maskin-monotonic on the extended outcome space, and hence by Theorem 4, f must be evidence-monotonic.

[23] Note that one has to make a small and obvious modification to the definition of Rule 2, since the message spaces are no longer identical across agents.

he has announced, hence even if evidence were inalienable, every equilibrium at any state θ would result in an outcome $f(\theta)$ and some costless evidence profile, i.e., we have achieved implementation with inalienable evidence. Indeed, given that $\hat{e}(\cdot)$ is single-valued by hypothesis here, this is precisely the mechanism used in proving Theorem 2.

Corollary 2 provides additional insight as to why the economic environment condition in Theorem 2 cannot be weakened to the SCF that satisfies no veto power. By Corollary 2, what is sufficient to implement a SCF f that is evidence-monotonic (given $n \geq 3$) is the existence of a costless extension that satisfies standard no veto power on the extended outcome space. This is a more demanding requirement than f satisfying no veto power. It is this gap that drives Example 4: the SCF f defined there has a unique costless extension $\hat{f}$ given by $\hat{f}(X) = (b, x)$ and $\hat{f}(Y) = (c, x)$. While f satisfies no veto power, $\hat{f}$ does not satisfy standard no veto power on the extended outcome space because in state Y, (c, y) is top-ranked by players 2 and 3 but involves costly evidence for player 1. Such a gap is not special; rather, it is to be expected in noneconomic environments, even when preferences are separable.[24]

The following example shows that the equivalence identified in Corollary 2 fails without the assumption of an economic environment.

Example 5. Assume $\Theta = \{\theta_1, \theta_2\}$ and $n = 3$, but only player 1 has non-cheap-talk evidence, so we ignore the evidence of players 2 and 3. Evidence is hard evidence with $E_1 = \{x, y_1, y_2\}$, $E_1^{\ell}(\theta_1) = \{x, y_1\}$, and $E_1^{\ell}(\theta_2) = \{x, y_2\}$. Additionally assume $A = \{a_1, a_2\}$ and all players have identical preferences over outcomes such that they both strictly prefer a_1 to a_2 in state θ_1 while they both strictly prefer a_2 to a_1 in state θ_2. The SCF f is given by $f(\theta_1) = a_2$ and $f(\theta_2) = a_1$.

It is straightforward that f cannot be implemented: given any mechanism (M, g), pick any message profile m. We must have $g(m, x) \in \{a_1, a_2\}$. If $g(m, x) = a_1$, then (m, x) would be an undesirable Nash equilibrium at state θ_1; if $g(m, x) = a_2$, then (m, x) would be an undesirable Nash equilibrium at state θ_2.

Now consider the costless extension of f, $\hat{f}$ defined by $\hat{f}(\theta_1) = (a_2, y_1)$ and $\hat{f}(\theta_2) = (a_1, y_2)$. On the extended outcome space, $\hat{f}$ can be implemented by just asking player 1 to send a cheap-talk direct message about the state and choosing (a_2, y_1) if he reports θ_1 while choosing (a_1, y_2) if he reports θ_2. Notice that $\hat{f}$ does not satisfy standard no veto power on the extended outcome space; indeed, there is no costless extension of f that does.[25] ◊

[24] Fix a noneconomic environment. Then for some θ, a, e, and $J \subseteq I$ with $|J| = n - 1$, we have that $\forall i \in J : (a, e_i) \in \arg\max_{b, e_i'} U_i(b, e_i', \theta)$. Without loss, let $n \notin J$. As long as n has some costly evidence at θ given outcome a (i.e., $E_n^{\ell}(a, \theta) \neq E_n$), no costless extension of f can satisfy standard no veto power on the extended outcome space.

[25] In this example, f can be implemented with inalienable evidence if the planner is allowed to forbid evidence, just by prohibiting player 1 from submitting evidence x. This observation suggests that the gap between implementation with alienable evidence and inalienable evidence can be narrowed even further than Corollary 2 if the planner has the ability to forbid some evidence when evidence is inalienable. We defer this topic to future research.

5. Hard evidence

In a variety of contexts, many models assume that players can partially prove the state of the world by providing evidence. We now study how our general results yield particular insights for implementation in such environments. We assume throughout this section that the setting is hard evidence as formalized in Definition 2: each player i's preferences are separable and represented by $u_i(a, \theta) - c_i(e_i, \theta)$, and in any state of the world, θ, any piece of evidence, e_i, has a zero cost for player i (i.e., $c_i(e_i, \theta) = 0$) or a sufficiently large cost (i.e., $c_i(e_i, \theta) > \sup_a u_i(a, \theta) - \inf_a u_i(a, \theta)$). Plainly, player i never submits evidence that is not costless at the true state, and thus evidence e_i is proof of the event $\{\theta : e_i \in E_i^\ell(\theta)\}$.[26] Without loss, we say that the set of evidence a player i has at state θ is $E_i^\ell(\theta)$—just his costless evidence—and the entire evidence structure is given by $\{E_i^\ell(\theta)\}_{i,\theta}$. In addition, when referring to a player's preferences in this section, we always mean outcome preferences, since together with the evidence structure, this contains all relevant information about preferences over the joint space of outcomes and evidence.

5.1 *A characterization of evidence monotonicity*

We begin by providing an alternative characterization of evidence monotonicity for hard evidence. This characterization is useful because it (partially) disentangles the role of hard evidence from that of preferences over outcomes in satisfying evidence monotonicity.

Proposition 2. *In a hard-evidence setting, a SCF is evidence-monotonic if and only if there exists $e^* : \Theta \to E$ such that*

(i) for all θ, $e^*(\theta) \in E^\ell(\theta)$ and

(ii) for all θ and θ', if

$$\forall i, a : \big[u_i(f(\theta), \theta) \geq u_i(a, \theta) \Longrightarrow u_i(f(\theta), \theta') \geq u_i(a, \theta')\big] \tag{*}$$

and

$$\big(e^*(\theta) \in E^\ell(\theta')\big) \quad \textit{and} \quad \Big(\forall i : E_i^\ell(\theta') \subseteq E_i^\ell(\theta) \textit{ or } f(\theta) \in \arg\max_a u_i(a, \theta')\Big), \tag{**}$$

then $f(\theta) = f(\theta')$.

