IMPLEMENTATION WITH EVIDENCE: COMPLETE INFORMATION

NAVIN KARTIK AND OLIVIER TERCIEUX

Abstract. We study full-implementation in Nash equilibrium under complete information. We generalize the canonical model (Maskin, 1977) by allowing agents to send *evidence* or discriminatory signals. A leading case is where evidence is hard information that proves something about the state of the world. In this environment, an implementable social choice rule need not be Maskin-monotonic. We formulate a weaker property, *evidence-monotonicity*, and show that this is a necessary condition for implementation. Evidence-monotonicity is also sufficient for implementation if there are three or more agents and the social choice rule satisfies two other properties—*no veto power* and *non-satiation*—that are reasonable in various settings, including "economic environments". We discuss how natural conditions on the cost of discriminatory signals yield possibility results, in contrast with traditional negative results. Additional results are provided for the case of one and two agents.

Columbia University and UC San Diego
Paris School of Economics and CNRS
E-mail addresses: nkartik@gmail.com, tercieux@pse.ens.fr.
Date: April 30, 2009.
For helpful comments, we thank David Ahn, Vince Crawford, Bob Evans, Jayant Vivek Ganguli, Vijay Krishna, Laurent Lamy, Eric Maskin, François Maniquet, Stephen Morris, Daijiro Okada, Mike Riordan, Hamid Sabourian, Joel Sobel, Joel Watson, and seminar participants at Paris School of Economics, UC San Diego, University of Cambridge, University of Michigan, and the 3rd workshop of Economic Design and Institutions in Brussel. 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).

1. Introduction

The theory of implementation is about designing decentralized mechanisms or game forms to achieve outcomes that are desired by a planner (who might be a benevolent government or social planner in a usual sense, but could also be a self-interested principal, or even represent the agents themselves at some constitutional stage). Abstractly, the planner's objectives are represented by a *social choice rule* (SCR), which specifies a set of desired outcomes in each possible state of the world. By *implementation* we mean in this paper *full-implementation*, where the concern is with ensuring that *every* decentralized outcome is desirable.[1] In order to determine which outcomes can emerge from any mechanism, one must adopt a game-theoretic solution concept. Here, we use the traditional benchmark of *Nash equilibrium*. Moreover, this paper belongs to the literature on implementation with *complete information*, where there are no information asymmetries amongst agents, so that only the planner does not know the true state.

In his seminal contribution, Maskin (1999, in circulation from 1977) provided a necessary property that any SCR must satisfy for it to be implementable in Nash equilibrium. (Hereafter, implementation without qualification refers to Nash implementation.) He also showed that this property, which we call *Maskin-monotonicity*, is sufficient for implementation when there are three or more agents and the SCR satisfies a further mild condition. While intuitive, Maskin-monotonicity is known to be quite demanding in many settings, particularly if the SCR is single-valued.[2] This has lead to a substantial research program that examines implementation under solution concepts that refine Nash equilibrium.[3] For some such refinements, permissive results have been obtained under complete information.[4] However, the robustness of these results to the introduction of small amounts of incomplete information has been recently questioned. In particular, if one requires these mechanisms to implement in environments with "almost" complete information, Maskin monotonicity is again a necessary condition.[5]

A maintained assumption in almost all of this literature is that agents can manipulate information about the state without restraint. Put differently, all messages that are or can be made available to an agent in any mechanism are assumed to be state independent, and moreover, directly payoff irrelevant. In this sense, all messages are "cheap talk", and only matter insofar as they affect the outcome determined by the mechanism. This paper aims to focus attention on this aspect of

[1]A substantial literature addresses partial-or weak-implementation, where the goal is only to ensure that at least one decentralized outcome is desirable.

[2]For instance, Saijo (1987) shows that over an unrestricted domain of preferences, a social choice function is Maskin-monotonic if and only if it is constant. In addition, if utility can be measured on a cardinal scale, SCRs that make interpersonal comparisons (such as those based on utilitarian or related welfare criteria) are generally not Maskin-monotonic, even on restricted domains.

[3]An alternative approach to the use of refinements of Nash equilibrium is to weaken the notion of implementation, in particular to approximate or virtual implementation (Matsushima, 1988; Abreu and Sen, 1991). A now standard critique of this approach is that — even if it only occurs with small probability — the mechanism may provide an outcome arbitrarily inefficient, unfair, or, in any meaningful sense, far from the desired alternative.

[4]For example, Moore and Repullo (1988) and Abreu and Sen (1990) analyze subgame perfection while Palfrey and Srivastava (1991) study undominated Nash equilibrium.

[5]See Aghion et al. (2007) and Kunimoto and Tercieux (2009) on the robustness of mechanisms using subgame perfection and Chung and Ely (2003) for mechanisms using undominated Nash equilibrium.

the implementation problem, and to generalize the set of environments under consideration. To motivate our treatment, here are two examples:

(1) A principal is concerned with dividing a fixed sum of money between agents as some function of their individual output. If only asked to send cheap-talk messages about their output, agents could claim anything they want. But the principal could also request agents to provide physical verification of their output. In this case, an agent would be unable to claim that his output is larger than it in fact is, but he could claim that it is less, just by not providing all of it.[6] If it is costless to provide the physical verification, the setting is one of certifiability or hard information. If instead agents bear costs as a function of how much output they carry to the principal's court (so to speak), then we have a more complicated costly signaling instrument.

(2) Consider an income taxation problem, where the planner cannot observe agents' income. Agents who are legally employed have a document that stipulates their income (a W2 form). If requested by a mechanism to submit such a document, a legally employed agent can either costlessly submit his true document, or fabricate a false document at some cost. The cost may take the form of a fixed cost or possibly vary with the degree of falsification. Illegally employed agents must fabricate a document, or admit to not possessing one. In this example, the informational signal of a submitted document is heterogeneous across agents.

The key element in these examples is that some messages are only feasible in some states of the world, or messages have differential costs in different states. Since this theme naturally arises in numerous settings, we believe it is important to study implementation in a framework that accommodates this feature. Accordingly, this paper adds "*evidence*" to an otherwise standard implementation environment. The crucial feature that distinguishes evidence from standard (cheap-talk) messages is that a piece of evidence is a *discriminatory signal* about the state of the world. A mechanism in this setting can not only request agents to send cheap-talk messages as usual, but also to submit evidence. Naturally, the planner will generally benefit from the availability of evidence; our interest is in understanding exactly how and to what extent. In particular, which SCRs are implementable given some evidence structure, and what evidence structure is needed to make a particular SCR implementable?

We begin in Section 2 by considering environments where evidence takes the form of hard or non-manipulable information: an agent cannot produce evidence that she does not in fact possess. Formally, in each state θ, each agent i has a set of evidence, E_i^θ, and can submit evidence e_i if and only if $e_i \in E_i^\theta$. Providing any feasible evidence is costless. We place no restriction on the evidence structure, $\left\{E_i^\theta\right\}_{i,\theta}$, so that the standard environment is a special case where a player's evidence set does not vary with the state.

A simple but significant observation is that SCRs that are not Maskin-monotonic can be implementable under some evidence structures. We identify a necessary condition, *evidence-monotonicity*,

[6] Postlewaite and Wettstein (1989) study implementation of the Walrasian correspondence in an exchange economy setting where agents' messages about their endowments have a similar assumption.

that a SCR must satisfy for it to be implementable. Evidence-monotonicity is a joint condition on the SCR, agents' preferences, and the evidence structure $\{E_i^\theta\}_{i,\theta}$. To provide some sense of the condition, suppose the SCR is single-valued and the evidence structure is *normal* in that if a player can submit evidence a or b, he can also submit both a and b (Bull and Watson, 2007). This can be interpreted as a "no time constraints" assumption on the evidence structure. In this case, evidence-monotonicity is satisfied if and only if for any pair of states over which the SCR violates Maskin-monotonicity, there is some player whose evidence set differs between the two states. Intuitively, if all agents' evidence is identical in two states, then a SCR that does not respect Maskin-monotonicity over these two states cannot be implemented. (Evidence-monotonicity is more demanding when the restriction to normal evidence is dropped.) Thus, evidence-monotonicity weakens and generalizes Maskin-monotonicity, reducing to Maskin-monotonicity if, and generally only if, there is no evidence available.

A central result of this paper is a partial converse: any evidence-monotonic SCR is implementable if there are three or more players and the SCR satisfies two properties called *no veto power* and *non-satiation*. No veto power was introduced by Maskin (1999) and requires that at any state, if an outcome is top-ranked by all-but-one or all agents, then the outcome is chosen by the SCR at that state. Non-satiation requires in any state, no player's most-preferred outcome should be in the range of the SCR, i.e. no outcome ever chosen by the SCR should ever be a player's top-ranked outcome.[7] While these two properties are restrictive, they are satisfied in many settings of interest. For example, both hold whenever the planner can make arbitrarily small transfers of a private good between players and the SCR does not give all of the private good to a single player (and there are three or more players). Thus, in *economic environments*, as defined by Moore and Repullo (1988) (and three or more players), evidence-monotonicity is a tight characterization of when a SCR is implementable.

That implementability of a SCR turns on evidence-monotonicity produces a number of corollaries. We introduce a notion of when a state and an event (i.e., a set of states) are *distinguishable* under a given evidence structure. We show that a SCR satisfies evidence-monotonicity if and only if appropriate state-event pairs are distinguishable. This yields a partial order on evidence structures in terms of how *informative* they are, so that a "more informative" evidence structure implies that a larger set of SCRs are evidence-monotonic. There is (an equivalence class of) maximally informative evidence structures, under any of which every SCR is evidence-monotonic, hence implementable under the aforementioned conditions. It is shown that a normal evidence structure is maximally informative if and only if for every pair of states there is some player whose evidence set differs between the two states.

We also use the notion of distinguishability to create a partial ranking of SCRs, so that a SCR is "easier" to implement if it is evidence-monotonic for a larger set of evidence structures. Naturally, Maskin-monotonic SCRs are the easiest to implement.

[7] A significantly weaker version of this condition is in fact sufficient, as we make clear in the analysis, but is more involved to state.

In Section 3, we generalize the environment so that a player can fabricate any piece of evidence in any state of the world. We posit a cost function $c_i(e_i, \theta)$ which specifies the cost for player i of producing evidence e_i when the true state is θ. A player's preferences are represented by $u_i(a, \theta) - c_i(e_i, \theta)$, where a is the outcome. The case of hard information essentially corresponds to the special case where $c_i(\cdot, \cdot) \in \{0, \infty\}$, with $c_i(e_i, \theta) = 0$ iff $e_i \in E_i^\theta$,[8] whereas now we consider arbitrary cost functions. A generalization of evidence-monotonicity to *cost-monotonicity* is formulated, and we derive a sense in which cost-monotonicity is necessary for a SCR to be implementable in this setting. We also show that if the planner can make sufficiently large transfers off the equilibrium path, any cost-monotonic SCR is implementable when there are three or more players.

Under some cost structures, any SCR is cost-monotonic. One striking case is where at least one player has a "small preference for honesty." Roughly, this is an assumption on the cost structure so that an honest player would prefer to tell the truth (in terms of a direct message about the state) rather than lying if he knows that the outcome is not much affected by this decision. A result with a similar flavor is proved by Matsushima (2008) using a different approach that follows the virtual implementation literature.

Our analysis makes clear that the precise form of evidence fabrication costs significantly affect the possibility of implementation. In many settings, the planner may have the power to design the cost structure (or part of it), perhaps at social cost. For instance, the degree of difficulty or technological cost of falsifying the income document in the second motivating example above could be modified by the planner. From the perspective that influencing evidence fabrication costs is an additional instrument for the planner, our analysis yields insight into how useful such an instrument can be. Moreover, in the case of hard information, our ranking of evidence structures in terms of their informativeness may also be useful in this vein.

In proving sufficiency of evidence- or cost-monotonicity (in conjunction with the other conditions) for implementation, we explicitly construct implementing mechanisms.[9] The mechanisms we use build upon the "integer games" that are standard in the Nash implementation literature. This reflects our central objective to revisit the classic results of Maskin (1999) in richer environments with evidence. While we acknowledge that integer games are unappealing in some respects, the usual justification for considering such mechanisms applies in the present context: it puts an upper bound on what is implementable, and moreover, permits the construction of a single mechanism that achieves the goal in every environment in which implementation is possible. We conjecture that the way in which we utilize evidence to achieve implementation could be adapted to, for example, "bounded" mechanisms (Jackson, 1992).

This paper contributes to the relatively small but growing literature on mechanism design with evidence. Most of these studies concern partial-implementation and hard information. An early

[8] There is one subtlety though, which concerns dynamic mechanisms, discussed in Section 2.9.

[9] Kartik and Tercieux (2008b) prove that if for each player and each state of the world, utility functions over outcomes are bounded, then implementation in mixed strategy Nash equilibria is possible under the usual assumptions in the standard setting with no evidence. Although only pure strategy equilibria are considered here, we can show that under a similar boundedness assumption, the current results would go through when considering mixed strategy Nash equilibria.

reference is Green and Laffont (1986), and a sample of more recent work is Bull (2008), Bull and Watson (2004, 2007), Deneckere and Severinov (2008), Glazer and Rubinstein (2004, 2006), Sher (2008), and Singh and Wittman (2001). We should also note the work of Myerson (1982), who studies partial-implementation in a setting where agents have private decision domains that the principal/planner cannot control, which is related to our interest in costly evidence fabrication.

Closest to our work is a recent paper by Ben-Porath and Lipman (2008), who also tackle complete information full-implementation with evidence. While our results were 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 and complementary. In terms of setting, they restrict attention to hard information and assume that the evidence technology satisfies *normality* (Bull and Watson, 2007) or *full reports* (Lipman and Seppi, 1995). We do not impose this condition on the evidence structure, and also consider costly evidence fabrication. In terms of results, their primary interest is in subgame perfect implementation, whereas we are entirely concerned with Nash implementation. Since the central difficulty in implementation theory is eliminating "bad" equilibria, our results with a weaker solution concept are stronger in this regard. The tradeoff is that our sufficiency results are proved using integer games, as already noted.[10]

Stepping outside mechanism design, there is a large literature on strategic communication games where evidence plays an important role. The introduction of hard evidence into implementation may be considered analogous to moving from communication games of cheap talk (Crawford and Sobel, 1982) to those verfiable/certifiable information (Milgrom, 1981; Okuno-Fujiwara et al., 1990; Lipman and Seppi, 1995; Forges and Koessler, 2005). Costly evidence fabrication has been studied in communication games by Kartik et al. (2007) and Kartik (2008), and in contract settings by, for example, Maggi and Rodriguez-Clare (1995).

2. Hard Evidence

2.1. **The Model.** There is a non-empty set of *agents* or players, $I = \{1, \dots, n\}$, a set of allocations or *outcomes*, A, a set of *states* of the world, Θ, and a vector of payoff functions, $\{u_i\}_{i=1}^n$, where each $u_i : A \times \Theta \to \mathbb{R}$.[11] To avoid trivialities, $|A| > 1$ and $|\Theta| > 1$. All agents are assumed to know the value of the state, whereas the planner does not. The planner's objectives are given by a *social choice rule* (SCR), which is a correspondence $f : \Theta \rightrightarrows A$; a social choice function is a single-valued SCR.

So far, the setting is standard. Let us now describe how evidence enters the model. In each state of the world, θ, agent i is endowed with a set of evidence, E_i^θ, which we assume to be non-empty

[10] Ben-Porath and Lipman's Theorem 2, which provides sufficient conditions for one-stage subgame perfect implementation (hence, Nash implementation) also uses an integer game and is strictly subsumed by our results.

[11] It is common to focus on just ordinal preferences in each state. We could do with ordinal preferences insofar as only pure strategy Nash equilibria are considered. For the extension to mixed Nash equilibria (see Remark 5), we would require that in any state θ, $u_i(\cdot, \theta)$ is an expected utility representation of player i's preferences over lotteries on outcomes. Moreover, our formulation allows the possibility of agents having cardinal valuations over outcomes.

without loss of generality.[12] Denote $E_i := \bigcup_\theta E_i^\theta$, $E^\theta := E_1^\theta \times \cdots \times E_n^\theta$ and $E := E_1 \times \cdots \times E_n$. We refer to $\mathcal{E} := \left\{E_i^\theta\right\}_{i,\theta}$ as the *evidence structure*. The interpretation is that any $e_i \in E_i$ is a document, piece of evidence, non-falsifiable claim, etc., that is only available to agent i in the set of states $\{\theta : e_i \in E_i^\theta\}$. The traditional setting is the special case where for all i, $E_i^\theta = E_i$ for all θ. For convenience, we will often refer to this case as a setting with "no evidence", even though E_i is non-empty for all i.

In standard implementation theory, a mechanism consists of a (cheap-talk) message space and an outcome function which specifies an outcome for every profile of messages. In the current setting, a mechanism can also take advantage of the evidence that agents might possess by conditioning the outcome on the evidence that is submitted. 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.[13]

Remark 1. Our definition of a mechanism entails that every agent must submit some evidence. This is without loss of generality: if one wants to allow agent i to have the choice of not submitting evidence, we would just add a new element, say e_i, to E_i^θ for all θ, and interpret this e_i as submitting no evidence.

Remark 2. In some cases, it may be reasonable to assume that the planner can prohibit agents from submitting certain evidence. Our definition does not allow *forbidding evidence*. This issue is discussed in Section 2.9.