Compared to the definition of evidence monotonicity, the difference in the characterization above is that (2) has been replaced by the conjunction of (*) and (**). Notice that (*) is the usual condition in Maskin monotonicity that refers only to preferences over outcomes. This makes it transparent that evidence monotonicity in a hard-evidence setting is a generalization of Maskin monotonicity: without (**), it would reduce to Maskin monotonicity, whereas the presence of (**) makes it a weaker requirement.[27]

[26] Indeed, with hard evidence, one could work directly with the "proof structure" induced by the evidence structure; to preserve continuity of notation and exposition, we do not do so.

[27] Since a hard-evidence setting is one of separable preferences, this is also implied by Proposition 1; the current characterization just underscores the point.

Proposition 2 says that for a SCF f to be evidence-monotonic, one must find a function $e^*(\cdot)$ such that condition (**) is *falsified* for every ordered pair of states (θ, θ') over which f violates Maskin monotonicity (i.e., for which $f(\theta) \neq f(\theta')$ but (*) is satisfied). Plainly, this is not possible unless some player's evidence set varies across such a pair of states. Fix some $e^*(\cdot)$ and a pair of states θ and θ' such that $f(\theta) \neq f(\theta')$. If $f(\theta)$ does not go down in any agent's preference ordering when the state changes from θ to θ' (i.e., condition (*) is satisfied), then we know from Maskin (1999) that to implement f, a mechanism has to exploit evidence. Condition (**) says that for this to be possible, either (i) the evidence profile being submitted at θ, $e^*(\theta)$, is not available at θ' (negating the first part of (**)) or (ii) some player must have evidence at θ' that is not available at θ, *and* outcome $f(\theta)$ should not be this player's most preferred outcome at θ' (together, negating the second part of (**)). The latter preference requirement is essential, because otherwise a player cannot be given incentives to disprove θ when the true state is θ' and he has the ability to submit evidence supporting θ; this is why the requirement enters only in the second part of (**) and not the first part.

The existential quantifier over $e^*(\cdot)$ in Proposition 2 raises the possibility that it may be tedious to verify whether a given SCF is evidence-monotonic. Subsequently, we show how the task can be simplified in particular domains. For now, let us note that, loosely speaking, one should choose $e^*(\theta)$ to be an evidence profile that is "most informative" about state θ with respect to the other states that cause a problem for Maskin monotonicity. In particular, if there is an evidence profile in state θ that proves more about the state than any other evidence profile available at θ, then one can take $e^*(\theta)$ to be this evidence profile. The idea can be illustrated by returning to the first motivating example of the Introduction, as follows.

Example 6. A principal is concerned with dividing a fixed sum of money, say $M > 0$, to agents as some function of their individual production. The outcome space is $A = \{(a_1, \dots, a_n) \in \mathbb{R}^n_+ : \sum_i a_i \leq M\}$. A state θ is a vector of units of output, i.e., $\theta = (\theta_1, \dots, \theta_n) \in \Theta = \mathbb{R}^n_+$. Each agent can show his true output or some subset of it, hence an agent is unable to claim that his output is greater than it in fact is, but he can claim that it is less. Formally, $E_i = \Theta$ and $E_i^\ell(\theta) = [0, \theta_i]$ for all i, θ. Assume that $u_i((a_1, \dots, a_n), \theta)$ is strictly increasing in a_i.

For any SCF f, write $f(\theta) =: (f_1(\theta), \dots, f_n(\theta))$. It follows that any SCF f that satisfies $\forall i, \theta : f_i(\theta) < M$ is evidence-monotonic. To see this, fix any such f. For any θ, let $e^*(\theta) = \theta$. It suffices to argue that for any $\theta' \neq \theta$, condition (**) is violated. If $\theta' \neq \theta$, there exists an agent i such that $\theta'_i \neq \theta_i$. First, if $\theta'_i < \theta_i$, then $e^*_i(\theta) = \theta_i \notin [0, \theta'_i] = E_i^\ell(\theta')$ and so the first part of (**) is violated. Second, if $\theta'_i > \theta_i$, then $\theta'_i \in E_i^\ell(\theta')$ but $\theta'_i \notin [0, \theta_i] = E_i^\ell(\theta)$, hence $E_i^\ell(\theta') \nsubseteq E_i^\ell(\theta)$. Moreover, $f(\theta) \notin \arg\max_a u_i(a, \theta')$ because $f_i(\theta) < M$. Therefore, the second part of (**) is violated. $\diamondsuit$

The next example illustrates how the characterization in Proposition 2 can be applied and also provides more insight into the different elements of condition (**).

Example 7. Let $n = 4$, $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, and $A = \{w, x, y, z\}$. Player 1's evidence structure is given by $E_1^\ell(\theta_1) = \{\alpha, \beta\}$, $E_1^\ell(\theta_2) = \{\alpha, \beta\}$, $E_1^\ell(\theta_3) = \{\alpha\}$, and $E_1^\ell(\theta_4) = \{\alpha\}$. All other players have no evidence (i.e., for $i > 1$, $E_i^\ell(\theta)$ is constant across θ), so with some abuse of notation we ignore their evidence below. The (ordinal) preferences of players 1 and 2 are given in the left table below, while those of players 3 and 4 are given in the right table.

θ_1	θ_2	θ_3	θ_4
w	w	w	w
x	y	x	x
y	x	y	y
z	z	z	z

θ_1	θ_2	θ_3	θ_4
z	z	z	z
x	y	x	x
y	x	y	y
w	w	w	w

So, for example, $u_1(w,\theta_1) > u_1(x,\theta_1) > u_1(y,\theta_1) > u_1(z,\theta_1)$.

(i) Consider the SCF f, where $f(\theta_1) = f(\theta_4) = x$ and $f(\theta_2) = f(\theta_3) = y$. Since $E^\ell(\theta_3) = E^\ell(\theta_4) = \{\alpha\}$, any $e^*(\cdot)$ that verifies evidence monotonicity satisfies $e^*(\theta_3) = e^*(\theta_4) = \alpha$. Since no player's preferences change between states θ_3 and θ_4, both (*) and (**) are satisfied for $\theta = \theta_3$ and $\theta' = \theta_4$. Hence f is not evidence-monotonic and, by Theorem 1, not implementable.

(ii) Next consider the SCF f^*, where $f^*(\theta_1) = x$ and $f^*(\theta_2) = f^*(\theta_3) = f^*(\theta_4) = y$. Even though f^* is not Maskin-monotonic (because $f^*(\theta_1) \neq f^*(\theta_3)$ while preferences do not change between states θ_1 and θ_3), one can check that f^* is evidence-monotonic by using $e^*(\theta_1) = e^*(\theta_2) = \beta$ and $e^*(\theta_3) = e^*(\theta_4) = \alpha$ in Proposition 2. Note that $f^*(\theta_1) = x$ is not top-ranked for any agent in state θ_3 and $f(\theta_3) = y$ is not top-ranked for any agent in state θ_1; this is essential to the violation of (**) for $\theta = \theta_3$ and $\theta' = \theta_1$. Furthermore, f^* satisfies no veto power because no alternative is top-ranked in any state by more than two players. Since f^* is evidence-monotonic, it is implementable by Theorem 3.

(iii) Finally, consider the SCF $\tilde{f}$, where $\tilde{f}(\theta_1) = x$ and $\tilde{f}(\theta_2) = \tilde{f}(\theta_3) = \tilde{f}(\theta_4) = w$. Consider $\theta = \theta_3$ and $\theta' = \theta_1$. Since no player's preferences change between θ_1 and θ_3, (*) is satisfied. Since we must have $e^*(\theta_3) = \alpha$, hence $e^*(\theta_3) \in E^\ell(\theta_1)$, the first part of (**) is also satisfied. The second part of (**) is trivially satisfied for all $i > 1$ (they have no evidence); it is also satisfied for player 1 because w is top-ranked for him in all states. Thus, (**) is satisfied and $\tilde{f}$ is not evidence-monotonic, hence not implementable. The problem here is that even though $E_1^\ell(\theta_3) \subsetneq E_1^\ell(\theta_1)$, it is not possible to reward player 1 in state θ_1 for disproving θ_3, because $\tilde{f}(\theta_3) = w$ is player 1's most preferred outcome in state θ_1. ◊

5.2 *Evidence monotonicity and distinguishability*

Proposition 2 makes no assumption on preferences, the SCF, or the evidence structure (beyond a hard-evidence setting). We now discuss how the characterization can be considerably simplified. The key is to completely disentangle properties of the evidence structure from properties of outcome preferences in verifying evidence monotonicity. In general, condition (**) shows that these are inextricably linked, so some domain restriction is needed for this approach. The following condition of *nonsatiation* proves useful.

Definition 7 (Nonsatiation). A SCF f satisfies nonsatiation if for all i, θ, and $a \in \bigcup_{\theta'} f(\theta')$, there exists $\tilde{a}$ such that $u_i(\tilde{a}, \theta) > u_i(a, \theta)$.

Intuitively, if a SCF satisfies nonsatiation, it is always possible to reward players. For example, nonsatiation is satisfied if there is a divisible private good that is positively valued by players and the SCF never allocates all of the private good to any player.[28] The important implication is that if f satisfies nonsatiation, the preference requirement in condition (**) can be ignored because the requirement is satisfied independently of the evidence structure. Hence, under nonsatiation, Proposition 2 holds if we replace (**) with the following weaker condition that depends only on the evidence structure:

$$e^*(\theta) \in E^\ell(\theta') \quad \text{and} \quad E^\ell(\theta') \subseteq E^\ell(\theta). \tag{***}$$

Plainly, nonsatiation is essential for this reduction. This can be seen by returning to the SCF $\tilde{f}$ in Example 7. Condition (***) is falsified for $\theta' = \theta_1$ and $\theta = \theta_3$ because $E^\ell(\theta_1) \nsubseteq E^\ell(\theta_3)$. But, as argued in the example, $\tilde{f}$ is not evidence-monotonic and thus not implementable: $\tilde{f}$ does not satisfy nonsatiation.

Since nonsatiation allows us to completely separate the roles of preferences and evidence structure in evidence monotonicity (respectively captured by (*) and (***)), the remainder of this section derives a number of simplifications and implications of evidence monotonicity within the domain of SCF's that satisfy nonsatiation.

A central question is what hard evidence structures permit non-Maskin-monotonic SCF's to be implemented. We provide a sharp answer by introducing a notion of *distinguishability*. Recall that an ordered pair of states (θ, θ') violates Maskin monotonicity if $f(\theta) \neq f(\theta')$ but (*) is satisfied. For any θ, define

$$T^f(\theta) := \{\theta' \in \Theta : (\theta, \theta') \text{ violates Maskin monotonicity}\}.$$

In other words, given that the planner wishes to implement the outcome $f(\theta)$ in state θ, $T^f(\theta)$ is the set of states that causes a problem for implementation of f in the absence of evidence. In particular, f is implementable without evidence only if $\bigcup_{\theta \in \Theta} T^f(\theta) = \varnothing$. Naturally, to implement f in a setting with evidence, θ and $T^f(\theta)$ should be appropriately "distinguishable" using evidence. This notion is made precise by the following definition.

Definition 8 (Distinguishability). For any θ and $\Omega \subseteq \Theta$, θ and Ω are distinguishable if for any $\Omega' \subseteq \Omega$, $E^\ell(\theta) \neq \bigcup_{\theta' \in \Omega'} E^\ell(\theta')$.

Thus, a state θ is distinguishable from an event or set of states Ω if for *every* subset Ω' of Ω, either some player can disprove Ω' when θ is the true state (which requires $E_i^\ell(\theta) \nsubseteq \bigcup_{\theta' \in \Omega'} E_i^\ell(\theta')$) or some player can disprove θ when the true state state is in Ω'

[28] Recall that the environment is economic if there is a divisible private good that is positively valued by all players. Nonsatiation further requires that the SCF be such that no player receives all of the private good in any state. More generally, nonsatiation does not imply an economic environment, but the conjunction of nonsatiation and no veto power does.

(which requires $E_i^\ell(\theta) \not\supseteq \bigcup_{\theta' \in \Omega'} E_i^\ell(\theta')$). Notice that if θ is distinguishable from Ω, then θ is distinguishable from any subset of Ω. Consequently, if θ and Ω are distinguishable, then θ must be "pairwise distinguishable" from every $\theta' \in \Omega$ (in particular, $\theta \notin \Omega$). However, such pairwise distinguishability need not be enough for distinguishability. The following example illustrates this point and also shows how distinguishability is connected to evidence monotonicity and implementability.

Example 8. There are two propositions: a and b. Each member of a group of three or more experts knows which of the two propositions are true, if any. Due to time or space limitations, however, each one can provide a proof of at most one proposition. This problem can be represented by $\Theta = \{\varphi, a, b, ab\}$, and for all i, $E_i^\ell(\varphi) = \{\varphi\}$, $E_i^\ell(a) = \{\varphi, a\}$, $E_i^\ell(b) = \{\varphi, b\}$, and $E_i^\ell(ab) = \{\varphi, a, b\}$, where φ represents "neither proposition is true" or "no proof provided."