A mechanism (M, g) induces a strategic game form in each state of the world, where a pure strategy for player i in state θ is $s_i \in M_i \times E_i^\theta$. Let $NE(M, g, \theta)$ be the set of pure strategy Nash equilibria (NE hereafter) of this game and $O(M, g, \theta) := \{a \in A : \exists (m, e) \in NE(M, g, \theta) \text{ s.t. } g(m, e) = a\}$ be the set of equilibrium allocations. For simplicity, we restrict attention to pure strategy equilibria in the main text; our results can be extended to mixed strategy Nash equilibria as discussed in Remark 5 following Theorem 2.

We are interested in (full-)implementation in Nash equilibria, defined as follows.

Definition 1 (Implementation)**.** *A mechanism (M, g) implements the SCR f if $\forall \theta$, $f(\theta) = O(M, g, \theta)$. A SCR is implementable if there exists a mechanism that implements it.*

Before turning to the analysis, we wish to emphasize that the current framework allows for different states to have identical profiles of agents' preferences. In contrast, in standard implementation theory, it is common to identify a state of the world with a profile of preferences. This is because in the standard model the planner cannot implement different outcomes unless some agent's preference ranking changes. Introducing evidence into the picture changes this perspective quite

[12] Since we will permit the planner to make available cheap-talk messages, one can always assume that for each i, there is some $e_i \in \bigcap\limits_\theta E_i^\theta$, which is just a cheap-talk message.

[13] Strictly speaking, letting the domain of g be $M \times E$ is excessive in the sense that some $e \in E$ may not be feasible in any state. It will be irrelevant what the outcome $g(\cdot, e)$ is for any such e; for later purposes, it is convenient to let the entire evidence structure be included in the domain of g.

dramatically (as will be seen), so that a state can reflect anything relevant to the planner beyond just the agents' preferences. For example, consider legal settings where neither a plaintiff's nor defendant's preferences over outcomes may vary with the state, but the state of world reflects the degree of damages that has been caused by the defendant and is thus relevant to the planner.[14]

In what follows, the symbols $\wedge$ and $\vee$ are sometimes used to respectively stand for "and" and "or".

2.2. **Evidence-Monotonicity: Necessity.** Maskin (1999) showed that in settings without evidence, SCRs must satisfy a monotonicity condition to be implementable. We will refer to his now classic condition as *Maskin-monotonicity.*[15] In settings with non-trivial evidence, Maskin-monotonicity is no longer necessary for a SCR to be implementable.

Example 1. *Let $E_i = \Theta$ and $E_i^\theta = \{\theta\}$ for all i, θ. This can be interpreted as agents never being able to misrepresent the state of the world. In this case, no matter the specification of agents' preferences, any social choice function f can be implemented by a mechanism with an arbitrary message space, M, and the outcome function $g(m, (\theta, \dots, \theta)) = f(\theta)$.*

The key to our analysis is the following generalization of Maskin-monotonicity.

Definition 2 (Evidence-monotonicity)**.** *A SCR f is evidence-monotonic provided that for all θ and all $a \in f(\theta)$, there exists $e^*_{\theta,a} \in E^\theta$ such that for any θ', if*

$$\forall i \in I, \forall b \in A\text{: } [u_i(a,\theta) \geq u_i(b,\theta)] \implies [u_i(a,\theta') \geq u_i(b,\theta')] \tag{*}$$

and

$$e^*_{\theta,a} \in E^{\theta'} \text{ and } E^{\theta'} \subseteq E^\theta, \tag{**}$$

then $a \in f(\theta')$.

To get some intuition for the definition, note that without (**), it is just Maskin-monotonicity. In a setting without evidence, evidence-monotonicity simplifies to Maskin-monotonicity because when $E_i^\theta = E_i$ for all θ, i, condition (**) is satisfied no matter the choice of $\left\{e^*_{\theta,a}\right\}_{\theta,a}$. Plainly, any Maskin-monotonic SCR is evidence-monotonic regardless of the evidence structure. In general, (**) will make it possible for non-Maskin-monotonic SCRs to be evidence-monotonic: for instance, in Example 1, (**) is not satisfied for any $\theta \neq \theta'$ (since $E^\theta = \{(\theta, \dots, \theta)\}$ for all θ), hence *every* SCR is evidence-monotonic, regardless of agents' preferences.

For any θ and $a \in f(\theta)$, $e^*_{\theta,a}$ can be thought of as the evidence profile that agents are supposed to provide the planner in state θ to indicate that a should be the outcome. Loosely speaking, one

[14] We should also note that if two states differ *only* in agents' preferences (and all other relevant aspects are identical), one might argue that no agent's evidence set should differ between the two states. This depends on whether an agent can provide some proof about his preferences.

[15] A SCR is Maskin-monotonic provided that for all θ, θ' and $a \in f(\theta)$, if

$$\forall i \in I, \forall b \in A\text{: } [u_i(a,\theta) \geq u_i(b,\theta)] \implies [u_i(a,\theta') \geq u_i(b,\theta')]$$

then $a \in f(\theta')$.

should choose $e^*_{\theta,a}$ to be an evidence profile that is "most informative" about state θ with respect to the other states that outcome a is also socially desirable in. 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 for any $a \in f(\theta)$, one can take $e^*_{\theta,a}$ to be this evidence profile. This is illustrated in the following example:

Example 2. *Consider the first motivating example in the introduction. A principal is concerned with dividing a fixed sum of money to agents as some function of their individual production. A state of nature θ is a profile of units of output, i.e. $\theta = (\theta_1, ..., \theta_n) \in \Theta = \mathbb{R}^n_+$. Each agent can explicitly show his true output or some subset of it, hence an agent is unable to claim that his output is larger that it in fact is, but he can claim that it is less. Formally, $E_i = \Theta$ and $E_i^\theta = [0, \theta_i]$ for all i, θ. We will now show that any SCR is evidence-monotonic in this setting.*

For any θ and $a \in f(\theta)$, let $e^*_{\theta,a} = \theta$. 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 $e^*_{i,\theta,a} = \theta_i \notin [0, \theta'_i] = E_i^{\theta'}$ and so (**) is violated. Second, if $\theta'_i > \theta_i$ then $\theta'_i \in E_i^{\theta'}$ but $\theta'_i \notin [0, \theta_i] = E_i^\theta$, hence (**) is violated.

Our first result is that evidence-monotonicity is a necessary condition for implementation.

Theorem 1. *If a SCR is implementable, it is evidence-monotonic.*

Proof. Assume f is implemented by a mechanism (M, g). Then for each θ and $a \in f(\theta)$, there exists $\left(m^*_{\theta,a}, e^*_{\theta,a}\right) \in M \times E^\theta$ that is a NE at θ such that $g\left(m^*_{\theta,a}, e^*_{\theta,a}\right) = a$. Consider any θ' satisfying (*) and (**). Since $\left(m^*_{\theta,a}, e^*_{\theta,a}\right)$ is a NE at θ,

$$u_i\left(g\left(m^*_{\theta,a}, e^*_{\theta,a}\right), \theta\right) \geq u_i\left(g\left(m'_i, m^*_{-i,\theta,a}, e'_{i,\theta}, e^*_{-i,\theta,a}\right), \theta\right)$$

for all i, $m'_i \in M_i$, $e'_{i,\theta} \in E_i^\theta$. By (*),

$$u_i\left(g\left(m^*_{\theta,a}, e^*_{\theta,a}\right), \theta'\right) \geq u_i\left(g\left(m'_i, m^*_{-i,\theta,a}, e'_{i,\theta}, e^*_{-i,\theta,a}\right), \theta'\right)$$

for all i, $m'_i \in M_i$, $e'_{i,\theta} \in E_i^\theta$. Since (**) stipulates $e^*_{\theta,a} \in E^{\theta'}$ and $E^{\theta'} \subseteq E^\theta$, we have

$$u_i\left(g\left(m^*_{\theta,a}, e^*_{\theta,a}\right), \theta'\right) \geq u_i\left(g\left(m'_i, m^*_{-i,\theta,a}, e'_{i,\theta}, e^*_{-i,\theta,a}\right), \theta'\right)$$

for all i, $m'_i \in M_i$, $e'_{i,\theta} \in E_i^{\theta'}$. Therefore, $\left(m^*_\theta, e^*_{\theta,a}\right)$ is a NE at θ', and $g\left(m^*_{\theta,a}, e^*_{\theta,a}\right) = a \in f(\theta')$. □

The example below illustrates why the Theorem rules out implementation of certain SCRs also and provides further insight into evidence-monotonicity.

Example 3. *Let $n = 4$, $\Theta = \{w, x, y, z\}$, $A = \{a, b, c, d\}$, and for all i, $E_i^w = \{\alpha, \beta\}$, $E_i^x = \{\alpha, \beta\}$, $E_i^y = \{\alpha\}$, and $E_i^z = \{\alpha\}$. For all i, preferences satisfy: $u_i(b, w) > u_i(c, w)$; $u_i(c, x) > u_i(b, x)$; $u_i(b, y) > u_i(c, y)$; $u_i(b, z) > u_i(c, z)$. In addition, for agents 1 and 2, in all states, a is* *the unique top-ranked alternative while d is the unique bottom-ranked alternative; analogously, for agents 3 and 4, in all states, d is the unique top-ranked alternative while a is uniquely bottom-ranked.*

Consider the SCR f where $f(w) = f(z) = b$ and $f(x) = f(y) = c$. f is not evidence-monotonic, because in Definition 2 we must set $e^*_{y,c} = e^*_{z,c} = \alpha\alpha\alpha\alpha$, and since no agent's preferences between b and c switch from state y to state z, evidence-monotonicity requires that $c \in f(z)$ because $f(y) = c$. Thus, by Theorem 1, f is not implementable.

On the other hand, consider the SCR f^* where $f^*(w) = b$ and $f^*(x) = f^*(y) = f^*(z) = c$. Even though f^* is not Maskin-monotonic (because $b = f^*(w) \neq f^*(y) = c$), one can see that f^* is evidence-monotonic by using $e^*_{w,b} = e^*_{x,b} = \beta\beta\beta\beta$ and $e^*_{y,c} = e^*_{z,c} = \alpha\alpha\alpha\alpha$ in Definition 2. We will show subsequently that f^* is in fact implementable.

2.3. **Evidence-Monotonicity: Sufficiency.** In this section, we show that under some conditions, evidence-monotonicity is also sufficient for implementation when $n \geq 3$. First, we observe that evidence-monotonicity by itself is not generally sufficient, since it reduces to Maskin-monotonicity without evidence, and it is well-known that Maskin-monotonicity is not sufficient for implementation in the standard setting (even with $n \geq 3$). Maskin introduced the following notion.

Definition 3 (No veto Power)**.** *A SCR f satisfies No veto power (NVP) if for all $\theta \in \Theta$ and $a \in A$,*

$$\left[\left|\{i : a \in \arg\max_{b\in A} u_i(b,\theta)\}\right| \geq n-1\right] \implies a \in f(\theta).$$

NVP says that if at state θ an outcome is top ranked by $n-1$ individuals, then the last individual cannot prevent this outcome from being in the socially desired set $f(\theta)$, i.e. it cannot be "vetoed".

From Maskin (1999), we know that when $n \geq 3$ and $E_i^\theta = E_i$ for all i, θ (a setting without evidence), a SCR is implementable if it satisfies evidence-monotonicity and NVP. This is no longer the case with non-trivial evidence structures:

Example 4. *$n = 4$. $\Theta = \{\theta_1, \theta_2\}$. $E_1^{\theta_1} = \{x\}$, $E_1^{\theta_2} = \{x, y\}$; for $i \neq 1$, $E_i^{\theta_1} = E_i^{\theta_2} = \{z\}$. $A = \{a, b\}$. For all θ and $i \in \{1, 2\} : u_i(b, \theta) > u_i(a, \theta)$. For all θ and $i \in \{3, 4\} : u_i(a, \theta) > u_i(b, \theta)$. The SCR is $f(\theta_1) = b$ and $f(\theta_2) = a$. This SCR is evidence-monotonic, as seen by using $e^*_{\theta_1,b} = (xzzz)$ and $e^*_{\theta_2,a} = (yzzz)$ in Definition 2. No veto power is also satisfied since for all θ, there is no outcome that is most-preferred by 3 or more players. Yet, the SCR is not implementable. To see this, suppose that it is implemented by some mechanism (M, g). Then $g(m^*, e^*) = b$ for some $(m^*, e^*) \in NE(M, g, \theta_1)$, because $b = f(\theta_1)$. This implies that there is no unilateral deviation for either player 3 or player 4 from (m^*, e^*) that leads to outcome a. Since $(m^*, e^*) \in M \times E^{\theta_1} \subseteq M \times E^{\theta_2}$, (m^*, e^*) is a strategy profile available at state θ_2. This implies that $(m^*, e^*) \in NE(M, g, \theta_2)$, hence $b \in O(M, g, \theta_2)$, a contradiction.*

Therefore, some additional condition is needed to guarantee implementability. To get some insight on the assumption we make below and the role it plays, notice that the difficulty here is that even though player 1 is able to disprove θ_1 when the true state is θ_2 (by submitting evidence y), he has no incentive to do so. Suppose that we modify the setting so that player 1 is "non-satiated" by any outcome in the range of f. For instance, assume the existence of an outcome c such that

$u_1(c,\theta) > u_1(b,\theta)$ for all θ. Now the mechanism could have player 1 getting outcome c when he deviates from (m^*, e^*) by submitting $e_1 = y$. In this case, $(m^*, e^*) \in NE(M, g, \theta_1)$, will not be an equilibrium anymore at state θ_2, since at this state player 1 will indeed deviate and announce y to get outcome c. Note that such a deviation is not available at state θ_1.

Accordingly, we introduce the following condition.

Definition 4 (Non-satiation)**.** *A SCR f satisfies non-satiation if for all $i \in I$, $\theta \in \Theta$, and $a \in \bigcup_{\theta' \in \Theta} f(\theta')$, there exists $\tilde{a} \in A$ such that $u_i(\tilde{a}, \theta) > u_i(a, \theta)$.*

Non-satiation requires that the SCR never choose an outcome that is any agent's ideal outcome in some state of the world. Note that Non-satiation does not preclude Pareto-efficiency (fn. 18 gives an example). As we will make clear, substantially weaker versions of this condition will suffice for our analysis, but for expositional simplicity it is convenient to begin with the above statement.

Both no veto power and non-satiation are reasonable in many settings with three or more players.[16] In particular, they are both automatically satisfied in so-called "economic environments" where there is a divisible private good which is positively valued by all agents, and the SCR does not ever allocate all of the private good to a single agent (cf. Moore and Repullo, 1988, condition EE1). They are also satisfied if the planner can augment outcomes with arbitrarily small transfers that are never to be used in equilibrium, even if he must maintain off-the-equilibrium-path budget balance (cf. Benoît and Ok, 2008; Ben-Porath and Lipman, 2008; Sanver, 2006).[17] There are also interesting pure public goods problems without transfers where both conditions hold.[18] On the other hand, if transfers are not feasible, there are important environments that would not satisfy the two conditions. For example, classic voting problems would violate non-satiation (but cf. Remark 7 below).

Remark 3. When non-satiation holds, NVP can only hold if in any state, there is no alternative that is top-ranked by $n-1$ or more agents. Thus, non-satiation and NVP together require "enough disagreement" in any state.

[16] Instead of NVP, it suffices to assume Moore and Repullo's (1990) weaker condition of *restricted veto power*, but we retain NVP here for ease of exposition. We should note that both non-satiation and, especially, NVP are overly demanding when there are only one or two agents. Sections 2.7 and 2.8 address implementability when $n < 3$ without invoking these conditions.

[17] To be more precise: consider an underlying outcome space, $\tilde{A}$, with each agent having a payoff function $\tilde{u}_i : \tilde{A} \times \Theta \to \mathbb{R}$, and a SCR $\tilde{f}$. (Note that $\tilde{A}$ itself may include transfers.) Suppose that the planner can impose an additional vector of transfers $(t_1, \ldots, t_n) \in X \subseteq \mathbb{R}^n$, and each agent values his personal transfer quasi-linearly. Assume that the space of possible transfers satisfies three weak properties: i) $(0, \ldots, 0) \in X$; ii) for all i, there exists $(t_1, \ldots, t_i, \ldots, t_n) \in X$ with $t_i > 0$; iii) 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 would be $X = \{(t_1, \ldots, t_i, \ldots, t_n) \in \mathbb{R}^n : \sum_i t_i = 0, |t_i| \leq k\}$ 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 extended outcome space $A = \tilde{A} \times \mathbb{R}^n$, an extended payoff function for each agent $u_i : A \times \Theta \to \mathbb{R}$ where $u_i(\tilde{a}, t_1, \theta) = u_i(a, \theta) + x$, and an extended SCR f derived from $\tilde{f}$ by setting $f(\theta) = \left(\tilde{f}(\theta), 0, \ldots, 0\right)$. Assuming that $n \geq 3$, this extended environment satisfies both non-satiation and no veto power.

[18] For example, the decision concerns how much of a public good to produce, $a \in [0,1]$. Agents are of one of two types, either they want $a = 1$ or $a = 0$. But, motivated by fairness, the planner's objective is to choose the mean of the desired levels. So long as $n \geq 4$ and there is always at least two agents of each preference type, both non-satiation and no veto power are satisfied.