Suppose now that the experts' preferences over outcomes are state-independent, so that (*) is always satisfied. Then, for any choice of $\{e_i^*(ab)\}_{i=1}^n$, there exists $\theta' \in \{a, b\}$ such that (**) is satisfied with $\theta = ab$. It follows from Proposition 2 that not every SCF is evidence-monotonic, and hence that not every SCF is implementable. In particular, implementability requires $f(ab) \in \{f(a), f(b)\}$. In the other direction, by choosing $e_i^*(\varphi) = \varphi$, $e_i^*(a) = a$, and $e_i^*(b) = b$, we see that $f(ab) \in \{f(a), f(b)\}$ is also sufficient for the SCF f to falsify (***) and thus for f to be implementable under no veto power and nonsatiation.

Hence, although this evidence structure satisfies "pairwise distinguishability" (i.e., $\theta \neq \theta' \Longrightarrow E(\theta) \neq E^\ell(\theta')$), evidence monotonicity is not guaranteed even under nonsatiation. The reason is that although state ab is distinguishable from any other state $\theta \in \{\varphi, a, b\}$, state ab is not distinguishable from the event $\{\varphi, a, b\}$ because of the assumption that an expert can provide a proof of at most one proposition.[29] $\Diamond$

The following result characterizes evidence monotonicity in terms of distinguishability.

Proposition 3. *Assume f satisfies nonsatiation. Then f is evidence-monotonic if and only if for all θ, θ and $T^f(\theta)$ are distinguishable.*

An immediate implication of Proposition 3 is that a simple condition on the evidence structure guarantees that any SCF that satisfies nonsatiation is evidence-monotonic, no matter what the agents' preferences are:

$$\forall \theta : \theta \text{ is distinguishable from } \Theta \setminus \{\theta\}. \tag{UD}$$

Condition (UD), short for *universal distinguishability*, requires that each state must be distinguishable from any event that does not contain it. As seen in Example 8, this is generally a stronger requirement than every state being distinguishable from every other state.

[29]If each expert could prove both a and b when the two statements are both true, then $E_i^\ell(ab)$ would be augmented by ab, in which case the state ab would be distinguishable from the event $\{\varphi, a, b\}$ and any SCF that satisfies nonsatiation would be evidence-monotonic.

Corollary 3. *Assume $n \geq 3$. Any SCF that satisfies both no veto power and nonsatiation can be implemented if (UD) holds.*

Proof. Let f be an arbitrary SCF that satisfies nonsatiation and no veto power. Pick any θ. Since $\theta \notin T^f(\theta)$, (UD) implies that θ is distinguishable from $T^f(\theta)$. By Proposition 3, f is evidence-monotonic. Theorem 3 yields the desired conclusion. $\square$

Corollary 3 is tight in the sense that if (UD) is violated, there exists a profile of utility functions (for $n \geq 3$) and a SCF that satisfies no veto power and nonsatiation such that the SCF is not implementable. We illustrate Corollary 3 with the following example.

Example 9. Let $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $n = 3$, and let $E_1^\ell(\theta_1) = \{x, y\}$, $E_1^\ell(\theta_2) = \{x\}$, $E_1^\ell(\theta_3) = \{y\}$, $E_2^\ell(\theta_1) = \{x, y\}$, $E_2^\ell(\theta_2) = \{y\}$, $E_2^\ell(\theta_3) = \{x\}$, and $E_3^\ell(\theta) = \{z\}$ for all θ. Then (UD) holds because $E^\ell(\theta_1) = \{(x, x, z), (x, y, z), (y, x, z), (y, y, z)\}$, $E^\ell(\theta_2) = \{(x, y, z)\}$, and $E^\ell(\theta_3) = \{(y, x, z)\}$. Hence, by Corollary 3, any SCF that satisfies no veto power and nonsatiation is implementable. $\Diamond$

Many models with hard evidence, both in mechanism design and beyond, assume that the structure of hard evidence satisfies a property known as *normality*, which captures the idea that agents face no constraints on time, effort, space, etc. in providing evidence.[30] The formal definition follows.

Definition 9 (Normality). The evidence structure is normal or satisfies normality if for all i and θ, there is some $\bar{e}_i(\theta) \in E_i^\ell(\theta)$ such that $[\bar{e}_i(\theta) \in E_i^\ell(\theta') \Longrightarrow E_i^\ell(\theta) \subseteq E_i^\ell(\theta')]$.

The formulation above follows Bull and Watson (2007). It says that for any player i and state θ, there is some evidence $\bar{e}_i(\theta)$ that can be interpreted as maximal or summary evidence because it proves by itself what agent i could prove by jointly sending all his available evidence. The condition is equivalent to the *full reports* condition of Lipman and Seppi (1995) or the *minimal closure* condition of Forges and Koessler (2005), and is somewhat weaker than Green and Laffont's (1986) *nested range condition* in their "direct mechanism" setting.[31]

To illustrate the property, consider Example 6 again, where $\Theta = \mathbb{R}_+^n$, $E_i = \Theta$, and $E_i^\ell(\theta) = [0, \theta_i]$ for all i, θ. This evidence structure is seen to be normal by setting $\bar{e}_i(\theta) = \theta_i$ for all i, θ: if $\bar{e}_i(\theta) = \theta_i \in E_i^\ell(\theta') = [0, \theta_i']$, then it must be that $\theta_i \leq \theta_i'$ and so $E_i^\ell(\theta) = [0, \theta_i] \subseteq [0, \theta_i'] = E_i^\ell(\theta')$, as required. It is straightforward to also check that Example 7 satisfies normality, whereas Example 8 does not.

[30]Exceptions include Bull and Watson (2007), Glazer and Rubinstein (2001, 2004, 2006), Lipman and Seppi (1995), and Sher (2010).

[31]Green and Laffont (1986) take $E_i = \Theta$ and assume that $\theta \in E_i^\ell(\theta)$. The nested range condition says that if $\theta' \in E_i^\ell(\theta)$ and $\theta'' \in E_i^\ell(\theta')$, then $\theta'' \in E_i^\ell(\theta)$. This implies normality because for all i and and θ, one can set $\bar{e}_i(\theta) = \theta$. To see that normality is strictly weaker, consider the example $\Theta = \{\theta_1, \theta_2, \theta_3\}$: for all i, $E_i^\ell(\theta_1) = \{\theta_1, \theta_2\}$ and $E_i^\ell(\theta_2) = E_i^\ell(\theta_3) = \{\theta_2, \theta_3\}$. Normality is verified by choosing, for all i, $\bar{e}_i(\theta_1) = \theta_1$ and $\bar{e}_i(\theta_2) = \bar{e}_i(\theta_3) = \theta_3$. By contrast, the nested range condition is violated because for all i, $\theta_2 \in E_i^\ell(\theta_1)$ yet $\theta_3 \in E_i^\ell(\theta_2)$ and $\theta_3 \notin E_i^\ell(\theta_1)$.

For SCF's that satisfy nonsatiation, normality implies that the characterization in Proposition 2 can be significantly simplified: a SCF that satisfies nonsatiation is evidence-monotonic if and only if for all θ and θ', if (*) and

$$E^{\ell}(\theta) = E^{\ell}(\theta') \tag{5}$$

hold, then $f(\theta) = f(\theta')$.

The reason is that under normality, no matter what $e^*(\cdot)$ is, (***) is equivalent to (5). Thus the existential qualifier in Proposition 2 can be dropped altogether.

A related observation is that when the evidence structure is normal, distinguishability of any state θ and event Ω is equivalent to the distinguishability of θ from each $\theta' \in \Omega$.[32] Combining this with Proposition 3 yields the following corollary.

Corollary 4. *Assume the evidence structure is normal. A SCF that satisfies nonsatiation is evidence-monotonic if and only if for any θ and $\theta' \in T^f(\theta)$, $E^{\ell}(\theta) \neq E^{\ell}(\theta')$.*

Corollary 4 identifies exactly which normal evidence structures permit implementation of non-Maskin-monotonic SCF's (under $n \geq 3$, nonsatiation, and no veto power). It can be combined with our earlier results to derive additional corollaries, such as that which follows.

Corollary 5. *Assume $n \geq 3$ and that the evidence structure is normal. A SCF f that satisfies both no veto power and nonsatiation is implementable if*

$$\forall \theta, \theta' : [E^{\ell}(\theta) = E^{\ell}(\theta') \Longrightarrow f(\theta) = f(\theta')]. \tag{6}$$

Remark 3. Ben-Porath and Lipman (2011) study implementation with hard evidence. They refer to condition (6) as "measurability" and show in their Proposition 1 that in this setting, a SCF is implementable when preferences are state-independent only if it satisfies (6). Our Proposition 2 in fact implies a stronger necessary condition for implementation with state-independent preferences when the hard-evidence structure is normal:

$$\forall \theta, \theta' : \begin{bmatrix} f(\theta) = f(\theta') \\ \text{if} \\ \forall i : \text{either } E_i^{\ell}(\theta) = E_i^{\ell}(\theta') \text{ or } (f(\theta) \in \arg\max_a u_i(a, \theta') \text{ and } E_i^{\ell}(\theta) \subseteq E_i^{\ell}(\theta')) \end{bmatrix}. \tag{7}$$

To verify this, observe first that under state-independent preferences, (*) is satisfied for all θ and θ'; second, if the conditional in (7) holds for some θ and θ', then normality implies that (**) is satisfied regardless of the choice of $e^*(\cdot)$ (in particular, when $e^*(\theta) = \bar{e}(\theta)$ for all θ).

Condition (7) is obviously stronger than (6) and emphasizes the necessity of being able to reward players for evidence submission. Under nonsatiation, the two conditions

[32] To see this, fix θ and $\Omega \subseteq \Theta$, and assume that $\forall \theta' \in \Omega : E^{\ell}(\theta) \neq E^{\ell}(\theta')$. Suppose, per contra, that for some $\Omega' \subseteq \Omega : E^{\ell}(\theta) = \bigcup_{\theta' \in \Omega'} E^{\ell}(\theta')$. Then for all $\theta' \in \Omega'$, $\bar{e}(\theta') \in E^{\ell}(\theta)$ (where $\bar{e}$ is from the definition of normality) and, moreover, for some $\tilde{\theta} \in \Omega'$, $\bar{e}(\theta) \in E^{\ell}(\tilde{\theta})$. By normality, $E^{\ell}(\tilde{\theta}) = E^{\ell}(\theta)$, which is a contradiction.

are equivalent. Ben-Porath and Lipman (2011) also independently prove a result similar to Corollary 5. Note that even under normality, nonsatiation and no veto power, (6) is not necessary for implementation when preferences are not state-independent, for instance the SCF f^* in Example 7. The reason is that when hard evidence is the same across two states, a mechanism can still exploit preference reversals to implement different outcomes, just as in the standard environment without evidence.

6. Conclusion

This paper generalizes the implementation problem to incorporate agents' ability to provide discriminatory signals or evidence about the state. The central theme of our results is that the planner can use either agents' preferences over outcomes or their evidentiary technology to discriminate between states of the world, even though evidence submission is inalienable. We study both hard evidence, where players can prove that the state lies in some subset of all possible states, and the costly production of evidence, where evidentiary costs are nonprohibitive but vary across states. The results we obtain may be useful in terms of both necessary conditions—in particular, the finding that the ability to reward players is sometimes needed—and sufficient conditions that demonstrate how a wide class of social choice functions are implementable as a function of the evidence structure. In particular, we identify an appropriate generalization and weakening of Maskin monotonicity—evidence monotonicity—and show that this is the key to implementation with evidence.

There are a number of directions in which this research can be developed. Our analysis here substantially exploits the complete-information setting, and it is obviously important to understand how the arguments can be extended when agents have private information. We conjecture that a weakening of Jackson's (1991) Bayesian monotonicity condition in a manner similar to how evidence monotonicity weakens Maskin monotonicity will be central to Bayesian implementation, in conjunction with standard conditions like incentive compatibility.

Within the complete-information framework, it would also be useful to understand how evidence changes the implementation problem when attention is restricted to "nice" mechanisms, for example, "bounded mechanisms" (Jackson 1992). In a related vein, the presence of evidence generally allows greater scope for implementation with weaker solution concepts, such as in dominant strategies. These are likely to be fruitful avenues for further study.

Finally, we note that our notion of implementation in this paper is that no evidentiary costs should be incurred in equilibrium. This is without loss of generality in a hard-evidence setting, but is not when there are nonprohibitive evidentiary costs. For example, the literature on screening shows that it may be possible to design a mechanism that induces information revelation from an agent (in a unique equilibrium) at the cost of incurring inefficient signaling distortions; this would apply in versions of our Example 3. It would be interesting to extend our analysis to full implementation in a general framework that allows for costly evidence provision in equilibrium.

Appendix: Omitted proofs

The following lemma is used in the proof of Theorem 4.