The next result shows that if there are three or more players, and the two aforementioned conditions hold, evidence-monotonicity is sufficient for implementation. The mechanism we use in the proof builds on the integer game construction that is standard in the literature.

Theorem 2. *Assume $n \geq 3$ and let f be a SCR satisfying no veto power and non-satiation. If f is evidence-monotonic, it is implementable.*

Proof. For each θ and $a \in f(\theta)$, let $e^*_{\theta,a}$ be the evidence profile per Definition 2. We will construct a mechanism that implements f. Fix, for all i, $M_i = \Theta \times A \times \mathbb{N}$, and define g via the following three rules:

(1) If $m_1 = \cdots = m_n = (\theta, a, k)$ with $a \in f(\theta)$, and $e = e^*_{\theta,a}$, then $g(m,e) = a$.

(2) If $\exists i$ s.t. $\forall j \neq i$, $e_j = e^*_{j,\theta,a}$ and $m_j = (\theta, a, k)$ with $a \in f(\theta)$, and either $m_i = (\theta', b, l) \neq (\theta, a, k)$ or $e_i \neq e^*_{i,\theta,a}$ then

   a) if $e_i \in E_i^\theta$, then $g(m,e) = \begin{cases} b \text{ if } u_i(a,\theta) \geq u_i(b,\theta) \\ a \text{ if } u_i(b,\theta) > u_i(a,\theta) \end{cases}$.

   b) if $e_i \notin E_i^\theta$, then $g(m,e) = b$.

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

Step 1. We first show that for any θ and any $a \in f(\theta)$, $a \in O(M, g, \theta)$. It suffices to show that at state θ, for any $k \in \mathbb{N}$, each agent i playing $m_i = (\theta, a, k)$ and sending $e_i = e^*_{i,\theta,a}$ is a NE, since by rule (1) of the mechanism, this results in outcome a. If some agent deviates from this strategy profile, rule (2) of the mechanism applies. There is no profitable deviation for any player to rule (2a) since any such deviation yields a weakly worse outcome. A deviation to rule (2b) is not feasible.

For the remainder of the proof, suppose the true state is θ', and (m,e) is a NE. We must show that $g(m,e) \in f(\theta')$.

Step 2. Assume (m,e) falls into rule (2). Note that any player $j \neq i$ can deviate and by announcing an integer large enough get any outcome he wants. Since (m,e) is a NE, $g(m,e)$ must be player j's most-preferred outcome in state θ for any $j \neq i$. NVP then implies that $g(m,e) \in f(\theta')$. A similar argument applies if (m,e) falls into rule (3) as well.

Step 3. It remains to consider (m,e) falling into rule (1). Here $m_1 = \cdots = m_n = (\theta, a, k)$ with $g(m,e) = a \in f(\theta)$ and for all i, $e_i = e^*_{i,\theta,a}$. Observe that any player i can deviate into rule (2a) while producing the same evidence (say, by changing his integer announcement), and hence can induce any outcome in the set $L(a,\theta) := \{b : u_i(a,\theta) \geq u_i(b,\theta)\}$. Since (m,e) is a NE, this implies that for any agent i and outcome $b \in L(a,\theta)$, $u_i(a,\theta') \geq u_i(b,\theta')$, and consequently condition (*) is satisfied. If (**) also holds, evidence-monotonicity implies that $a \in f(\theta')$, and we are done. To see that (**) must hold, suppose to contradiction that it does not. Then there is some agent i such that $\left(e^*_{i,\theta,a} \notin E_i^{\theta'}\right)$ or $\left(E_i^{\theta'} \nsubseteq E_i^\theta\right)$. Since $e_i = e^*_{i,\theta,a}$ and the true state is θ', we have $e^*_{i,\theta,a} \in E_i^{\theta'}$, and consequently $E_i^{\theta'} \nsubseteq E_i^\theta$. This implies that there is some $\tilde{e}_i \in E_i^{\theta'}$ such that $\tilde{e}_i \notin E_i^\theta$, and by non-satiation, agent i can profitably deviate into rule (2b), contradicting (m,e) being a NE. □

To see an illustration of the Theorem, return to Example 3. The SCR f^* defined there satisfies both NVP (because there is no alternative that is top-ranked in any state by more than 2 players) and non-satiation (because for each player and state, either a or d is uniquely top-ranked, but a and d are not in the range of f^*). Since we argued that f^* is evidence-monotonic, it is implementable by Theorem 2.

Remark 4. For evidence structures as in Examples 1 and 2, where every SCR is evidence-monotonic, Theorem 1 implies that any SCR is implementable if $n \geq 3$ and non-satiation and no veto power hold.

Remark 5. The mechanism used in the proof of Theorem 2 does not work when mixed Nash equilibria are considered. Kartik and Tercieux (2008b) show that in the standard setting without evidence, any Maskin-monotonic SCR can be implemented in mixed Nash equilibria (under NVP and $n \geq 3$) so long as for each player i and state θ, $u_i(\cdot, \theta)$ is bounded from below; see also Maskin and Sjöström (2002, Section 4.3). These arguments can be adapted to the current setting with evidence, extending Theorem 2 to mixed Nash equilibria so long as utilities are bounded in each state.

Remark 6. Neither no veto power nor non-satiation are necessary for implementability (even when $n \geq 3$). This is readily seen by considering evidence structures as in Example 1.

Remark 7. As stated, Theorem 2 is not a strict generalization of Maskin's (1999) classic sufficiency result, because our non-satiation assumption restricts the domain even when evidence is not available. This can be remedied as follows.

For each ordered pair of states (θ, θ'), let $D(\theta, \theta')$ be the set of agents who can disprove θ at θ'. Formally,

$$D(\theta, \theta') := \left\{ i \in I : E_i^{\theta'} \nsubseteq E_i^{\theta} \right\}.$$

We say that a SCR f satisfies *weak non-satiation* if for all $\theta, \theta' \in \Theta$ and $a \in f(\theta)$,

$$\left[D(\theta, \theta') \neq \emptyset \right] \Longrightarrow \left[\exists i \in D(\theta, \theta') \text{ and } \tilde{a} \in A \text{ s.t. } u_i(\tilde{a}, \theta') > u_i(a, \theta') \right].$$

It is immediate to verify that the proof of Theorem 2 goes through if we substitute non-satiation with weak non-satiation — indeed, weak non-satiation is precisely the property needed in Step 3 of the proof. Note that weak non-satiation trivially holds in standard setting without evidence, since $D(\theta, \theta') = \emptyset$ for all θ, θ' in the standard setting. Thus, this version of the Theorem subsumes Maskin's sufficiency result. Moreover, it also points out that the full strength of the non-satiation assumption is unnecessary. What is needed is that there be some outcome that can be used to reward a player who disproves what others' claim is the true state.[19]

2.4. **Evidence-monotonicity and Distinguishability.**

[19]Consider a modification of Example 3, where outcome d is deleted from the outcome space, but everything else stays the same. Non-satiation no longer holds because for players 3 and 4, b and c are top-ranked in some state and also are also in the range of the f^*. So Theorem 2 would not apply per se. But, weak non-satiation holds because for all θ, θ', if $D(\theta, \theta') \neq \emptyset$ then $1 \in D(\theta, \theta')$, and for player 1, outcome a is always uniquely top-ranked while $a \notin range(f^*)$. Thus, f^* is implementable.

2.4.1. *Normal Evidence Structures.* Theorems 1 and 2 place no restriction on the evidence structure. The key condition they identify for implementation, evidence-monotonicity, is therefore somewhat abstract. It can be simplified considerably if the evidence structure has the following property that is often assumed in settings with hard information.

Definition 5 (Normality)**.** *The evidence structure is normal or satisfies normality if for all i and θ, there is some $\bar{e}_{i,\theta} \in E_i^\theta$ such that $\left[\bar{e}_{i,\theta} \in E_i^{\theta'} \implies E_i^\theta \subseteq E_i^{\theta'}\right]$.*

The formulation of normality 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 a maximal or summary evidence because it proves by itself what agent i could prove by jointly sending all his available evidence. Thus it is equivalent to the *full reports condition* of Lipman and Seppi (1995), and is somewhat weaker than Green and Laffont's (1986) *nested range condition* in their "direct mechanism" setting.[20] Ben-Porath and Lipman (2008) assume a slightly stronger condition than normality, viz. that if an agent i can prove that the true state lies in either set $X \subseteq \Theta$ or $Y \subseteq \Theta$, then he can also prove that it lies in $X \cap Y$.[21] The literature on strategic communication with hard information following Milgrom (1981) and Grossman (1981) typically assumes much stronger conditions than normality.

Example 5 (Example 2 continued)**.** *Consider Example 2 where $\Theta = \mathbb{R}_+^n$, $E_i = \Theta$ and $E_i^\theta = [0, \theta_i]$ for all i, θ. This evidence structure satisfies normality, as seen by setting $\bar{e}_{i,\theta} = \theta_i$ for all i, θ. Observe that if $\bar{e}_{i,\theta} = \theta_i \in E_i^{\theta'} = [0, \theta_i']$ then it must be that $\theta_i \leq \theta_i'$ and so $E_i^\theta = [0, \theta_i] \subseteq [0, \theta_i'] = E_i^{\theta'}$, as required.*

Example 6 (Example 3 continued)**.** *It is also straightforward to check that Example 3 satisfies normality: for all i, set $\bar{e}_{i,w} = \bar{e}_{i,x} = \beta$ and $\bar{e}_{i,y} = \bar{e}_{i,z} = \alpha$.*

Under normality, the definition of evidence-monotonicity can be simplified by replacing condition (**) with

$$E^\theta = E^{\theta'}. \tag{1}$$

The reason is that under normality, we can take $e^*_{i,\theta,a}$ in Definition 2 to equal $\bar{e}_{i,\theta}$ of Definition 5, because normality implies that for any $e_i \in E^\theta$, if $\bar{e}_{i,\theta} \in E_i^{\theta'}$, then $e_i \in E^{\theta'}$. This observation leads to an alternative characterization of evidence-monotonicity under normality. Say that a triplet (θ, a, θ') where $a \in f(\theta)$ *violates Maskin-monotonicity* if $a \notin f(\theta')$ and (*) is satisfied. For any θ

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

[21] It is clear that Ben-Porath and Lipman's (2008) assumption implies normality; to see that is strictly stronger, consider the following example: $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, $E_i^{\theta_1} = \{x, y, z\}$, $E_i^{\theta_2} = \{\tilde{x}, y, z\}$, $E_i^{\theta_3} = \{y\}$, $E_i^{\theta_4} = \{z\}$. This structure is normal (set $\bar{e}_{i,\theta_1} = x$ and $\bar{e}_{i,\theta_2} = \tilde{x}$) but in state θ_1, while any agent can prove that the state lies in $\{\theta_1, \theta_2, \theta_3\}$ by sending y or that the state lies in $\{\theta_1, \theta_2, \theta_4\}$ by sending z, he cannot prove that the state lies in $\{\theta_1, \theta_2\}$.

and $a \in f(\theta)$, we define

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

and for any θ,

$$T^f(\theta) := \bigcup_{a \in f(\theta)} T^f(\theta, a).$$

In other words, given that one wishes to achieve the outcomes $f(\theta)$ in state θ, $T^f(\theta)$ are precisely the states that cause a problem for implementation of f in the absence of evidence. In particular, a SCR f is implementable without evidence only if $\bigcup_{\theta \in \Theta} T^f(\theta) = \emptyset$.

Proposition 1. *Assume that the evidence structure is normal. Then a SCR f is evidence-monotonic if and only if*

$$\forall \theta : \theta' \in T^f(\theta) \implies E^\theta \neq E^{\theta'}. \tag{2}$$

Proof. Assume normality. For the "only if" part of the result, suppose that (2) fails. Then there is some θ, θ', and a such that $E^\theta = E^{\theta'}$, $a \in f(\theta)$, $a \notin f\left(\theta'\right)$, and (*). From Definition 2 and the simplification of (**) to $E^\theta = E^{\theta'}$ under normality, f is not evidence-monotonic. The "if" part is similarly straightforward and omitted. □

To see an application of this result, we return to Example 3.

Example 7 (Example 3 continued.)**.** *Consider the SCR f from Example 3, where the evidence structure is normal. One can check that $T^f(w) = T^f(z) = y$, $T^f(x) = \emptyset$, and $T^f(y) = \{w, z\}$. Since $E^z = E^y$, Proposition 1 confirms that f is not evidence-monotonic (as argued directly before). On the other, consider the SCR f^* from the example. We have $T^{f^*}(w) = \{y, z\}$, $T^{f^*}(x) = \emptyset$, $T^{f^*}(y) = T^{f^*}(z) = w$. Since $E^w \neq E^y = E^z$, (2) is satisfied and hence f^* is evidence-monotonic (as argued directly before).*

Proposition 1 highlights what evidence structure is needed to implement non-Maskin-monotonic SCRs. It can be combined with our earlier results to yield useful corollaries.

Corollary 1. *Assume $n \geq 3$ and that the evidence structure is normal. A SCR f that satisfies both no veto power and non-satiation is implementable if*

$$\forall \theta, \theta' : E^\theta = E^{\theta'} \implies f(\theta) = f(\theta'). \tag{3}$$

Proof. Plainly, (3) implies (2). The result follows from Proposition 1 and Theorem 2. □

Remark 8. Ben-Porath and Lipman (2008) call condition (3) "measurability." In their Proposition 1, they show that under assumptions of normality and state-independent preferences for all agents (i.e., $\forall i \in I$, $\forall \theta, \theta' \in \Theta$, $\forall a, b \in A$, $u_i(a, \theta) \geq u_i(b, \theta) \implies u_i\left(a, \theta'\right) \geq u_i\left(b, \theta'\right)$), a SCR is implementable only if it satisfies (3). This is a special case of our Theorem 1 because under state-independent preferences, condition (*) is satisfied for all a, θ, and θ'; hence because (**) simplifies to (1) under normality, Theorem 1 reduces to requiring (3) for an implementable SCR. Conversely, Corollary 1 says that when $n \geq 3$, no veto power, non-satiation and (3) are sufficient for implementation under normality. Ben-Porath and Lipman (2008) have also independently proved a similar result. Note that even under normality, (3) is not necessary for implementation when preferences are not state-independent, for instance the SCR f^* in Example 3. The reason is when hard evidence is the same between two states, the planner can exploit preference reversals to implement different outcomes.

Motivated by the previous result, consider the following condition on the evidence structure:

$$\forall \theta : \theta' \neq \theta \implies E^{\theta'} \neq E^{\theta}. \tag{UPD}$$

Condition (UPD), short for *Universal Pairwise Distinguishability*, requires that any pair of distinct states be distinguishable via evidence, i.e. there must be at least one player whose available evidence differs in the two states. Obviously, if two states θ and θ' cannot be distinguished in this sense, implementation requires that no triplet (θ, a, θ') violate Maskin-monotonicity. On the other hand, since (UPD) implies that (3) is trivially satisfied, we also have:

Corollary 2. *Assume $n \geq 3$ and that the evidence structure is normal.* Any *SCR that satisfies both no veto power and non-satiation can be implemented if* (UPD) *holds.*

The above sufficient condition is tight: if (UPD) is violated, we can specify a profile of utility functions and a SCR (satisfying no veto power and non-satiation) such that this SCR is not evidence-monotonic and therefore not implementable.

2.4.2. *Non-Normal Evidence Structures.* Despite being a prevalent assumption, normality is a fairly demanding requirement.[22] One interpretation is that it assumes that there is no constraint on time, effort, space, etc., in providing evidence. The next example introduces some of the issues that arise with non-normal evidence structures.

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 $\Theta = \{\phi, a, b, ab\}$, and for all i, $E_i^{\phi} = \{\phi\}$, $E_i^a = \{\phi, a\}$, $E_i^b = \{\phi, b\}$, and $E_i^{ab} = \{\phi, a, b\}$, where ϕ represents "neither proposition is true" or "no proof provided." Although this evidence structure satisfies* (UPD)*, it is not normal, because for any i and $x \in E_i^{ab}$, there exists $\theta' \in \{a, b\}$ such that $x \in E_i^{\theta}$ but $E_i^{ab} \nsubseteq E_i^{\theta}$. Note that if each expert could prove both a and b when both are true, then we would augment ab to E_i^{ab}, and normality would be satisfied.*

Suppose now that the preferences of the experts over outcomes are state-independent, so that (*) *is always satisfied. Since any choice of $\{e^*_{i,ab}\}_{i=1}^n$ will create some $\theta' \in \{a, b\}$ such that* (**) *is satisfied with $\theta = ab$, it follows from Theorem 1 that not every SCR is implementable. In particular, if the SCR is a function, then implementability requires $f(ab) \in \{f(a), f(b)\}$. On the other hand, by choosing $e^*_{i,\phi} = \phi, e^*_{i,a} = a$, and $e^*_{i,b} = b$, Theorem 2 implies that $f(ab) \in \{f(a), f(b)\}$ is also sufficient for the social choice function to be implementable under no veto power and non-satiation.*

[22] Bull and Watson (2007), Glazer and Rubinstein (2001, 2004, 2006), Lipman and Seppi (1995), and Sher (2008) study problems where normality does not hold.

This motivates the question of what evidence structures permit non-Maskin-monotonic SCRs to be implemented even when the evidence structure is not normal. A sharp answer can be provided by using a general notion of distinguishability.

Definition 6. *For any θ and $\Omega \subseteq \Theta$, θ and Ω are distinguishable if for any $\Omega' \subseteq \Omega : E^{\theta} \neq \bigcup_{\theta' \in \Omega'} E^{\theta'}$.*

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 $E_i^{\theta} \nsubseteq \bigcup_{\theta' \in \Omega'} E_i^{\theta'}$) or some player can disprove θ when some state in Ω' is the true state (which requires $E_i^{\theta} \nsupseteq \bigcup_{\theta' \in \Omega'} E_i^{\theta'}$). 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, $\theta \notin \Omega$). The following result establishes the converse for normal evidence structures.

Proposition 2. *Assume the evidence structure is is normal. For any $\theta \in \Theta$ and $\Omega \subseteq \Theta$, if θ is distinguishable from each $\theta' \in \Omega$ then θ is distinguishable from Ω.*

Proof. Fix θ and Ω and assume that $\forall \theta' \in \Omega : E^{\theta} \neq E^{\theta'}$. Suppose, per contra, that for some $\Omega' \subseteq \Omega : E^{\theta} = \bigcup_{\theta' \in \Omega'} E^{\theta'}$. Then for all $\theta' \in \Omega'$, $\bar{e}_{\theta'} \in E^{\theta}$ (where $\bar{e}$ is from the definition of normality), and moreover for some $\tilde{\theta} \in \Omega'$, $\bar{e}_{\theta} \in E^{\tilde{\theta}}$. By normality, $E^{\tilde{\theta}} = E^{\theta}$, a contradiction. □

Example 8 shows why normality is key to the above result: in the example, state ab is pairwise distinguishable from every other state but not distinguishable from the event $\{\phi, a, b\}$.

We can now state a general characterization of evidence-monotonicity in terms of distinguishability.

Proposition 3. *A SCR f is evidence-monotonic if and only if for all θ and $a \in f(\theta)$, θ and $T^f(\theta, a)$ are distinguishable.*

Proof. See Appendix. □

Remark 9. Proposition 1 can be seen as a corollary of Proposition 3, because when the evidence structure is normal, distinguishability of any θ and Ω reduces to distinguishability of θ from each $\theta' \in \Omega$ (Proposition 2).

An immediate implication of Proposition 3 is that the following condition on the evidence structure guarantees that every SCR is evidence-monotonic, no matter what the agents' preferences are:

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

Condition (UD), short for *Universal Distinguishability*, requires that each state must be distinguishable from any event that does not contain it. In general, this will be a stronger requirement than (UPD), but Proposition 2 implies that they are equivalent under normality.

Corollary 3. *If* $n \geq 3$. Any *SCR that satisfies both No veto power and non-satiation can be implemented if* (UD) *holds.*

Proof. Let f be an arbitrary SCR. Pick any θ and $a \in f(\theta)$. Since $\theta \notin T^f(\theta, a)$, (UD) implies that θ is distinguishable from $T^f(\theta, a)$. By Proposition 3, f is evidence-monotonic. Theorem 2 yields the desired conclusion. □

The Corollary is tight in the sense that if (UD) is violated, we can specify a profile of utility functions and a SCR (satisfying no veto power and non-satiation) such that the SCR is not evidence-monotonic and hence not implementable. We illustrate the Corollary with the following example.

Example 9. $\Theta = \{\theta_1, \theta_2, \theta_3\}$; $n = 3$; $E_1^{\theta_1} = \{x, y\}$, $E_1^{\theta_2} = \{x\}$, $E_1^{\theta_3} = \{y\}$, $E_2^{\theta_1} = \{x, y\}$, $E_2^{\theta_2} = \{y\}$, $E_2^{\theta_3} = \{x\}$, *and* $E_3^{\theta} = \{z\}$ *for all* θ. *It is easy to see that normality does not hold.* (UD) *holds because* $E^{\theta_1} = \{(x, x, z), (x, y, z), (y, x, z), (y, y, z)\}, E^{\theta_2} = \{(x, y, z)\}$, *and* $E^{\theta_3} = \{(y, x, z)\}$. *Hence, by Corollary 3, any SCR satisfying no veto power and non-satiation is implementable.*

2.5. **Ranking Evidence Structures.** Since the structure of evidence crucially affects the possibility of implementation, it is natural to ask whether some evidence structures are "more informative" than others from a planner's perspective because they allow for a larger set of implementable SCRs. Not surprisingly, the notion of distinguishability is key.

Definition 7. *Evidence structure* $\tilde{\mathcal{E}}$ *is more informative than* $\mathcal{E}$, *denoted* $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$, *if any* $\theta \in \Theta$ *and* $\Omega \subseteq \Theta$ *that are distinguishable under* $\mathcal{E}$ *are also distinguishable under* $\tilde{\mathcal{E}}$.

Thus, $\blacktriangleright$ is partial order on the set of possible evidence structures. At one end, if $\mathcal{E}$ satisfies (UD), then for any $\tilde{\mathcal{E}}$, $\mathcal{E} \blacktriangleright \tilde{\mathcal{E}}$. At the other end, if $\mathcal{E}$ represents no evidence ($\forall i, \theta : E_i^{\theta} = E_i$), then for any $\tilde{\mathcal{E}}$, $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$.

Remark 10. If $\tilde{\mathcal{E}}$ is normal, then $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$ if and only if

$$\{(\theta, \theta') : E^{\theta} \neq E^{\theta'}\} \subseteq \{(\theta, \theta') : \tilde{E}^{\theta} \neq \tilde{E}^{\theta'}\}.$$

Proof. See Appendix. □

The following result shows why $\blacktriangleright$ appropriately captures when one evidence structure is "better" than another.

Proposition 4. *Assume that* $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$. *If a SCR is evidence-monotonic under* $\mathcal{E}$ *it is also evidence-monotonic under* $\tilde{\mathcal{E}}$.

Proof. Pick any SCR f that is evidence-monotonic under $\mathcal{E}$, and any θ and $a \in f(\theta)$. By the "only if" part of Proposition 3, θ is distinguishable from $T^f(\theta, a)$. Since $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$, θ and $T^f(\theta, a)$ are also distinguishable under $\tilde{\mathcal{E}}$. Since θ and $a \in f(\theta)$ are arbitrary, the "if" part of Proposition 3 implies that f is evidence-monotonic under $\tilde{\mathcal{E}}$. □

Remark 11. The above result is tight in the sense that if $\tilde{\mathcal{E}}$ is not more informative than $\mathcal{E}$, then there exist preferences for agents and a SCR such that the SCR is evidence-monotonic under $\mathcal{E}$ but not under $\tilde{\mathcal{E}}$. See the Appendix for an explicit construction.

Proposition 4 and Theorem 2 imply the following.

Corollary 4. *Assume that $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$ and $n \geq 3$. Let f be a SCR satisfying no veto power and non-satiation. If f is implementable under $\mathcal{E}$ then f is also implementable under $\tilde{\mathcal{E}}$.*

2.6. **Ranking Social Choice Rules.** Just as the notion of distinguishability allowed us to create a partial order on evidence structures, it also allows a partial order on SCRs ranking them in terms of how easy they are to implement. We begin with a definition capturing when a SCR is "more Maskin-monotonic".

Definition 8. *For SCRs f and h, f is more Maskin-monotonic than h, denoted $f \trianglerighteq h$, if*

$$\forall \theta, a \in f(\theta) : \exists a' \in h(\theta) \text{ s.t. } T^f(\theta, a) \subseteq T^h(\theta, a'). \tag{4}$$

Note that this definition is independent of any evidence structure. If f is Maskin-monotonic, then $f \trianglerighteq h$ for any h, and moreover, if h is not Maskin-monotonic, then [not $h \trianglerighteq f$]. The next result shows why the relation of more Maskin-monotonic is useful.

Proposition 5. *If $f \trianglerighteq h$, then if h is evidence-monotonic under evidence structure $\mathcal{E}$, f is also evidence-monotonic under $\mathcal{E}$.*

Proof. Suppose h is evidence-monotonic under $\mathcal{E}$. Pick any θ and $a \in f(\theta)$; by the "if" part Proposition 3, it suffices to show that θ is distinguishable from $T^f(\theta, a)$. By (4), there exists $a' \in h(\theta)$ such that $T^f(\theta, a) \subseteq T^h(\theta, a')$. By the "only if" part of Proposition 3, θ is distinguishable from $T^h(\theta, a')$, and hence distinguishable from every subset of $T^h(\theta, a')$, including $T^f(\theta, a)$. □

Proposition 5 and Theorem 2 imply the following corollary, which justifies why $\trianglerighteq$ ranks SCRs in terms of their implementability.

Corollary 5. *Assume $n \geq 3$ and f and h are SCRs satisfying NVP and non-satiation such that $f \trianglerighteq h$. If h is implementable under evidence structure $\mathcal{E}$, then f is also implementable under $\mathcal{E}$.*

2.7. **Two Agents.** As noted earlier, non-satiation and no veto power are each overly demanding when there are only two agents; indeed, the two cannot be jointly satisfied (unless there are an infinite number of alternatives and neither player has a top-ranked alternative in any state).[23] Yet the two agent case is quite important in problems of contracting, bargaining, and dispute resolutions. We thus give it a separate treatment.

In the standard setting without evidence, implementation with only two players is known to be significantly more demanding than with many players. Maskin (1999) shows that with two players,

[23]With an infinite number of alternatives, it is not hard to construct an example of a two-player setting where a SCR is not implementable even though non-satiation, no veto power and evidence-monotonicity are satisfied. Thus, Theorem 2 does not extend to $n = 2$.

a Pareto-efficient SCR defined over an unrestricted domain of preferences is implementable if and only if it is dictatorial.[24] The major complication with two-player implementation is that the planner faces a severe challenge in identifying which player has deviated when a non-equilibrium message profile is observed. Dutta and Sen (1991) and Moore and Repullo (1990) provide a full characterization of implementable SCRs for the two-player case without evidence by introducing a condition known as *condition-β*. In particular, Moore and Repullo (1990, Corollary 4) provide conditions under which Maskin-monotonicity is necessary and sufficient for implementation in settings without evidence. In this section, we derive an analogous result in our setting with evidence.[25]

Let $f(\Theta) := \bigcup_\theta f(\theta)$ denote the range of f. The following condition of *restricted veto power* is due to Moore and Repullo (1990):

Definition 9 (Restricted Veto Power). *A SCR f satisfies restricted veto power (RVP) if for all i, θ, a and $b \in f(\Theta)$,*

$$\left[\left(a \in \bigcap_{j \neq i} \arg\max_{c \in A} u_j(c, \theta)\right) \wedge (u_i(a, \theta) \geq u_i(b, \theta))\right] \implies a \in f(\theta).$$

In words, RVP says that if an outcome a is top-ranked at state θ by all players $j \neq i$, and if there is an outcome b in the range of f such that i weakly prefers a to b at state θ, then a must belong to $f(\theta)$. Thus, agent i does not have the power to veto outcome a unless it is strictly worse for him than every outcome in the range of f. Plainly, RVP is a weakening of NVP, trivially coinciding with NVP when $f(\Theta) = A$. Intuitively, the smaller the range of f, the less demanding is RVP.

We say that there is a *bad outcome* (relative to a SCR f) if there is some outcome $z \notin f(\Theta)$ such that for all $i, \theta : u_i(z, \theta) < u_i(a, \theta)$ for all $a \in f(\Theta)$. In other words, in any state, outcome z is strictly worse for all players than any outcome in the the range of the SCR. We are now in position to generalize a result of Moore and Repullo (1990, Corollary 4) to settings with evidence.

Theorem 3. *Assume $n = 2$ and the existence of a bad outcome. A SCR f satisfying RVP is implementable if and only if it is evidence-monotonic.*

Proof. See Appendix. □

Thus, when $n = 2$, there is a bad outcome, and RVP holds, evidence-monotonicity precisely characterizes what is implementable. Note that non-satiation plays no role in the above Proposition.[26]

[24]In our framework, unrestricted domain of preferences means that for any profile of complete and transitive binary relations over outcomes, $(\precsim_1, ..., \precsim_n)$, there exists a state of the world, θ, in which for each player i, $x \precsim_i y \Leftrightarrow u_i(x, \theta) \leq u_i(y, \theta)$. A SCR f is dictatorial if there is some i such that for all a and θ, $a \in f(\theta) \Leftrightarrow a \in \arg\max_b u_i(b, \theta)$.

[25]In a previous version of this paper, we extended *condition-β* to *condition evidence-β* and showed that this property is necessary and sufficient for implementation of a SCR satisfying the condition of restricted veto-power discussed below. This characterization is available upon request.

[26]Indeed, for $n = 2$, an implication of RVP is that for $i \in \{1, 2\}$:

$$\left[\left(a \in \arg\max_{c \in A} u_i(c, \theta)\right) \wedge (a \in f(\Theta))\right] \implies a \in f(\theta).$$

This condition is weaker than non-satiation (since non-satiation requires that the antecedent of the above condition is not satisfied for any i) but would actually be sufficient to prove Theorem 2. Thus with only two players, RVP

We also remark that while restrictive, both RVP and existence of a bad outcome are naturally satisfied in some economic environments where there is sufficient conflict between agents.[27] Example 10 below provides a non-economic environment satisfying both assumptions.

Combining Proposition 3 and Theorem 3 yields a two-agent counterpart of Corollary 3.

Corollary 6. *Assume $n = 2$ and the existence of a bad outcome.* Any *SCR that satisfies RVP can be implemented if* (UD) *holds.*

The following example illustrates an application of this result.

Example 10. *Suppose the outcome is how much of some public good to produce, so that $A = \mathbb{R}_+$. There are two agents, each of whom has single-peaked preferences with a bliss point. Specifically, if an agent has bliss point x, his payoff from amount a of the public good is $-(x-a)^2$. Agent 1's possible bliss points are $[2, 3]$ whereas agent 2's possible bliss points are $[5, 6]$. A state of the world is a pair of bliss points, i.e. $\Theta = [2, 3] \times [5, 6]$. The SCR f is given by the midpoint of the two agents' bliss points.*

Observe that $f(\Theta) = [3.5, 4.5]$. Consequently, 0 plays the role of a bad outcome. RVP is also satisfied because in any state, the bliss point for one agent is strictly worse for the other agent than any outcome in $f(\Theta)$. Note that neither NVP nor non-satiation are satisfied, however.

Now let the evidence structure be such that in any state $\theta = (\theta_1, \theta_2)$, $E_i^\theta = \{[x, y] : x \leq \theta_i \leq y\}$; in other words, each agent submit evidence that his bliss point lies in any interval containing his bliss point. (UD) *is satisfied because $\theta \in E^{\theta'}$ if and only if $\theta' = \theta$.*

By Corollary 6, f is implementable, even though f is not Maskin-monotonic.

Remark 12. We should note that both the existence of a bad outcome and RVP are essential to Theorem 3 and Corollary 6. Examples 16 and 17 in the Appendix show that without either of the two conditions, the results need not hold.

2.8. **One Agent.** Finally, we consider the case where there is only one agent. This case is interesting from the point of view of the "principal-agent" literature. To ease notation, we drop the agent index in the following theorem, which provides a complete characterization of implementable SCRs when there is only one agent.

Theorem 4. *Assume $n = 1$. A SCR f is implementable if and only if*

$$[a, b \in f(\theta)] \Longrightarrow [u(a, \theta) = u(b, \theta)], \tag{5}$$

and for all θ and $a \in f(\theta)$, there exists $e^*_{\theta,a} \in E^\theta$ such that for all θ', $b \in f(\theta')$

$$\left[\left(u(a, \theta') \geq u(b, \theta')\right) \wedge \left(e^*_{\theta,a} \in E^{\theta'}\right)\right] \Longrightarrow \left[a \in f(\theta')\right], \tag{6}$$

(and *a fortiori* NVP) implies the portion of non-satiation that is essential for our sufficiency argument, but this is not the case when there are three or more players.

[27]For example, if the outcome is how much of a divisible private good each agent gets and the SCR provides a strictly positive amount to each agent, then under usual preferences, $(0, 0)$ is a bad outcome and RVP holds if there is a feasibility constraint on the total amount of the private good that can be provided.

and for any $T \subseteq \Theta$ *such that* $\bigcap_{\theta'' \in T} E^{\theta''} \neq \emptyset$, *there exists* $c \in A$ *such that for any* $\theta \in T$,

$$[u(c,\theta) < u(b,\theta)\ \forall b \in f(\theta)] \vee [c \in f(\theta)]. \tag{7}$$

Proof. See Appendix. ☐

Though seeming complicated, the conditions of the Theorem are intuitive. Condition (5) says that at any state, the agent must be indifferent between all the socially desired outcomes at that state. Condition (6) says that if outcome $b \in f(\theta')$ is attainable at θ (hence the evidence $e^*_{\theta',b} \in E^\theta$), but $b \notin f(\theta)$, then the agent must strictly prefer the social desired outcomes at θ to b. It is fairly straightforward why both these conditions are necessary for implementation. The third condition, (7), says that if some evidence is common to a set of states, then there is some outcome such that at any of these states, either the outcome is socially desired, or it is strictly worse for the agent than the social desired ones. Loosely, this condition is necessary for implementation because otherwise, any way in which the planner responds to this evidence will lead to an undesired outcome (Example 11 below illustrates).