Lemma 1. *A SCF f is evidence-monotonic if and only if there exists a nonempty-valued correspondence $e^{**}:\Theta \rightrightarrows E$ such that*

(i) for all θ, $e^{**}(\theta)\subseteq E^{\ell}(\theta, f(\theta))$, and

(ii) for all θ, θ', and $e\in e^{**}(\theta)$, if

$$\forall i, b, e_i' : \big[U_i(f(\theta), e_i, \theta) \geq U_i(b, e_i', \theta) \Rightarrow U_i(f(\theta), e_i, \theta') \geq U_i(b, e_i', \theta')\big], \tag{8}$$

then $f(\theta)=f(\theta')$ and $e\in e^{**}(\theta')$.

Notice two differences between the definition of evidence monotonicity and the condition given in Lemma 1: first, the mapping e^{**} is a correspondence, whereas the e^* in Definition 4 is a function; second, part (ii) places a requirement on the relationship between $e^{**}(\theta)$ and $e^{**}(\theta')$ that is not required by Definition 4.

Proof of Lemma 1. The "if" direction is straightforward because any single-valued selection from the correspondence $e^{**}(\cdot)$ in the lemma's statement verifies Definition 4.

So consider the "only if" direction. Assume that f is evidence-monotonic. Fix any e^* that verifies Definition 4. Define a binary relation R on $\Theta\times\Theta$ as follows: $\theta R\theta'$ whenever condition (2) in Definition 4 holds, which we reproduce here as

$$\forall i, a, e_i' : \big[U_i(f(\theta), e_i^*(\theta), \theta) \geq U_i(a, e_i', \theta) \Rightarrow U_i(f(\theta), e_i^*(\theta), \theta') \geq U_i(a, e_i', \theta')\big]. \tag{9}$$

Define the correspondence $e^{**}:\Theta \rightrightarrows E$ as follows: for any θ', $e^{**}(\theta')=\bigcup_{\theta:\theta R\theta'} e^*(\theta)$. Note that because R is reflexive, $e^*(\theta')\in e^{**}(\theta')$ for any θ'.

We show that this correspondence $e^{**}(\cdot)$ satisfies the lemma's requirements. To check the first requirement, pick any $e\in e^{**}(\theta')$. By construction, this means that $e=e^*(\theta)$ for some θ such that $\theta R\theta'$. Thus, condition (9) holds, and hence evidence monotonicity implies that $f(\theta')=f(\theta)$. Moreover, by using $e_i'=e_i^*(\theta')$ and $a=f(\theta)$ in condition (9), and the fact that $e^*(\theta')\in E^{\ell}(\theta', f(\theta'))=E^{\ell}(\theta', f(\theta))$ (by part (i) of Definition 4), it follows that $e=e^*(\theta)\in E^{\ell}(\theta', f(\theta))=E^{\ell}(\theta', f(\theta'))$. Since e was an arbitrary choice from $e^{**}(\theta')$, it follows that for all θ', $e^{**}(\theta')\subseteq E^{\ell}(\theta', f(\theta'))$, which is the first requirement of the lemma.

We now show that $e^{**}(\cdot)$ also satisfies part (ii) of the lemma's requirements. Pick any θ, θ', and $e\in e^{**}(\theta)$ that satisfy (8). We must prove that $f(\theta)=f(\theta')$ and $e\in e^{**}(\theta')$. Note that since $e\in e^{**}(\theta)$, there must exist θ'' such that $e=e^*(\theta'')$ and $\theta'' R\theta$. By definition of R,

$$\forall i, a, e_i' : \big[U_i(f(\theta''), e_i^*(\theta''), \theta'') \geq U_i(a, e_i', \theta'') \Longrightarrow U_i(f(\theta''), e_i^*(\theta''), \theta) \geq U_i(a, e_i', \theta)\big].$$

That f is evidence-monotonic implies $f(\theta'')=f(\theta)$. Hence, the preceding line is equivalent to

$$\forall i, a, e_i' : \big[U_i(f(\theta), e_i^*(\theta''), \theta'') \geq U_i(a, e_i', \theta'') \Longrightarrow U_i(f(\theta), e_i^*(\theta''), \theta) \geq U_i(a, e_i', \theta)\big]. \tag{10}$$

Furthermore, from condition (8) and $e = e^{**}(\theta'')$, we have

$$\forall i, a, e'_i : \big[U_i(f(\theta), e^*_i(\theta''), \theta) \geq U_i(a, e'_i, \theta) \Longrightarrow U_i(f(\theta), e^*_i(\theta''), \theta') \geq U_i(a, e'_i, \theta')\big]. \quad (11)$$

Combining (10) and (11) yields

$$\forall i, a, e'_i : \big[U_i(f(\theta), e^*_i(\theta''), \theta'') \geq U_i(a, e'_i, \theta'') \Longrightarrow U_i(f(\theta), e^*_i(\theta''), \theta') \geq U_i(a, e'_i, \theta')\big].$$

Since $f(\theta'') = f(\theta)$, the preceding line is equivalent to $\theta'' R \theta'$. That f is evidence-monotonic now implies that $f(\theta'') = f(\theta')$ and so $f(\theta) = f(\theta')$. Finally, observe that since $\theta'' R \theta'$, the construction of $e^{**}(\cdot)$ implies that $e = e^*(\theta'') \in e^{**}(\theta')$. □

Proof of Theorem 4. For the "only if" direction, assume f is evidence-monotonic. Pick any correspondence $e^{**}(\cdot)$ that satisfies the conditions in Lemma 1 and let $\hat{f} = (f, e^{**})$. We show that $\hat{f}$ is Maskin-monotonic on the extended outcome space. To prove this, fix any θ, $(a, e) \in \hat{f}(\theta)$, and θ'. We must show that if

$$\forall i, b, e' : \big[U_i(a, e_i, \theta) \geq U_i(b, e'_i, \theta) \Longrightarrow U_i(a, e_i, \theta') \geq U_i(b, e'_i, \theta')\big], \quad (12)$$

then $(a, e) \in \hat{f}(\theta')$. So assume (12). Then (8) is satisfied (since $a = f(\theta)$), hence Lemma 1 implies that $f(\theta') = f(\theta)$ and $e \in e^{**}(\theta')$, which together imply that $(a, e) \in \hat{f}(\theta')$, as required.