Remark 13. If there is a bad outcome, call it z, then (7) is satisfied for all θ by setting $c = z$. Regarding (6), we observe that for an arbitrary specification of preferences and SCR, (UD) is not generally sufficient. Instead, what is needed is the much stronger requirement of *complete provability* (Lipman and Seppi, 1995): $\forall\theta$, $\exists \hat{e}^\theta \in E^\theta$ s.t. $\hat{e}^\theta \notin \bigcup_{\theta' \neq \theta} E^{\theta'}$.[28] We conclude that under the demanding conditions of there existing a bad outcome and complete provability, any single-valued SCR (which trivially satisfies (5)) is implementable with only one agent.

2.9. **Discussion.** In this section, we address two notable restrictions in the class of mechanisms we have considered to this point.

2.9.1. *Forbidding Evidence.* We have assumed that a mechanism cannot constrain the evidence that agents can submit. In some settings, this assumption may be a priori restrictive. For example, certain kinds of evidence, such as hearsay, could be prohibited in legal proceedings. Accordingly, a more general definition of a mechanism would be to specify $(M, \hat{E}, g)$, where $\hat{E} = \hat{E}_i \times \cdots \times \hat{E}_n$ with for each i, $\hat{E}_i \subseteq E_i$, and $g : M \times \hat{E} \to A$. Here, agent i is only allowed to submit evidence in the set $\hat{E}_i$, i.e. the evidence player i can send when the state is θ is $\hat{E}_i^\theta = \hat{E}_i \cap E_i^\theta$. If for some i, $\hat{E}_i \neq E_i$, then the mechanism is said to be *forbidding evidence.*

The following example shows that forbidding evidence can be useful.

Example 11. *Suppose* $\Theta = \{\theta_1, \theta_2\}$, $A = \{b, c\}$, *and for all i:* $E_i^{\theta_1} = \{x, y\}$, $E_i^{\theta_2} = \{x, z\}$, *and preferences are represented by* $u_i(b,\theta_1) > u(c,\theta_1)$ *and* $u_i(c,\theta_2) > u(b,\theta_2)$. *Consider the SCR* $f(\theta_1) = c$ *and* $f(\theta_2) = b$. *If a mechanism* (M, g) *does not forbid evidence, it cannot implement* f, *by the following logic: we must have* $g(m, (x, \ldots, x)) = c$ *for all* $m \in M$, *otherwise* $b \notin f(\theta_1)$ *would be a Nash equilibrium outcome in state* θ_1*; but then* $(m, (x, \ldots, x))$ *is a Nash equilibrium at state*

[28] Complete provability is necessary to guarantee (6) in the sense that if complete provability fails, we can write down preferences and a SCR such that (6) is violated.

θ_2 which leads to outcome c, hence f is not implemented. Intuitively, the problem is that since agents can produce evidence x in both states, it is impossible to achieve the outcome they like less in both states. But if a mechanism can forbid evidence, it is straightforward to implement f: let agents only submit evidence in $\hat{E}_i = \{y, z\}$, and let the outcome be c if they each submit y, and b otherwise.

Nevertheless, Theorems 1 and 2 remain valid even when mechanisms that forbid some evidence are considered. Thus, even if evidence can be forbid, evidence-monotonicity remains necessary for implementation and is also sufficient when $n \geq 3$ and no veto power and non-satiation hold. (Indeed, Example 11 satisfies evidence-monotonicity but crucially violates NVP.[29])

2.9.2. *Dynamic Mechanisms.* Our treatment of mechanisms thus far implicitly assumes that the induced game by any mechanism is *static*, i.e. entails one round of simultaneous moves. In the standard environment without evidence, sequential or *dynamic mechanisms*—ones that induce extensive-form games richer than one round of simultaneous moves—cannot improve on static mechanisms so long as the solution concept remains Nash equilibrium. However, this is not the case with hard evidence, as was pointed out in a partial-implementation framework by Bull and Watson (2007). Consider the following example borrowed from their paper:

Example 12. *There are $n = 2$ players, $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and $A = \{a, b\}$. Preferences are state-independent: for each θ, $u_1(b, \theta) > u_1(a, \theta)$ whereas $u_2(a, \theta) > u_2(b, \theta)$. Player 1 has only trivial evidence available in any state, i.e. $E_1^{\theta_1} = E_1^{\theta_2} = E_1^{\theta_3} = E_1$. Player 2's evidence sets are given by $E_2^{\theta_1} = \{x\}$, $E_2^{\theta_2} = \{x, y\}$ and $E_2^{\theta_3} = \{y\}$. This evidence structure is not normal. The SCR is given by $f(\theta_1) = f(\theta_3) = \{b\}$ and $f(\theta_2) = \{a\}$. This SCR is not evidence-monotonic, because $T^f(\theta_2) = \{\theta_1, \theta_3\}$ and $E^{\theta_2} = E_1 \times \{x, y\} = E^{\theta_1} \cup E^{\theta_3}$, violating the condition for evidence-monotonicity in Proposition 3. Thus, by Theorem 1, f is not implementable.*[30]

The point we now turn to is that whether dynamic mechanisms can be useful or not depends critically on the interpretation one has of evidence "not being available" at some state. Thus far, we have assumed that any $e_i \notin E_i^\theta$ is physically unavailable and cannot be produced by agent i at state θ. An alternative interpretation is that agents can always fabricate or produce any evidence in any state, but some evidence is more costly than others — hard evidence being an extreme case where any $e_i \in E_i^\theta$ is costless, whereas any $e_i \notin E_i^\theta$ is infinitely costly. (We consider less extreme cost structures in the subsequent section.) Let us call these two interpretations the *feasibility* and *cost* interpretations respectively. Insofar as static mechanisms are concerned, the distinction is irrelevant, since in equilibrium, a player would never produce evidence that is either unavailable or infinitely costly, and there are no out-of-equilibrium considerations.

With dynamic mechanisms, the issue is more subtle. Broadly, the reason that the analysis with dynamic mechanisms can depend on the feasibility versus cost interpretation is as follows: under the cost interpretation, at state θ any player i *will* be able to send any $e_i \notin E_i^\theta$ off the equilibrium

[29]As constructed, the example also violates non-satiation, but that is easily accommodated by adding an outcome that is uniquely top-ranked in every state by every agent.

[30]This has already been shown by Bull and Watson (2007) using a different argument.

path, and this will be a standard incredible threat; on the other hand, such a behavior is precluded under the feasibility interpretation. Let us illustrate by returning to the example.

Example 13 (Example 12 continued)**.** *Consider the following dynamic mechanism:*

1st Stage: *player 1 can announce any state $\theta \in \Theta$. Denote this (cheap-talk) announcement m_1.*

2nd Stage: *after observing m_1, player 2 has to submit evidence, e_2.*

Outcomes: *$g(m_1, e_2) = a$ if $m_1 = \theta_2$; else*

$$g(m_1, e_2) = \begin{cases} b \text{ if } e_2 \in E_2^{m_1} \\ a \text{ otherwise.} \end{cases}$$

Thus the mechanism chooses outcome a if player 1 claims (via cheap talk) that the true state is θ_2; if he claims anything else, the mechanism chooses a if player 2 submits evidence contradicting player 1's claim, otherwise the mechanism chooses b.

Start with the feasibility interpretation, i.e. assume that when the true state is θ, any $e_i \notin E_i^\theta$ is not available to player i (even at an infinite cost). Then, if the true state is θ_1, player 2 can only send evidence x, hence player 1 can ensure that his preferred outcome b is chosen by announcing $m_1 = \theta_2$, and this is the unique Nash equilibrium outcome. If the true state is θ_3, a similar reasoning shows that the only NE outcome is again b. Now assume the true state is θ_2, in which case player 2 has evidence x and y available. By playing the strategy of submitting evidence x if $m_1 \neq \theta_1$ and submitting evidence y if $m_1 = \theta_1$, player 2 guarantees that the outcome is a no matter what player 1 announces. Since a is player 2's preferred outcome, every Nash equilibrium must yield outcome a (and a subgame perfect NE exists). Therefore, under the feasibility interpretation, the SCR is implementable with a dynamic mechanism when it is not with any static mechanism.

Now consider the cost interpretation, where in any state θ, player 2 can send any $e_i \in E_2 = \{x, y\}$, but incurs an infinite cost if $e_i \notin E_2^\theta$ and no cost otherwise. We will argue that the SCR is no longer implementable by the given dynamic mechanism. If the true state is θ_1, then the following is a Nash equilibrium: player 1 announces θ_2; player 2 submits evidence y if $m_1 = \theta_1$ and submits evidence x otherwise. This NE yields outcome $a \neq f(\theta_1)$. Note that this NE is not subgame perfect.

We have thus proved:

Proposition 6. *Under the feasibility interpretation, a non-normal evidence structure can permit implementation with a dynamic mechanism even if evidence-monotonicity fails.*

However, it turns out that the necessity of evidence-monotonicity for implementation is preserved under either: (i) the cost interpretation; or (ii) the feasibility interpretation but with a normal evidence structure. We formally state and prove these two results in Appendix B.[31]

[31]Importantly, these results are proved assuming that extensive forms are restricted to those in which each player can only send evidence once in the game, as in Bull and Watson (2007). This avoids artificially "transforming" a

3. Costly Evidence Fabrication

In this section, we consider a more general setting where agents can fabricate any evidence in any state of the world, but potentially bear differential costs of fabrication across states. Specifically, in each state of the world, θ, agent i can create evidence $e_i \in E_i$ but thereby incurs a disutility of $c_i(e_i, \theta) \in \mathbb{R}_+ \cup \{+\infty\}$. We refer to $(c_i(\cdot,\cdot))_{i \in I}$ as the *cost structure* and we denote $E_i^\theta = \{e_i \in E_i : c_i(e_i, \theta) = 0\}$. Without loss of generality, assume that for all θ, i, $E_i^\theta \neq \emptyset$.[32] This defines a general costly signaling model. Note in particular that it reduces to the hard evidence model (with cost interpretation) if the range of each $c_i(\cdot,\cdot)$ is $\{0, +\infty\}$.

We now explicitly allow the planner to make (monetary) transfers among agents. For the remainder of this section, a mechanism refers to a *mechanism with transfers*, which is a tuple (M, g, τ), where $M = M_1 \times \cdots \times M_n$ is the (cheap-talk) message space, $g : M \times E \to A$ is the outcome mapping, and $\tau : M \times E \to \mathbb{R}^n$ is a *transfer mapping* that is a novel instrument compared to the setting in Section 2. A mechanism (M, g, τ) induces a strategic game in each state of the the world, where a pure strategy for player i in state θ is $s_i \in M_i \times E_i$. The payoff for player i from a profile (m, e) in state θ is

$$u_i(g(m, e), \theta) - c_i(e_i, \theta) + \tau_i(m, e).$$

Let $NE(M, g, \tau, \theta)$ be the set of pure strategy Nash equilibria of this game at state θ, and

$$O(M, g, \tau, \theta) := \{a \in A : \exists (m, e) \in NE(M, g, \tau, \theta) \text{ s.t. } g(m, e) = a\}$$

be the set of equilibrium allocations at state θ. Similarly, let

$$T(M, g, \tau, \theta) := \{t \in \mathbb{R}^n : \exists (m, e) \in NE(M, g, \tau, \theta) \text{ s.t. } \tau(m, e) = t\}$$

be the set of equilibrium transfer vectors at state θ.

We are interested in (full-)implementation in Nash equilibria where no transfer occurs at equilibrium. We will also assume that in addition to no transfers, the planner's objective is to avoid any signaling costs, i.e. in equilibrium only costless messages can be used.

Definition 10 (Implementation with transfers)**.** *A mechanism (M, g, τ) implements the SCR f if $\forall \theta$, $f(\theta) = O(M, g, \tau, \theta)$ and $T(M, g, \tau, \theta) = \{(0, \ldots, 0)\}$. In addition, we require that $(m, e) \in NE(M, g, \tau, \theta) \implies e \in E^\theta$. A SCR f is implementable if there exists a mechanism that implements it.*

We now define the relevant notion of monotonicity in this setting.

Definition 11 (Cost-monotonicity)**.** *A SCR f is cost-monotonic provided that for all θ and all $a \in f(\theta)$, there exists $e^*_{\theta,a} \in E^\theta$ such that for all θ': if*

$$\forall i \in I, b \in A, \alpha \in \mathbb{R} : [u_i(a, \theta) \geq u_i(b, \theta) - \alpha] \implies [u_i(a, \theta') \geq u_i(b, \theta') - \alpha] \quad \text{(CM*)}$$

non-normal evidence structure into a normal one, which would happen if an agent could send evidence an unlimited number of times.

[32]This is without loss of generality because we are going to allow the planner to add costless (cheap-talk) messages.

and

$$\forall i \in I : \left(e^*_{i,\theta,a} \in E_i^{\theta'}\right) \wedge \left(c_i(\cdot,\theta') \geq c_i(\cdot,\theta)\right),^{33} \tag{CM**}$$

then $a \in f(\theta')$*.*

Remark 14. Cost-monotonicity is more general than evidence-monotonicity. To see this, note first that (CM*) implies (*) by considering $\alpha = 0$. Second, (CM**) implies (**) because $c_i(\cdot,\theta') \geq c_i(\cdot,\theta)$ implies $E_i^{\theta'} \subseteq E_i^{\theta}$. Therefore, any evidence-monotonic SCR is cost-monotonic. Moreover, we note that if $c_i(\cdot,\cdot) \in \{0,+\infty\}$ for all i—the case of hard evidence with cost interpretation—(CM**) reduces to (**), so that if in addition there were no transfers, cost-monotonicity would be equivalent to evidence-monotonicity.

Example 14. *Consider a setting where players have a small "preference for honesty." To be precise, assume that for each player i, $E_i = \Theta$ and the cost of sending evidence 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. Note that in this case, for each θ and θ', $c_i(\theta',\theta') = 0 < \varepsilon = c_i(\theta',\theta)$ and so condition (CM**) is violated. In fact, the existence of just one player having a preference for honesty is enough for (CM**) to be violated. In this case, any SCR is cost-monotonic.*

The following necessity result is analogous to Theorem 1.

Theorem 5. *If f is implementable then f is cost-monotonic.*

Proof. Since f is implementable, there exists a mechanism (M,g,τ) that implements f. Hence, for each θ, for each $a \in f(\theta)$, there exists $(m_{\theta,a}, e_{\theta,a}) \in M \times E^\theta$ Nash equilibrium at θ such that $g(m_{\theta,a}, e_{\theta,a}) = a$ and $\tau(m_{\theta,a}, e_{\theta,a}) = 0$. Set for each $\theta : e^*_{\theta,a} = e_{\theta,a}$. Now consider any θ' satisfying (CM*) and (CM**) above. Note that because $(m_{\theta,a}, e_{\theta,a}) \in M \times E^\theta$ is a (costless) Nash equilibrium at θ :

$$u_i(g(m_{\theta,a}, e_{\theta,a}),\theta) \geq u_i(g(m'_i, e'_i, m_{-i,\theta,a}, e_{-i,\theta,a}),\theta) - \tau_i(m'_i, e'_i, m_{-i,\theta,a}, e_{-i,\theta,a}) - c_i(e'_i,\theta)$$

for any $(m'_i, e'_i) \in M_i \times E_i$. By (CM*),

$$u_i(g(m_{\theta,a}, e_{\theta,a}),\theta') \geq u_i(g(m'_i, e'_i, m_{-i,\theta,a}, e_{-i,\theta,a}),\theta') - \tau_i(m'_i, e'_i, m_{-i,\theta,a}, e_{-i,\theta,a}) - c_i(e'_i,\theta)$$

for any $(m'_i, e'_i) \in M_i \times E_i$. Now by (CM**):$e_{i,\theta} = e^*_{i,\theta,a} \in E_i^{\theta'}$ and $c_i(.,\theta') \geq c_i(.,\theta)$, this implies that $(m_\theta, e_\theta) \in M \times E^{\theta'}$ and

$$u_i(g(m_{\theta,a}, e_{\theta,a}),\theta') \geq u_i(g(m'_i, e'_i, e_{-i,\theta,a}, m_{-i,\theta,a}),\theta') - \tau_i(m'_i, e'_i, m_{-i,\theta,a}, e_{-i,\theta,a}) - c_i(e'_i,\theta')$$

for any $(m'_i, e'_i) \in M_i \times E_i$ and so $(m_{\theta,a}, e_{\theta,a})$ is a Nash equilibrium at θ'. Hence $a \in f(\theta')$ as claimed. □

To see how Theorem 5 has bite, consider the following example of education signaling.

[33]To avoid any confusion: $c_i(\cdot,\theta') \geq c_i(\cdot,\theta)$ means that for all $e_i \in E_i$, $c_i(e_i,\theta') \geq c_i(e_i,\theta)$.

Example 15. *There are n workers, each with an ability level measuring her marginal productivity. The state of the world is a vector of abilities. For convenience, assume no two workers have the same ability, i.e.* $\Theta = \{\theta \in \mathbb{R}_+^n : i \neq j \implies \theta_i \neq \theta_j\}$. *There is one job that must be allocated to a single worker along with a wage, so that* $A = \{1, \dots, n\} \times \mathbb{R}_+$. *The SCR is* $f(\theta) = (i^*, \theta_{i^*})$ *where* $\{i^*\} = \arg\max_i \theta_i$, *i.e. allocate the job to the most able worker and pay her marginal product. Agent i's utility from an outcome $a = (a_1, a_2)$ is state-independent: $u_i(a, \theta) = 0$ if $a_1 \neq i$ (he does not get the job), and $u_i(a, \theta) = a_2$ otherwise. Note that consequently, (CM*) is satisfied for all a.*

*Suppose now that agents can signal their ability through a choice of education, $e_i \in \mathbb{R}_+$. The cost of education for a worker only depends on his individual type, so that we abuse notation and write $c_i(e_i, \theta_i)$. Assume the cost of education is non-negative and satisfies two properties: for all i and θ_i, $c_i(e_i, \theta_i) = 0$ if and only if $e_i = 0$; and for all i and e_i, $c_i(e_i, \theta_i) \geq c_i(e_i, \theta_i')$ if and only if $\theta_i \leq \theta_i'$. (While it may be natural to also impose the Spencian property that the marginal cost is decreasing in individual ability, it is not required.) Then $E_i^\theta = \{0\}$ for all i, θ, and consequently, (CM**) is satisfied if and only if $\theta' \leq \theta$ (in the sense of the usual vector order). Since we can make all workers' ability strictly decrease but change the socially desired outcome, it follows that f is not cost-monotonic, and by Theorem 5, not implementable.*[34]

*On the other hand, consider a modified environment where workers intrinsically enjoy or tolerate different levels of education as a function of their ability. Specifically, for all i,e_i, and θ_i, $c_i(e_i, \theta_i) = 0$ if and only if $e_i \leq \theta_i$. In other words, a type θ_i worker finds education costly only if he must acquire more than θ_i of it. Further assume, naturally, that if $c_i(e_i, \theta_i) > 0$ and $\theta_i' < \theta_i$, then $c_i(e_i, \theta_i') > c_i(e_i, \theta_i)$. It is straightforward to see that by setting $e_{i,\theta,a}^* = \theta_i$, (CM**) is never satisfied, so that f—or any SCR, for that matter—is now cost-monotonic.*

The next result shows that if there are three or more players, cost-monotonicity is also sufficient for implementation. This is analogous to Theorem 2, noting that because we now allow explicitly for transfers that must not be used in equilibrium, No veto power and non-satiation are automatically satisfied.

Theorem 6. *Assume $n \geq 3$. If f is cost-monotonic then f is implementable.*

Proof. The proof is by construction of a mechanism that implements any cost-monotonic SCR f when $n \geq 3$. Fix, for all i, $M_i = \Theta \times A \times \mathbb{N} \times \mathbb{R}$. The first three components are the same as used in the proof of Theorem 2, the fourth is a new element where the player requests a transfer to him. Now define g and τ according to the following rules:

(1) If $m_1 = \cdots = m_n = (\theta, a, k, \alpha)$ with $a \in f(\theta)$, and $e = e_{\theta,a}^*$, then $g(m, e) = a$ and $\tau(m, e) = (0, \dots, 0)$.
(2) If $\exists i$ s.t. $\forall j \neq i$, $e_j = e_{j,\theta,a}^*$ and $m_j = (\theta, a, k, \alpha)$ with $a \in f(\theta)$, and either $m_i = (\theta', b, l, \hat{\alpha}) \neq (\theta, a, k, \alpha)$ or $e_i \neq e_{i,\theta,a}^*$ then

[34] Note that this is true even if the allocation only consists of choosing which worker to assign the job to, maintaining a fixed wage that is independent of ability.

a) if $e_i \in E_i^{\theta}$, then

$$\tau_i(m,e) = \begin{cases} -\hat{\alpha} \text{ if } u_i(a,\theta) \geq u_i(b,\theta) - \hat{\alpha} \\ 0 \text{ if } u_i(b,\theta) - \hat{\alpha} > u_i(a,\theta), \end{cases}$$

and

$$g(m,e) = \begin{cases} b \text{ if } u_i(a,\theta) \geq u_i(b,\theta) - \hat{\alpha} \\ a \text{ if } u_i(b,\theta) - \hat{\alpha} > u_i(a,\theta). \end{cases}$$

In addition, $\tau_{-i}(m,e) = \left(-\frac{\tau_i(m,e)}{n-1}, ..., -\frac{\tau_i(m,e)}{n-1}\right)$.

b) if $e_i \notin E_i^{\theta}$, then $\tau_i(m,e) = c_i(e_i,\theta)$, and $g(m,e) = a$. Here again $\tau_{-i}(m,e) = \left(-\frac{\tau_i(m,e)}{n-1}, \ldots, -\frac{\tau_i(m,e)}{n-1}\right)$,

(3) For any other (m,e), letting $m_i = (\theta_i, a_i, k_i, \alpha_i)$ and $i^* = \min_{i \in I} \arg\max_{j \in I} k_j$, $g(m,e) = a_{i^*}$, $\tau_{i^*}(m,e) = \alpha_{i^*}$ and $\tau_{-i^*}(m,e) = \left(-\frac{\tau_{i^*}(m,e)}{n-1}, \ldots, -\frac{\tau_{i^*}(m,e)}{n-1}\right)$.

Step 1. We wish to show that for any θ and any $a \in f(\theta)$, $a \in O(M,g,\theta)$. It suffices to show that at state θ, for any $(k,\alpha) \in \mathbb{N} \times \mathbb{R}$, all agents i playing $m_i = (\theta, a, k, \alpha)$ with $a \in f(\theta)$ and $\forall i$, $e_i = e^*_{i,\theta}$, is a NE. This strategy profile results in outcome a and moreover, since $e^*_{i,\theta} \in E_i^{\theta}$: $c_i(e_i,\theta) = 0$. If some agent deviates, rule (2) of the mechanism applies. There is no profitable deviation for any player to rule (2a) since any such deviation yields a weakly worse outcome and the same evidence cost. There is no profitable deviation to rule (2b) since this yields the same outcome and exactly compensates the agent for the evidence falsification cost.

For the remainder of the proof, suppose the true state is θ', and (m,e) is a NE. We must show that $g(m,e) \in f(\theta')$.

Step 2. No NE can fall into rule (3). For if it does, note that all players other than i^* are receiving a transfer of at most 0, and hence any player $j \neq i^*$ can deviate and induce the same outcome at the same evidence cost and increase his monetary transfer to any $\alpha_j > 0$.[35] In the same way, no NE can fall into rule (2). For if it does, any agent $j \neq i$ can deviate and play a strategy that results in rule (3), induce the same outcome at the same evidence cost and increase his monetary transfer.

Step 3. It remains to consider the NE falling into rule (1). Here $g(m,e) = a$. By the usual Maskin-mechanism argument, condition (CM*) must hold, since a player can always deviate into rule (2a) while producing the same evidence, a player can always get any outcome b and transfer $\hat{\alpha}$ such that $u_i(a,\theta) \geq u_i(b,\theta) - \hat{\alpha}$, because (m,e) is a NE, this implies that $u_i(a,\theta') \geq u_i(b,\theta') - \hat{\alpha}$ which yields condition (CM*). If condition (CM**) also holds, cost-monotonicity implies that $a \in f(\theta')$, and we are done. To see that condition (CM**) must hold, observe that if it were not the case, this can be due first to the fact that there is some player $i : e^*_{i,\theta,a} \notin E_i^{\theta'}$, in this case player i can profitably deviate by announcing $e^*_{i,\theta',a} \in E_i^{\theta'}$. This is true since, by rule (2a), this induces the same outcome a and saves on evidence falsification cost and by rule (2b), this induces the same outcome and saves on evidence falsification cost, and strictly benefits (in case the deviation induces

[35] Since their is no bound on the set of possible transfers, even player i^* has an incentive to deviate to a higher monetary transfer, but this is inessential. This part of the proof would work with an arbitrarily small positive upper bound on transfers.

rule (2b) $e^*_{i,\theta',a} \notin E^\theta_i$ and so: $c_i\left(e^*_{i,\theta',a}, \theta\right) > 0$) in monetary transfer. The failure of (CM**) can also be due to the fact that there is some i and $\tilde{e}_i$ such that

$$c_i\left(\tilde{e}_i, \theta'\right) < c_i\left(\tilde{e}_i, \theta\right). \tag{8}$$

This implies that $\tilde{e}_i \notin E^\theta_i$, and since $e_i = e^*_{i,\theta,a} \in E^\theta_i$, $e_i \neq \tilde{e}_i$. Consider a deviation for agent i to $\tilde{e}_i$ and $m_i = (\theta, a, k, \alpha)$. By rule (2b), this generates a payoff to i of $u_i(a, \theta') + c_i\left(\tilde{e}_i, \theta\right) - c_i\left(\tilde{e}_i, \theta'\right)$, which by inequality (8) is greater than i's equilibrium payoff of $u_i(a, \theta')$, a contradiction.

Step 4. Finally, to show that no evidence cost is incurred in equilibrium, we will show $e \in E^{\theta'}$, so that for each $i : c_i\left(e_i, \theta'\right) = 0$. But, we know that any NE is in rule (1), thus, $(m, e) \in NE\left(M, g, \theta'\right) \implies e = e^*_{\theta,a}$ for some θ. Since we have shown above that (CM**) must hold, it follows that and so $e = e^*_{\theta,a} \in E^{\theta'}$, as needed. □

Remark 15. Even though the definition of implementation did not require it, the proof of Theorem 6 shows that implementation can be achieved while maintaining a global balanced budget, i.e. even off the equilibrium path.

Remark 16. The proof of the Theorem reveals that it is not essential that transfers be unbounded when evidence fabrication costs are not. Even if we assumed that for each player i and $(m, e) \in M \times E : 0 \leq \tau_i(m, e) \leq \max_{e_i,\theta} c_i\left(e_i, \theta\right)$, the argument would remain valid.[36]

Theorem 6 immediately implies the following sharp result.

Corollary 7. *If* $n \geq 3$ *and*

$$\forall \theta, \theta' : \theta' \neq \theta \implies \exists i, e_i \text{ s.t. } c_i(e_i, \theta') < c_i(e_i, \theta), \tag{9}$$

then any SCR is implementable.

Proof. (9) implies that (CM**) is violated for any θ, θ'. Hence, every SCR is cost-monotonic, and Theorem (6) delivers the conclusion. □

Condition (9) is satisfied if for every pair of distinct states θ, θ', there is some i with $E^\theta_i \neq E^{\theta'}_i$. Thus, when $n \geq 3$, any SCR can be implemented if players have a small preference for honesty as in Example 14, and the planner can make sufficient off-path transfers (which may only need to be quite small, following Remark 16). Matsushima (2008) obtains a result that is similar in spirit in a setting where the set of outcomes as well as the set of states is finite. The underlying argument is quite different because Matsushima (2008) uses proof techniques from the virtual implementation literature, in particular lotteries over outcomes.

Remark 17. One may wonder whether the permissiveness of Theorem 6 is driven by the assumption that the planner can make transfers off the equilibrium path or the presence of discriminatory signals. We refer to Benoît and Ok (2008) and Sanver (2006) for a detailed discussion of what is implementable with off-path transfers in the standard environment without evidence. For our purposes, let us note that if all evidence were costless to all players in all states—which corresponds

[36]Since $c_i(\cdot, \cdot)$ is defined over $\mathbb{R}_+ \cup \{+\infty\}$, note that the max can be equal to $+\infty$.

to the standard setting without evidence—then condition (CM**) is necessarily satisfied and can be dropped altogether from the definition of cost-monotonicity. Theorems 5 and 6 continue to apply, but evidently this version of cost-monotonicity can be significantly more demanding than in the presence of evidence costs. For example, if every player's cardinal preferences are state-independent (i.e., for all i, a, b, θ, and θ', $u_i(a,\theta) - u_i(b,\theta) = u_i(a,\theta') - u_i(b,\theta')$), and there is no evidence, then only constant SCRs are cost-monotonic. On the other hand, by Corollary 7, the picture dramatically changes with a small preference for honesty or, more generally, arbitrarily small costs that satisfy (9).

4. Conclusion

This paper has generalized the implementation problem to incorporate agents' ability to signal something about the state of the world. The central idea is that the planner can use either agents' preferences over outcomes or their signaling technology to discriminate between states of the world. We have studied both hard evidence, when players can prove that the state lies in some subset of all possible states, and the costly fabrication of evidence, where the costs of fabrication can vary across states. The various results we have provided show precisely how a wider class of social choice rules are implementable as a function of the evidence structure. In particular, we have formulated an appropriate generalization and weakening of Maskin-monotonicity—evidence-monotonicity in the case of hard evidence, or even more generally, cost-monotonicity—and shown that this is the key to implementation in our framework. Natural evidence structures yield quite permissive results for implementation.

There are a number of directions that this research can be taken in. 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. In ongoing work, Kartik and Tercieux (2008a), we study implementation with hard evidence in Bayesian Nash equilibrium in incomplete information environments. We extend the notion of evidence-monotonicity to *Bayesian evidence-monotonicity*. Bayesian evidence-monotonicity generalizes Jackson's (1991) Bayesian monotonicity condition, and together with the usual incentive compatibility condition is the key for Bayesian implementation with hard evidence.

Within the complete information framework, it would also be useful to understand how evidence changes the implementation problem when one restricts attention to "nice" mechanisms, for example, "bounded mechanisms" (Jackson, 1992). We are optimistic that based on the current paper and Ben-Porath and Lipman's (2008) interesting results for subgame perfect implementation with hard evidence, this will be a fruitful avenue. A related question is the possibility of (Nash) implementation without using integer games or similar constructions. One possibility would be to consider quasi-linear settings where large transfers are available, as Goltsman (2009) has studied in the standard setting without evidence. We conjecture that when the evidence structure satisfies Universal Distinguishability (UD), any SCR is implementable in such environments.

Appendix A. Omitted Proofs

Proof of Proposition 3. We first note that f is evidence-monotonic if and only if

$$\forall \theta, a \in f(\theta),\ \exists e^*_{\theta,a} \in E^\theta \text{ s.t. } \forall \theta' \in T^f(\theta,a),\ \exists i \in N : \left(e^*_{i,\theta,a} \notin E_i^{\theta'}\right) \vee \left(E_i^{\theta'} \nsubseteq E_i^\theta\right). \tag{10}$$

In the sequel, we will use this equivalent formulation. For the "if part" of the result, assume that for each θ, $a \in f(\theta)$ and $\Omega \subseteq T^f(\theta,a) : E^\theta \neq \bigcup_{\theta' \in \Omega} E^{\theta'}$. Proceed by contradiction and assume that (10) is false. This implies that there exists θ and $a \in f(\theta)$ s.t. for all $e \in E^\theta$ there exists $\theta'(e) \in T^f(\theta,a)$, for which we have that for all $i \in N : \left(e_i \in E_i^{\theta'(e)}\right) \wedge \left(E_i^{\theta'(e)} \subseteq E_i^\theta\right)$. Set $\Omega := \bigcup_{e \in E^\theta} \theta'(e)$. Note that $\Omega \subseteq T^f(\theta,a)$. Since for each $e \in E^\theta$ it is the case that $e \in E^{\theta'(e)}$, this shows that $E^\theta \subseteq \bigcup_{\theta' \in \Omega} E^{\theta'}$. Finally, for each $e \in E^\theta : E^{\theta'(e)} \subseteq E^\theta$, hence we have $\bigcup_{\theta' \in \Omega} E^{\theta'} \subseteq E^\theta$, and so $E^\theta = \bigcup_{\theta' \in \Omega} E^{\theta'}$, a contradiction with the assumption. For the "only if part", assume that f satisfies (10) and proceed again by contradiction assuming that for some θ, some $a \in f(\theta)$ and some $\Omega \subseteq T^f(\theta,a) : E^\theta = \bigcup_{\theta' \in \Omega} E^{\theta'}$. This implies that for some θ and some $a \in f(\theta)$: (i) for all $e \in E^\theta$, there exists $\theta'(e) \in \Omega \subseteq T^f(\theta,a)$ s.t. $e \in E^{\theta'(e)}$ and (ii) $E^{\theta'(e)} \subseteq E^\theta$, which contradicts the assumption that (10) is true. □

Proof of Remark 10. Necessity is immediate, so we prove sufficiency. Assume

$$\{(\theta,\theta') : E^\theta \neq E^{\theta'}\} \subseteq \{(\theta,\theta') : \tilde{E}^\theta \neq \tilde{E}^{\theta'}\}. \tag{11}$$

Suppose θ and $\Omega \subseteq \Theta$ are distinguishable under $\mathcal{E}$. Then θ is distinguishable from any $\theta' \in \Omega$ under $\mathcal{E}$, any by (11), θ is distinguishable from any $\theta' \in \Omega$ under $\tilde{\mathcal{E}}$. Since $\tilde{\mathcal{E}}$ is normal, Proposition 2 implies that θ and Ω are distinguishable under $\tilde{\mathcal{E}}$. □

Proof of Remark 11. Suppose [not $\tilde{\mathcal{E}} \blacktriangleright \mathcal{E}$]. Then there exists a state θ^* and an event $\Omega^* \subseteq \Theta \setminus \{\theta^*\}$ that are distinguishable under $\mathcal{E}$ but not under $\tilde{\mathcal{E}}$. Let outcomes $x, y \in A$. Consider state independent preferences for all players except player 1, whose preferences are as follows: for any $\theta \in \Omega^* \cup \{\theta^*\}$, $u_1(y,\theta) > u_1(x,\theta) > u_1(a,\theta)$ for all $a \notin \{x,y\}$; for any $\theta \in \Omega \setminus \Omega^* \setminus \{\theta^*\}$, $u_1(x,\theta) > u_1(y,\theta) > u_1(a,\theta)$ for all $a \notin \{x,y\}$. Consider the SCR f given by $f(\theta) = \{y\}$ for all $\theta \in \Omega^*$ and $f(\theta) = \{x,y\}$ for all $\theta \notin \Omega^*$. It can be verified that $T^f(\theta^*,x) = \Omega^*$ and $T^f(\cdot,\cdot) = \emptyset$ otherwise. By Proposition 3, f is evidence-monotonic under $\mathcal{E}$ but not under $\tilde{\mathcal{E}}$. □

Proof of Theorem 3. (Necessity.) This is implied by Theorem 1.