For the "if" direction, suppose $\hat{f} = (f, e^{**})$ is a costless extension of f that is Maskin-monotonic on the extended outcome space. To show that f is evidence-monotonic, it suffices to show that $e^{**}(\cdot)$ satisfies the requirements of Lemma 1. The first requirement of Lemma 1 is obviously satisfied by the definition of a costless extension; to prove the second, fix θ, θ', and $e \in e^{**}(\theta)$, and assume (8). We must show that $f(\theta') = f(\theta)$ and $e \in e^{**}(\theta')$. Condition (8) implies that (12) holds when $a = f(\theta)$. Thus, Maskin monotonicity of (f, e^{**}) implies that $(f(\theta), e) \in \hat{f}(\theta')$, which implies that $f(\theta) = f(\theta')$ and $e \in e^{**}(\theta')$, as required. □

Proof of Proposition 2. Consider the hard-evidence setting, i.e., separable preferences and for all i, θ, and e_i, either $c_i(e_i, \theta) = 0$ or $c_i(e_i, \theta) > \sup_a u_i(a, \theta) - \inf_a u_i(a, \theta)$.

First we prove sufficiency. For this, it suffices to show that (2) $\Longrightarrow$ (*) and (**). So assume (2). By considering $e'_i = e^*_i(\theta)$ in (2), it is straightforward that (*) follows. For the first part of (**), consider (2) with $a = f(\theta)$ and $e'_i \in E^\ell_i(\theta')$. Since the antecedent within (2) is then satisfied, the consequent must be true, which yields $c_i(e^*_i(\theta), \theta') \leq 0$, which implies $e^*_i(\theta) \in E^\ell_i(\theta')$. As this is true for all i, the first part of (**) is shown. Now observe that the second part of (**) is equivalent to

$$\forall i : E^\ell_i(\theta') \not\subseteq E^\ell_i(\theta) \Longrightarrow \big[\forall a : u_i(f(\theta), \theta') \geq u_i(a, \theta')\big]. \quad (13)$$

Fix any i. It suffices to show that (13) is satisfied. For any $e'_i \in E^\ell_i(\theta') \setminus E^\ell_i(\theta)$, it is straightforward to check that the antecedent within (2) is always satisfied for any a (because $c_i(e'_i, \theta) > \sup_a u_i(a, \theta) - \inf_a u_i(a, \theta)$). Hence, for any a, the consequent of (2) must be true, and given that $e^*_i(\theta) \in E^\ell_i(\theta')$ and $e'_i \in E^\ell_i(\theta') \setminus E^\ell_i(\theta)$, it follows that the consequent of (13) holds.

Next we show necessity. For this it suffices to show that (*) and (**) $\Longrightarrow$ (2). Accordingly, assume (*) and (**), which in particular implies (13) for all i. If $e'_i \notin E^\ell_i(\theta')$, then $c_i(e'_i, \theta') > \sup_a u_i(a, \theta') - \inf_a u_i(a, \theta')$, which combines with the first part of (**) to imply that (2) is satisfied (because the consequent therein holds, regardless of the antecedent). Next, if $e'_i \in E^\ell_i(\theta)$, then by (*) and the first part of (**), it follows that (2) must hold. Finally, for all $e'_i \in E^\ell_i(\theta') \setminus E^\ell_i(\theta)$, (13) combined with the first part of (**) implies that the consequent in (2) is satisfied. □

Proof of Proposition 3. Assume nonsatiation. Recall that we can then replace (**) by (***) in Proposition 2. It follows from this version of Proposition 2 that f is evidence-monotonic if and only if

$$\forall \theta, \exists e^*(\theta) \in E^\ell(\theta) \text{ s.t. } \forall \theta' \in T^f(\theta), \exists i \text{ s.t. } (e^*_i(\theta) \notin E^\ell_i(\theta')) \text{ or } (E^\ell_i(\theta') \nsubseteq E^\ell_i(\theta)). \tag{14}$$

We work with this equivalent formulation.

For the "if" direction of the result, assume that

$$\forall \theta \text{ and } \Omega \subseteq T^f(\theta) : E^\ell(\theta) \neq \bigcup_{\theta' \in \Omega} E^\ell(\theta') \tag{15}$$

and, toward contradiction, that (14) is false. This implies that there exists θ such that for all $e \in E^\ell(\theta)$, there exists $\theta'(e) \in T^f(\theta)$ for which

$$\forall i : (e_i \in E^\ell_i(\theta'(e))) \text{ and } (E^\ell_i(\theta'(e)) \subseteq E^\ell_i(\theta)).$$

Set $\Omega := \bigcup_{e \in E^\ell(\theta)} \theta'(e)$ and note that $\Omega \subseteq T^f(\theta)$. Since $e \in E^\ell(\theta'(e))$ for each $e \in E^\ell(\theta)$, it follows that $E^\ell(\theta) \subseteq \bigcup_{\theta' \in \Omega} E^\ell(\theta')$. Finally, for each $e \in E^\ell(\theta)$, we have $E^\ell(\theta'(e)) \subseteq E^\ell(\theta)$, hence $\bigcup_{\theta' \in \Omega} E^\ell(\theta') \subseteq E^\ell(\theta)$, and so $E^\ell(\theta) = \bigcup_{\theta' \in \Omega} E^\ell(\theta')$, a contradiction to (15).

For the "only if" direction, assume that f satisfies (14). We proceed again by contradiction, assuming that for some θ and $\Omega \subseteq T^f(\theta)$, $E^\ell(\theta) = \bigcup_{\theta' \in \Omega} E^\ell(\theta')$. This implies that for some θ, (i) for all $e \in E^\ell(\theta)$, there exists $\theta'(e) \in \Omega \subseteq T^f(\theta)$ such that $e \in E^\ell(\theta'(e))$, and (ii) $E^\ell(\theta'(e)) \subseteq E^\ell(\theta)$. But this contradicts the assumption that f satisfies (14). □

References

Abreu, Dilip and Arunava Sen (1991), "Virtual implementation in Nash equilibrium." *Econometrica*, 59, 997–1021. [326]

Aghion, Philippe, Drew Fudenberg, Richard Holden, Takashi Kunimoto, and Olivier Tercieux (2009), "Subgame-perfect implementation under value perturbations and the hold-up problem." Unpublished paper, University of Chicago. [327]

Ben-Porath, Elchanan and Barton L. Lipman (2011), "Implementation with partial provability." Unpublished paper, Hebrew University of Jerusalem and Boston University. [327, 330, 335, 348, 349]

Benoît, Jean Pierre and Efe A. Ok (2008), "Nash implementation without no-veto power." *Games and Economic Behavior*, 64, 51–67. [335]