(Sufficiency.) For each θ and $a \in f(\theta)$, let $e^*_{\theta,a}$ be the evidence profile per Definition 2. We will construct a mechanism that implements f. Fix, for all i,

$$M_i = \{(\theta,a,b,r,k) \in \Theta \times A \times A \times \{F, NF\} \times \mathbb{N} : a \in f(\theta)\},$$

where $\{F, NF\}$ is the set comprising the 2 elements "flag" and "no flag". Define g via the following four rules:

(1) If $m_1 = (\theta, a, b_1, NF, k_1)$ and $m_2 = (\theta, a, b_2, NF, k_2)$ and $e = e^*_{\theta,a}$ then $g(m,e) = a$.
(2) If $m_1 = (\theta_1, a_1, b_1, NF, k_1)$ and $m_2 = (\theta_2, a_2, b_2, NF, k_2)$ and Rule 1 does not apply then $g(m,e) = z$ where z is the worst outcome.
(3) If $\exists i$ s.t. $m_i = (\theta_i, a_i, b_i, F, k_i)$ and for $j \neq i : m_j = (\theta_j, a_j, b_j, NF, k_j)$ then

$$\text{a) If } e_i \in E_i^{\theta_j} \text{ then } g(m,e) = \begin{cases} a_j \text{ if } u_i(a_i, \theta_j) \geq u_i(a_j, \theta_j) \\ a_i \text{ if } u_i(a_j, \theta_j) > u_i(a_i, \theta_j) \end{cases}.$$

b) if $e_i \notin E_i^{\theta_j}$, then $g(m,e) = b_i$.

(4) If $m_1 = (\theta_1, a_1, b_1, F, k_1)$ and $m_2 = (\theta_2, a_2, b_2, F, k_2)$, then $g(m,e) = a_{i^*}$ where $i^* = \min_{i \in \{1,2\}} \arg\max_{j \in \{1,2\}} k_j$.

Step 1. We first show that for any θ and any $a \in f(\theta)$, $a \in O(M, g, \theta)$. It suffices to show that at state θ, for any $k \in \mathbb{N}$, each agent i playing $m_i = (\theta, a, a, NF, k)$ and sending $e_i = e^*_{i,\theta,a}$ is a NE, since by rule (1) of the mechanism, this results in outcome a. If some agent i deviates from this strategy profile, then only cases (2) or (3a) of the mechanism can apply. There is no profitable deviation for any player to rule (2) or (3a) since any such deviation yields a (weakly) worse outcome.

For the remainder of the proof, suppose the true state is θ', and (m,e) is a NE. We must show that $g(m,e) \in f(\theta')$.

Step 2. Any (m,e) that falls into Rule (2) is not a NE. To see this, recall that $g(m,e) = z$ and note that from (m,e), each player i can deviate by raising a flag and reach Rule (3). So consider a deviation for player 1 to $\tilde{m}_1 = (\cdot, a_2, a_2, F, \cdot)$. This leads to outcome $a_2 \in f(\theta_2)$, which player 1 strictly prefers to z since z is a bad outcome.

Step 3. Assume (m,e) falls into rule (3a). Without loss of generality, assume that $m_1 = (\theta_1, a_1, b_1, F, k_1)$ and $m_2 = (\theta_2, a_2, b_2, NF, k_2)$. Note that player 2 can raise a flag inducing rule (4). Thus, announcing an integer strictly higher than the one of player 1, he will get any outcome he wants. Hence, $g(m,e) \in \arg\max_{c \in A} u_2(c, \theta')$. In addition, player 1 can deviate to $(\theta_1, a_2, b_1, F, k_1)$ while producing the same evidence and so get outcome $a_2 \in f(\theta_2)$. Hence, we get $u_1(g(m,e), \theta') \geq u_1(a_2, \theta')$ and so RVP ensures that $g(m,e) \in f(\theta')$. Now, if (m,e) falls into rule (3b), player i can get any outcome he wants, and his opponent get any outcome he wants by raising a flag and announcing an integer strictly higher than player i. Hence, $g(m,e) \in \bigcap_{i \in \{1,2\}} \arg\max_{c \in A} u_i(c, \theta')$. Restricted no veto power here again completes the proof. A similar argument applies if (m,e) falls into rule (4) as well.

Step 4. It remains to consider (m,e) falling into rule (1), i.e. $m_1 = (\theta, a, b_1, NF, k_1)$ and $m_2 = (\theta, a, b_2, NF, k_2)$ while $e = e^*_{\theta,a}$. The outcome is $g(m,e) = a$. Observe that any player i can raise a flag and deviate into rule (3a) while producing the same evidence, and hence can induce any outcome in the set $L(a,\theta) := \{b : u_i(a,\theta) \geq u_i(b,\theta)\}$. Since (m,e) is a NE, this implies that for any agent i and outcome $b \in L(a,\theta)$, $u_i(a,\theta') \geq u_i(b,\theta')$, and consequently condition (*) is satisfied. If (**) also holds, evidence-monotonicity implies that $a \in f(\theta')$, and we are done. To see that (**) must hold, suppose to contradiction that it does not. Then there is some agent i such that $\left(e^*_{i,\theta,a} \notin E_i^{\theta'}\right) \vee \left(E_i^{\theta'} \nsubseteq E_i^{\theta}\right)$. Since $e_i = e^*_{i,\theta,a}$ and the true state is θ', we have $e^*_{i,\theta,a} \in E_i^{\theta'}$, and consequently $E_i^{\theta'} \nsubseteq E_i^{\theta}$. This implies that there is some $\tilde{e}_i \in E_i^{\theta'}$ such that $\tilde{e}_i \notin E_i^{\theta}$, and agent i can deviate into rule (3b), and get any outcome he desires. Hence, in this case, $a \in \arg\max_{c \in A} u_i(c, \theta')$. Note also that $a \in f(\theta)$. Hence, RVP applies and we get that $g(m,e) \in f(\theta')$. □

Proof of Theorem 4. (Necessity.) Fix any implementing mechanism, (M, g). It is straightforward that condition (5) must hold. Fix θ and $a \in f(\theta)$. To see that (6) must also hold, let $e^*_{\theta,a}$ be the evidence sent when the state is θ at some NE $(m^*_{\theta,a}, e^*_{\theta,a})$ such that $g(m^*_{\theta,a}, e^*_{\theta,a}) = a$. Now pick any θ' and $b \in f(\theta')$ and assume that $u(a,\theta') \geq u(b,\theta')$, and $e^*_{\theta,a} \in E^{\theta'}$. Note that since $b \in f(\theta')$, there exists $(m^*_{\theta',b}, e^*_{\theta',b})$ NE at θ' so that $g(m^*_{\theta',b}, e^*_{\theta',b}) = b$. Since $(m^*_{\theta,a}, e^*_{\theta,a}) \in M \times E^{\theta'}$ it is easily checked that $u(a,\theta') \geq u(b,\theta')$ implies that $(m^*_{\theta,a}, e^*_{\theta,a})$ is a NE at θ'. Hence, $a \in f(\theta')$.

Let us now show that condition (7) is also necessary. Fix any $T \subseteq \Theta$ such that $\bigcap_{\theta'' \in T} E^{\theta''} \neq \emptyset$, and pick any $x \in \bigcap_{\theta'' \in T} E^{\theta''}$. Set $c := g(m,x)$ for some arbitrary $m \in M$. Fix any $\theta \in T$, $b \in f(\theta)$, and assume that $c \notin f(\theta)$. We must show that $u(b,\theta) > u(c,\theta)$. Since $c \notin f(\theta)$, $(m,x) \notin NE(M,g,\theta)$. Since by construction, $(m,x) \in M \times E^{\theta}$, there exists some $(\tilde{m},\tilde{x}) \in M \times E^{\theta}$ such that $u(g(\tilde{m},\tilde{x}),\theta) > u(c,\theta)$. Since $b \in O(M,g,\theta)$, we have $u(b,\theta) \geq u(g(\tilde{m},\tilde{x}),\theta)$. Hence, $u(b,\theta) > u(c,\theta)$, as claimed.

(Sufficiency.) Set $E^* := \bigcup_{\theta\in\Theta} \bigcup_{a \in f(\theta)} e^*_{\theta,a}$. Let $M = A$, and define the outcome function g according to the following rules:

(1) If $(m,e) \in A \times E^*$
  a) If $(m,e) = (a, e^*_{\theta,a})$ for some θ and $a \in f(\theta)$ then $g(m,e) = a$;
  b) Otherwise, pick any θ and $a \in f(\theta)$ such that $e = e^*_{\theta,a}$ and set $g(m,e) = a$;

(2) If $(m,e) \notin A \times E^*$, let $T := \{\theta \in \Theta : e \in E^{\theta}\}$; since $\bigcap_{\theta' \in T} E^{\theta'} \neq \emptyset$, set $g(e) = c$ where c is as in (7).

Step 1. We wish to show that for any θ and any $a \in f(\theta)$, $a \in O(M,g,\theta)$. To see this, we show that $\left(a, e^*_{\theta,a}\right)$ is an equilibrium strategy. Let us first show that there is no profitable deviation to $(m,e) \in A \times E^*$. Consider a deviation that falls into rule (1a) i.e. $(m,e) = (b, e^*_{\theta',b})$ for some θ' and $b \in f(\theta')$. If $b \in f(\theta)$ then, by (5), this is not profitable. If $b \notin f(\theta)$ then, note that $e^*_{\theta',b} \in E^{\theta}$ and so by condition (6), $u(a,\theta) > u(b,\theta)$ and so the deviation is not profitable. For deviations falling into rule (1b), a similar reasoning applies. Finally, consider a deviation falling into rule (2) i.e. $(m,e) \notin A \times E^*$. In this case $g(e) = c$. If $c \in f(\theta)$ then, by (5), the deviation is not profitable. If $c \notin f(\theta)$, by condition (7) and by the definition of T above (obviously, $e \in E^{\theta}$ and so $\theta \in T$) that $u(a,\theta) > u(c,\theta)$ and so here again the deviation is not profitable.

For the remainder of the proof, suppose the true state is θ, and (m,e) is a NE. We must show that $g(m,e) \in f(\theta)$.

Step 2. Suppose (m, e) falls into rule (1a) or (1b). Take any θ' and $b \in f(\theta')$ such that $e = e^*_{\theta',b}$. To show that $b \in f(\theta)$, observe that if it were not the case, then since $e^*_{\theta',b} \in E^\theta$ by condition (6), we get that for $a \in f(\theta) : u(a, \theta) > u(b, \theta)$ and so there would exist a profitable deviation to $(a, e^*_{\theta,a})$ where $a \in f(\theta)$ a contradiction.

Step 3. Suppose (m, e) falls into rule (2). In this case $g(m, e) = c$. Again to show that $c \in f(\theta)$, note that if it were not the case, by condition (7) and by the definition of T above (obviously, $e \in E^\theta$ and so $\theta \in T$) we would have $u(a, \theta) > u(c, \theta)$ for $a \in f(\theta)$ and so there would exist a profitable deviation to $(a, e^*_{\theta,a})$ where $a \in f(\theta)$ a contradiction. □

Proof of Remark 12. We provide two examples to show that if either existence of a bad outcome or RVP fails, then an evidence-monotonic SCR may not be implementable when $n = 2$. The examples also show that condition (UD) is not sufficient to implement a SCR if either there is no bad outcome or RVP fails.

Example 16. $\Theta = \{\theta_1, \theta_2\}$. *The evidence structure is as follows:* $E_1^{\theta_1} = \{\theta_1, \theta_2\}$ *and* $E_1^{\theta_2} = \{\theta_2\}$*; symmetrically,* $E_2^{\theta_1} = \{\theta_1\}$ *and* $E_2^{\theta_2} = \{\theta_1, \theta_2\}$*. The set of outcomes is* $A = \{a, b, c, d\}$*. Preferences are state independent: for all* θ, $u_1(c, \theta) > u_1(b, \theta) > u_1(a, \theta) > u_1(d, \theta)$ *and* $u_2(d, \theta) > u_2(a, \theta) > u_2(b, \theta) > u_2(c, \theta)$*. The SCR is given by* $f(\theta_1) = a$ *and* $f(\theta_2) = b$*. This SCR satisfies RVP (and non-satiation), but there is no bad outcome. Note that the evidence structure satisfies (UD), so that* f *is evidence-monotonic by Proposition 3.*

However, f *is not implementable. To see this, assume to contradiction that* f *is implementable. Let* $\left(s^*_{1,\theta_1}, s^*_{2,\theta_1}\right)$ *and* $\left(s^*_{1,\theta_2}, s^*_{2,\theta_2}\right)$ *be NE at state* θ_1 *and state* θ_2 *respectively. We must have* $g\left(s^*_{1,\theta_1}, s^*_{2,\theta_1}\right) = a$ *and* $g\left(s^*_{1,\theta_2}, s^*_{2,\theta_2}\right) = b$*. Since* $s^*_{1,\theta_2} \in M_1 \times E_1^{\theta_2} \subseteq M_1 \times E_1^{\theta_1}$ *and* $\left(s^*_{1,\theta_1}, s^*_{2,\theta_1}\right)$ *is a NE at state* θ_1*, we must have* $g\left(s^*_{1,\theta_2}, s^*_{2,\theta_1}\right) \in \{a, d\}$*; otherwise at state* θ_1*, player 1 would have an incentive to deviate from* $\left(s^*_{1,\theta_1}, s^*_{2,\theta_1}\right)$ *playing* s^*_{1,θ_2}*. But note now that* $s^*_{2,\theta_1} \in M_2 \times E_2^{\theta_1} \subseteq M_2 \times E_2^{\theta_2}$ *and so at state* θ_2*, player 2 can deviate from* $\left(s^*_{1,\theta_2}, s^*_{2,\theta_2}\right)$ *and get an outcome in* $\{a, d\}$ *by playing* s^*_{2,θ_1}*, which is strictly better for him than the equilibrium outcome* b*, a contradiction.*

Example 17. $n = 2$*;* $\Theta = \{\theta_1, \theta_2\}$*;* $A = \{a, b, c\}$*;* $f(\theta_1) = a$ *and* $f(\theta_2) = b$*; for* $i = \{1, 2\}$*:* $u_i(b, \theta_1) > u_i(a, \theta_1) > u_i(c, \theta_1)$ *and* $u_i(a, \theta_2) > u_i(b, \theta_2) > u_i(c, \theta_2)$*; for* $i = \{1, 2\}$*:* $E_i^{\theta_1} = \{\theta_1\}$ *and* $E_i^{\theta_2} = \{\theta_1, \theta_2\}$*. The evidence structure satisfies (UD), hence* f *is evidence-monotonic by Proposition 3. Outcome* c *is a bad outcome, since it is not in the range of the SCR and is the bottom-ranked outcome for both agents in both states.*

The SCR f *is not implementable by the following argument: if mechanism* (M, g) *implements* f*, there must exist some* $m^* \in M$ *such that* $g(m^*, (\theta_1, \theta_1)) = a$*, since* $f(\theta_1) = a$*. But then,* $(m^*, (\theta_1, \theta_1))$ *is a Nash equilibrium at state* θ_2 *since* a *is top-ranked by both agents in state* θ_2*, contradicting* $f(\theta_2) = b$*.*

□

Appendix B. Dynamic Mechanisms

Let us first formally define a dynamic mechanism, which we refer to as just "mechanism" hereafter. A mechanism is a tuple (M, H, g). For each player i, there is a set of cheap talk messages M_i and as before $M = M_1 \times \cdots \times M_n$. We introduce a new message available to the agent at each state $e_i^{\emptyset}$ that is interpreted as "player i sends no evidence". We will now define H the set of possible histories. Before doing so, let us define recursively the following set $\bar{H}$. First, $\emptyset \in \bar{H}$ and second if $h = (h_1, h_2, ..., h_k) \in \bar{H}$ then for any $X \subseteq I$ and $h_{k+1} \in \prod_{i \in X} \left(\left(E_i \cup \{e_i^{\emptyset}\} \right) \times M_i \right)$, $(h_1, h_2, ..., h_k, h_{k+1}) \in \bar{H}$. Given an history $h = (h_1, ..., h_k, h_{k+1}) \in \bar{H}$, we know that $h_{k+1} \in \prod_{i \in X} \left(\left(E_i \cup \{e_i^{\emptyset}\} \right) \times M_i \right)$ for some $X \subseteq I$. In line with this, we will write $h_{k+1} = \left\{ (e_i^{k+1}, m_i^{k+1}) \right\}_{i \in X}$.