Bergemann, Dirk and Stephen Morris (2008), "Ex post implementation." *Games and Economic Behavior*, 63, 527–566. [334]

Bull, Jesse (2008), "Mechanism design with moderate evidence cost." *B.E. Journal of Theoretical Economics (Contributions)*, 8, Arcticle 15. [327]

Bull, Jesse and Joel Watson (2004), "Evidence disclosure and verifiability." *Journal of Economic Theory*, 118, 1–31. [327]

Bull, Jesse and Joel Watson (2007), "Hard evidence and mechanism design." *Games and Economic Behavior*, 58, 75–93. [326, 327, 329, 331, 347]

Chung, Kim Sau and Jeffrey C. Ely (2003), "Implementation with near-complete information." *Econometrica*, 71, 857–871. [327]

Crawford, Vincent P. and Joel Sobel (1982), "Strategic information transmission." *Econometrica*, 50, 1431–1451. [327]

Dagan, Nir, Oscar Volij, and Roberto Serrano (1999), "Feasible implementation of taxation methods." *Review of Economic Design*, 4, 57–72. [327]

Deneckere, Raymond and Sergei Severinov (2007), "Optimal screening with costly misrepresentation." Unpublished paper, University of Wisconsin at Madison. [327]

Deneckere, Raymond and Sergei Severinov (2008), "Mechanism design with partial state verifiability." *Games and Economic Behavior*, 64, 487–513. [327]

Dutta, Bhaskar and Arunava Sen (2011), "Nash implementation with partially honest individuals." Unpublished paper, University of Warwick. [326, 336]

Emons, Winand and Claude Fluet (2009), "Accuracy versus falsification costs: The optimal amount of evidence under different procedures." *Journal of Law, Economics, and Organization*, 25, 134–156. [327]

Forges, Françoise and Frédéric Koessler (2005), "Communication equilibria with partially verifiable types." *Journal of Mathematical Economics*, 41, 793–811. [347]

Glazer, Jacob and Ariel Rubinstein (2001), "Debates and decisions: On a rationale of argumentation rules." *Games and Economic Behavior*, 36, 158–173. [347]

Glazer, Jacob and Ariel Rubinstein (2004), "On optimal rules of persuasion." *Econometrica*, 72, 1715–1736. [327, 347]

Glazer, Jacob and Ariel Rubinstein (2006), "A study in the pragmatics of persuasion: A game theoretical approach." *Theoretical Economics*, 1, 395–410. [327, 347]

Green, Jerry R. and Jean Jacques Laffont (1986), "Partially verifiable information and mechanism design." *Review of Economic Studies*, 53, 447–456. [327, 347]

Grossman, Sanford J. (1981), "The informational role of warranties and private disclosure about product quality." *Journal of Law and Economics*, 24, 461–483. [327]

Hurwicz, Leonid, Eric S. Maskin, and Andrew Postlewaite (1995), "Feasible Nash implementation of social choice rules when the designer does not know endowments or production sets." In *The Economics of Informational Decentralization* (John O. Ledyard, ed.), 367–433, Kluwer, Amsterdam. [327]

Jackson, Matthew O. (1991), "Bayesian implementation." *Econometrica*, 59, 461–477. [349]

Jackson, Matthew O. (1992), "Implementation in undominated strategies: A look at bounded mechanisms." *Review of Economic Studies*, 59, 757–775. [349]

Jackson, Matthew O., Thomas R. Palfrey, and Sanjay Srivastava (1994), "Undominated nash implementation in bounded mechanisms." *Games and Economic Behavior*, 6, 474–501. [326]

Kartik, Navin (2009), "Strategic communication with lying costs." *Review of Economic Studies*, 76, 1359–1395. [327]

Kartik, Navin, Marco Ottaviani, and Francesco Squintani (2007), "Credulity, lies, and costly talk." *Journal of Economic Theory*, 134, 93–116. [327]

Kartik, Navin and Olivier Tercieux (2012), "A note on mixed-Nash implementation." Unpublished paper, Columbia University. [336]

Lipman, Barton L. and Duane J. Seppi (1995), "Robust inference in communication games with partial provability." *Journal of Economic Theory*, 66, 370–405. [326, 347]

Maggi, Giovanni and Andrés Rodriguez-Clare (1995), "Costly distortion of information in agency problems." *Rand Journal of Economics*, 26, 675–689. [327]

Maskin, Eric S. (1999), "Nash equilibrium and welfare optimality." *Review of Economic Studies*, 66, 23–38. [323, 325, 331, 337, 338, 340, 343]

Maskin, Eric S. and Thomas Sjöström (2002), "Implementation theory." In *Handbook of Social Choice and Welfare*, Volume 1 (Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura, eds.), 237–288, North Holland, Amsterdam. [336]

Matsushima, Hitoshi (1988), "A new approach to the implementation problem." *Journal of Economic Theory*, 45, 128–144. [326]

Matsushima, Hitoshi (2008a), "Behavioral aspects of implementation theory." *Economics Letters*, 100, 161–164. [326, 336]

Matsushima, Hitoshi (2008b), "Role of honesty in full implementation." *Journal of Economic Theory*, 139, 353–359. [326, 336]

Milgrom, Paul R. (1981), "Good news and bad news: Representation theorems and applications." *Bell Journal of Economics*, 12, 380–391. [327]

Moore, John H. and Rafael Repullo (1988), "Subgame perfect implementation." *Econometrica*, 56, 1191–1220. [326]

Moore, John H. and Rafael Repullo (1990), "Nash implementation: A full characterization." *Econometrica*, 58, 1083–1099. [336]

Myerson, Roger B. (1982), "Optimal coordination mechanisms in generalized principal-agent problems." *Journal of Mathematical Economics*, 10, 67–81. [329]

Postlewaite, Andrew and David Wettstein (1989), "Feasible and continuous implementation." *Review of Economic Studies*, 56, 603–611. [327]

Sanver, M. Remzi (2006), "Nash implementing non-monotonic social choice rules by awards." *Economic Theory*, 28, 453–460. [335]

Sher, Itai (2010), "Persuasion and dynamic communication." Unpublished paper, University of Minnesota. [327, 347]

Singh, Nirvikar and Donald Wittman (2001), "Implementation with partial verification." *Review of Economic Design*, 6, 63–84. [327]

Spence, A. Michael (1973), "Job market signaling." *Quarterly Journal of Economics*, 87, 355–374. [324]

Submitted 2010-2-6. Final version accepted 2011-7-4. Available online 2011-7-4.