Now the set of possible histories H is a subset of $\bar{H}$ satisfying the following properties. First if $h \in H$ it is of finite lenght i.e. there is some k such that $h = (h_1, ..., h_k)$. Now denote H_T the set of terminal histories in H i.e. the set of histories $h = (h_1, ..., h_k) \in H$ such that there is no h_{k+1} such that $(h_1, h_2, ..., h_k, h_{k+1}) \in H$. We assume, as in Bull and Watson (2007), that extensive forms are restricted to those in which each player sends an evidence exactly once in every path through the tree[37] i.e. for any $h = (h_1, h_2, ..., h_k) \in H_T$, and for any i, there exists a unique $\tilde{k}$ such that $e_i^{\tilde{k}} \neq e_i^{\emptyset}$ (note that this implies that each player plays at least once along any history). We will also impose the following consistency condition that if $(h_1, h_2, ..., h_k, h_{k+1}) \in H$, then $(h_1, h_2, ..., h_k) \in H$. Finally, our assumption that players cannot be compelled or forbidden to present evidence takes the following form. Let $(h_1, h_2, ..., h_k, h_{k+1}) \in H$ and denote X the subset of I such that $h_{k+1} \in \prod_{i \in X} \left(\left(E_i \cup \{e_i^{\emptyset}\} \right) \times M_i \right)$. Pick any $\hat{h}_{k+1} \in \prod_{i \in X} \left(\left(E_i \cup \{e_i^{\emptyset}\} \right) \times M_i \right)$ consistent with the assumption that each player sends an evidence exactly once along any history, i.e. such that if there is $i \in X$ and $k' \leq k$ such that $e_i^{k'} \neq e_i^{\emptyset}$ then $\hat{e}_i^{k+1} = e_i^{\emptyset}$. We make the assumption that in this case, $(h_1, h_2, ..., h_k, \hat{h}_{k+1}) \in H$.

H implicitly defines a function $P : H \diagdown H_T \rightarrow 2^I$ where $P(h)$ denotes the set of players who choose messages at history $h \in H \diagdown H_T$ where the interpretation is that players in $P(h)$ move simultaneously after having observed history h. We also denote $\hat{P}(h)$ for the set of players in $P(h)$ who have not sent their piece of evidence yet.

In addition, we have a function $g : H_T \rightarrow A$ specifying the outcome selected by the mechanism for each terminal node. We now define two notions of Nash equilibrium each corresponding to one interpretation, i.e. feasibility or cost interpretation. Now for each player i, let $H_i = \{h \in H : i \in P(h)\}$ be the set of histories where player i plays. Define player i's strategy as a mapping $\sigma_i : H_i \rightarrow \left(E_i \cup \{e_i^{\emptyset}\} \right) \times M_i$. For consistency, we assume that a strategy must satisfy

[37] There is a sense in which this assumption is made for consistency since we may have that evidences $\{x\}$ and $\{y\}$ are available at θ but $\{x, y\}$ is not. If a player could submit $\{x\}$ first and then submit $\{y\}$ at a subsequent point, it is effectively as though we have enlarged the set of evidence available to a player, which we do not wish to do. Indeed, if we assumed that players can submit evidence an arbitrary number of times in a dynamic mechanism, then this would artificially make the evidence structure normal and thereby change the problem fundamentally.

that for any $h \in H$ and $i \in P(h) \setminus \hat{P}(h) : \sigma_i(h) \in \{e_i^{\emptyset}\} \times M_i$. Each profile of strategy $\sigma = (\sigma_i)_{i \in I}$ generates a terminal history $h(\sigma) \in H_T$ where $h(\sigma) = (h_1(\sigma), ..., h_k(\sigma))$ for some integer k. Recall that an history $h' = (h'_1, ..., h'_{k'}) \in H$ is a truncation of history $h = (h_1, ..., h_k) \in H$ if $k' \leq k$ and $h' = (h_1, ..., h_{k'})$.

Definition 12. *Given a mechanism (M, H, g), a profile of strategy $\sigma = (\sigma_i)_{i \in I}$ is a NE under the cost interpretation at state θ if*

(1) $u_i(g(h(\sigma), \theta) \geq u_i(g(h(\sigma'_i, \sigma_{-i}), \theta))$ *for any strategy σ'_i s.t. for any $h \in H_i : \sigma'_i(h) \in \left(\left(E_i^{\theta} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$, and*
(2) *for any truncation h' of $h(\sigma)$, $\sigma_i(h') \in \left(\left(E_i^{\theta} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$ for each $i \in P(h')$.*

The set of NE under the cost interpretation at θ is denoted $NE\left(\left(M, H, g\right), \theta\right)$. Condition (1) of the above definition says that one cannot have profitable deviations, in particular to strategies where the evidence sent is not in E_i^{θ}. This is consistent with the interpretation that sending such a message is infinitely costly. Condition (2) imposes no restrictions on histories that are not truncations of $h(\sigma)$, i.e. on out-of-equilibrium histories. In particular, for such an history, h, it may be the case that a player i in $P(h)$ plays $\sigma_i(h) \notin \left(E_i^{\theta} \times M_i\right)$. Again, this definition is consistent with the interpretation that evidence $e_i \notin E_i^{\theta}$ is always available but is infinitely costly.

Let $O\left(\left(M, H, g\right), \theta\right) := \{a \in A : \exists \sigma \in NE\left(\left(M, H, g\right), \theta\right) \text{ s.t. } g\left(h(\sigma)\right) = a\}$ and say that a mechanism (M, H, g) implements a SCR f under the cost interpretation if for all $\theta : f(\theta) = O\left(\left(H, g, M\right), \theta\right)$. A SCR is implementable under the cost interpretation if there exists a mechanism that implements it under the cost interpretation.

Proposition 7. *If a SCR is implementable (in a dynamic mechanism) under the cost interpretation, it is evidence-monotonic.*

Proof. Assume f is implemented under the cost interpretation by a mechanism (M, H, g). Then for each θ and $a \in f(\theta)$, there exists $\sigma^* \in NE\left(\left(M, H, g\right), \theta\right)$ s.t. $g\left(h(\sigma^*)\right) = a$. By definition, for each player i, there is exactly one piece of evidence that is sent along $h(\sigma^*)$, call the profile of such evidences $e^*_{\theta,a}$ and note that by (2) in the definition of a NE under the cost interpretation, $e^*_{\theta,a} \in E^{\theta}$. Consider any θ' satisfying (*) and (**). Since σ^* is a NE at θ,

$$u_i\left(g\left(h(\sigma^*)\right), \theta\right) \geq u_i\left(g\left(h(\sigma'_i, \sigma^*_{-i})\right), \theta\right)$$

for all i, σ'_i s.t. for any $h \in H_i : \sigma'_i(h) \in \left(\left(E_i^{\theta} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$. By (*),

$$u_i\left(g\left(h(\sigma^*)\right), \theta'\right) \geq u_i\left(g\left(h(\sigma'_i, \sigma^*_{-i})\right), \theta'\right)$$

for all i, σ'_i s.t. for any $h \in H_i : \sigma'_i(h) \in \left(\left(E_i^{\theta} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$. By (**) $E^{\theta'} \subseteq E^{\theta}$, thus we have

$$u_i\left(g\left(h(\sigma^*)\right), \theta'\right) \geq u_i\left(g\left(h(\sigma'_i, \sigma^*_{-i})\right), \theta'\right)$$

for all i, σ'_i s.t. for any $h \in H_i : \sigma'_i(h) \in \left(\left(E_i^{\theta'} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$. In addition, (**) stipulates that $e^*_{\theta,a} \in E^{\theta'}$ and so for any truncation h' of $h(\sigma)$, $\sigma^*_i(h') \in \left(\left(E_i^{\theta'} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$ for each $i \in P(h')$.

Therefore, σ^* is a NE (under the cost interpretation) at θ' and since $g(h(\sigma^*)) = a$, we get that $a \in f(\theta')$. □

By Theorem 2, evidence-monotonicity is sufficient for implementation in the above sense whenever there are three or more players and no veto power and non-satiation are satisfied. Indeed we do not need to go beyond static games in this case. Thus we get exactly the same type of characterization as in the main text when dynamic mechanisms are allowed but the cost interpretation is adopted.

Let us turn now to the feasibility interpretation. Here, at each state θ, it is impossible for any player i to send any evidence not in E_i^θ, even off-the-equilibrium path.

Definition 13. *Given a mechanism (M, H, g), a profile of strategy $\sigma = (\sigma_i)_{i \in I}$ is a NE under the feasibility interpretation at state θ if*

(1) $u_i(g(h(\sigma), \theta) \geq u_i(g(h(\sigma_i', \sigma_{-i}), \theta)$ *for any strategy* σ_i' *s.t. for any* $h \in H_i : \sigma_i'(h) \in \left(\left(E_i^\theta \cup \{e_i^\emptyset\}\right) \times M_i\right)$*, and*
(2) *for any history* $h' \in H$*,* $\sigma_i(h') \in \left(\left(E_i^\theta \cup \{e_i^\emptyset\}\right) \times M_i\right)$ *for each* $i \in P(h')$*.*

The set of NE under the feasibility interpretation at θ is denoted $NE^{\mathbb{F}}((M, H, g), \theta)$. Define $O^{\mathbb{F}}((M, H, g), \theta) := \{a \in A : \exists \sigma \in NE^{\mathbb{F}}((M, H, g), \theta) \text{ s.t. } g(h(\sigma)) = a\}$, and say that a mechanism (H, g, M) implements a SCR f under the feasibility interpretation if for all $\theta : f(\theta) = O^{\mathbb{F}}((M, H, g), \theta)$. A SCR is implementable under the feasibility interpretation if there exists a mechanism that implements it under the feasibility interpretation.

As shown in Section 2.9 of the main text, if we use this notion then dynamic mechanisms may be useful to achieve implementation of a SCR that is not evidence-monotonic. However, we prove next that if the evidence structure is normal, then again our central analysis remains unchanged.

Proposition 8. *Assume the evidence structure is normal. If a SCR is implementable (in a dynamic mechanism) under the feasibility interpretation, it is evidence-monotonic.*

Proof. Recall first that in the case of normality, a SCR f is evidence-monotonic provided that for all θ and all $a \in f(\theta)$, for any θ' satisfying (*) and $E^\theta = E^{\theta'}$ we have $a \in f(\theta')$. Assume f is implemented under the feasibility interpretation by a mechanism (H, g, M). Pick θ and $a \in f(\theta)$, and an equilibrium $\sigma^* \in NE^{\mathbb{F}}((M, H, g), \theta)$ s.t. $g(h(\sigma^*)) = a$. Consider any θ' satisfying (*) and $E^\theta = E^{\theta'}$. Since σ^* is a NE at θ,

$$u_i(g(h(\sigma^*)), \theta) \geq u_i\left(g\left(h(\sigma_i', \sigma_{-i}^*)\right), \theta\right)$$

for all i, σ_i' s.t. for any $h \in H_i : \sigma_i'(h) \in \left(\left(E_i^\theta \cup \{e_i^\emptyset\}\right) \times M_i\right)$. By (*),

$$u_i\left(g(h(\sigma^*)), \theta'\right) \geq u_i\left(g\left(h(\sigma_i', \sigma_{-i}^*)\right), \theta'\right)$$

for all i, σ_i' s.t. for any $h \in H_i : \sigma_i'(h) \in \left(\left(E_i^\theta \cup \{e_i^\emptyset\}\right) \times M_i\right)$. Because $E^{\theta'} = E^\theta$, we have

$$u_i\left(g(h(\sigma^*)), \theta'\right) \geq u_i\left(g\left(h(\sigma_i', \sigma_{-i}^*)\right), \theta'\right)$$

for all i, σ_i' s.t. for any $h \in H_i : \sigma_i'(h) \in \left(\left(E_i^{\theta'} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$. In addition, $E^{\theta'} = E^{\theta}$ implies that for any history $h' \in H$, $\sigma_i^*(h') \in \left(\left(E_i^{\theta'} \cup \{e_i^{\emptyset}\}\right) \times M_i\right)$ for each $i \in P(h')$. Therefore, σ^* is a NE (under the feasibility interpretation) at θ' and since $g\left(h(\sigma^*)\right) = a$, we get that $a \in f\left(\theta'\right)$. $\square$

References

Abreu, D. and A. Sen (1990): "Subgame perfect implementation: A necessary and almost sufficient condition," *Journal of Economic Theory*, 50, 285–299. 1

——— (1991): "Virtual Implementation in Nash Equilibrium," *Econometrica*, 59, 997–1021. 1

Aghion, P., R. Holden, and D. Fudenberg (2007): "Subgame Perfect Implementation with Almost Perfect Information," Mimeo, University of Harvard. 1

Ben-Porath, E. and B. L. Lipman (2008): "Implementation and Partial Provability," Mimeo, Hebrew University and Boston University. 5, 10, 13, 14, 15, 29

Benoît, J.-P. and E. Ok (2008): "Nash Implementation Without No-veto Power," *Games and Economic Behavior*, 51–67. 10, 28

Bull, J. (2008): "Mechanism Design with Moderate Evidence Cost," *The B.E. Journal of Theoretical Economics (Contributions)*, 8. 5

Bull, J. and J. Watson (2004): "Evidence Disclosure and Verifiability," *Journal of Economic Theory*, 118, 1–31. 5

——— (2007): "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, 58, 75–93. 3, 5, 13, 15, 22, 23, 34

Chung, K.-S. and J. Ely (2003): "Implementation with Near-Complete Information," *Econometrica*, 71, 857–871. 1

Crawford, V. and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451. 5

Deneckere, R. and S. Severinov (2008): "Mechanism Design with Partial State Verifiability," *Games and Economic Behavior*, 64, 487–513. 5

Dutta, B. and A. Sen (1991): "A Necessary and Sufficient Condition for Two-Person Nash Implementation," *Econometrica*, 58, 121–128. 19

Forges, F. and F. Koessler (2005): "Communication Equilibria with Partially Verifiable Types," *Journal of Mathematical Economics*, 41, 793–811. 5

Glazer, J. and A. Rubinstein (2001): "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, 36, 158–173. 15

——— (2004): "On Optimal Rules of Persuasion," *Econometrica*, 72, 1715–1736. 5, 15

——— (2006): "A study in the pragmatics of persuasion: a game theoretical approach," *Theoretical Economics*, 1, 395–410. 5, 15

Goltsman, M. (2009): "Nash Implementation using Simple Mechanisms without Undesirable Mixed-Strategy Equilibria," Mimeo, University of Western Ontario. 29

Green, J. R. and J.-J. Laffont (1986): "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, 53, 447–56. 5, 13

Grossman, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law & Economics*, 24, 461–483. 13

Jackson, M. O. (1991): "Bayesian Implementation," *Econometrica*, 59, 461–77. 29

——— (1992): "Implementation in Undominated.Strategies: A Look at Bounded Mechanisms," *Review of Economic Studies*, 59, 757–75. 4, 29

Kartik, N. (2008): "Strategic Communication with Lying Costs," Forthcoming in *Review of Economic Studies*. 5

Kartik, N., M. Ottaviani, and F. Squintani (2007): "Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, 134, 93–116. 5

Kartik, N. and O. Tercieux (2008a): "Implementation and Evidence: Incomplete Information," Manuscript in preparation. 29

——— (2008b): "Mixed-Nash and Correlated Equilibrium Implementation," Manuscript in preparation. 4, 12

Kunimoto, T. and O. Tercieux (2009): "Implementation with Near-Complete Information: The Case of Subgame Perfection," Mimeo, Paris School of Economics. 1

Lipman, B. L. and D. J. Seppi (1995): "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66, 370–405. 5, 13, 15, 21

Maggi, G. and A. Rodriguez-Clare (1995): "Costly Distortion of Information in Agency Problems," *RAND Journal of Economics*, 26, 675–689. 5

Maskin, E. (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23–38. 1, 3, 4, 7, 9, 12, 18

Maskin, E. and T. Sjöström (2002): "Implementation Theory," in *Handbook of Social Choice and Welfare*, ed. by K. J. Arrow, A. K. Sen, and K. Suzumura, Amsterdam: North Holland, vol. 1, 237–288. 12

Matsushima, H. (1988): "A New Approach to the Implementation Problem," *Journal of Economic Theory*, 45, 128–144. 1

——— (2008): "Role of Honesty in Full Implementation," *Journal of Economic Theory*, 139, 353–359. 4, 28

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391. 5, 13

Moore, J. and R. Repullo (1988): "Subgame Perfect Implementation," *Econometrica*, 56, 1191–1220. 1, 3, 10

——— (1990): "Nash Implementation: A Full Characterization," *Econometrica*, 58, 1083–1099. 10, 19

Myerson, R. B. (1982): "Optimal coordination mechanisms in generalized principal-agent problems," *Journal of Mathematical Economics*, 10, 67–81. 5

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura (1990): "Strategic Information Revelation," *Review of Economic Studies*, 57, 25–47. 5

Palfrey, T. R. and S. Srivastava (1991): "Nash Implementation Using Undominated Strategies," *Econometrica*, 59, 479–501. 1

Postlewaite, A. and D. Wettstein (1989): "Feasible and Continuous Implementation," *Review of Economic Studies*, 56, 603–11. 2

Saijo, T. (1987): "On Constant Maskin Monotonic Social Choice Functions," *Journal of Economic Theory*, 42, 382–386. 1

Sanver, M. R. (2006): "Nash implementing non-monotonic social choice rules by awards," *Economic Theory*, 28, 453–460. 10, 28

Sher, I. (2008): "Persuasion and Limited Communication," Mimeo, University of Minnesota. 5, 15

Singh, N. and D. Wittman (2001): "Implementation with Partial Verification," *Review of Economic Design*, 6, 63–84. 5